



International Conference on
**Social Media
& Society**

#SMSociety 2026

Book of Abstracts

14th International Conference on Social Media & Society

University of Glasgow

13–15 July 2026



Scan for online conference schedule

Table of contents

Session 1A - Algorithms and AI (1).....	20
1A.1 Coded Biases and Demographic Stereotyping in Facebook’s Curation of AI-Generated Images.	20
Aanila Kishwar (Indiana University Bloomington).....	20
1A.2 Hybrid AI Visuals and Deepfakes as instruments of political deception in the pre-election period of Bangladesh.	25
Arafat Al Yeasin (University of Liverpool)	25
1A.3 AI Fortune-Teller: Procedural Rhetoric and the Algorithmic Construction of Trust in AI Divination. Erika	30
Ningxin Wang (The Chinese University of Hong Kong, Shenzhen)	30
1A.4 Algorithm Adoration, Synergy, Discord, and Apathy: Categorizing Human–Algorithm Relationships in Social Media.	35
Biyang Wu-Ouyang (Education University of Hong Kong).....	35
Session 1B - Misinformation and Disinformation	38
1B.1 Seeing Is No Longer Believing: Audience Trust, Scepticism, and AI-Generated Visuals on Social Media.	38
Chisom Omeokachie (Independent Researcher).....	38
1B.2 Infodemics and Routine Vaccine Hesitancy in Low- and Middle-Income Countries: A Systematic Review of Social Media and Communication Environments.	43
Isabela Braga de Moura, Claudia Araujo (COPPEAD Graduate School of Business), Marina Martins Siqueira (Hospital Israelita Albert Einstein)	43
1B.3 “The solution to your pains!”: How scam ads exploit health disinformation on Meta platforms in Brazil.....	48
Nicole Sanchotene, Débora Salles (Federal University of Rio de Janeiro), Daphne Silva (Federal University of Alagoas), João Haddad, Marie Santini (Federal University of Rio de Janeiro)	48
1B.4 Combating Climate Misinformation via AI chatbots: How Interactivity and Lateral Reading Support Shape Trust, Perceived Response Credibility, and Further Verification Intentions.	54
Shirley S. Ho (Nanyang Technological University), Junru Huang (CNRS@CREATE), Chang He (Nanyang Technological University).....	54
Session 1C - Social Media Cultures and Everyday Life (1).....	61
1C.1 “Meaningful” or “meaningless” social interactions? The automation of commenting in live streams on Douyin.	61
Andreas Schellewald (Goldsmiths, University of London), Hui Lin (King's College London)	61
1C.2 Mediating ‘my London’ on social media: Knowledge migrants’ urban place experiences and perceptions through digital place memories.	66
Chenfei LI (King's College London)	66
1C.3 From Feed to Floor: AI-Powered Social Media as a Sensory Buffer in Neurodivergent Retail Experiences.	70

1C.4 Mainstreaming the Red Pill: The Popularization of Red Pill Ideology on TikTok, Instagram, and YouTube.....	76
Duygu Karatas (University of Siegen)	76
Session 1D - Discourse and Public Opinion (1)	81
1D.1 Exploring the temporal constructions of Facebook/Meta: presentist speed, progress, and exponentiality.....	81
Asher Kessler (London School of Economics)	81
1D.2 Children in Political Imagery: A Visual Values Analysis of Turkish Leaders' Instagram Posts.	85
Ayşenur Benevento (Bahcesehir University)	85
1D.3 Mongolian Stand-Up Comedy on YouTube as Public Discourse.	88
Borchuluun Yadamsuren (LUC), Undrah Baasanjav (Southern Illinois University), Iderjargal Dashdondog (National University of Mongolia).....	88
1D.4 A Deficit of Shared Reality? Youth, Fragmented Information Environments, Social Media, and Political Participation.	93
Catarina Feio, Lídia Oliveira (Universidade de Aveiro).....	93
Session 1E - WORKSHOP: Meta Content Library as a Research Tool	99
Yair Rubinstein (Research Partnerships - Meta), Evgenii Strepetov (Software Engineer - Meta) ..	99
Session 2A - Algorithms and AI (2).....	100
2A.1 Algorithmic Femininity on Instagram: A Case Study of the VI @fit_aitana.	100
Anna Pujadas-Gómez, Clara Virós-Martín (Universitat Pompeu Fabra).	100
2A.2 Algorithms for Academia: Mapping Users' Imaginaries of Academic Social Networks.	105
Guillaume Latzko-Toth (Laval University), Clémentine Fruchard (Université du Québec à Montréal), Elsa Fortant (INRS), Claudine Bonneau, Florence Millerand (Université du Québec à Montréal).....	105
2A.3 When Precarity Goes Viral: Entrepreneurial Ideology, Algorithmic Visibility and the Cross-Platform Reframing of Formal Employment.	111
Jessica Souza, Rosana Pinheiro-Machado (University College Dublin).....	111
2A.4 How do Terms of Service Influence Social Media Users Dynamics - A Privacy Anxiety Perspective.....	115
Jingyuan Liu (Boston University).	115
Session 2B - Health and Wellbeing (1)	119
2B.1 "I'm Crushed, Ashamed, and Shattered": Moral Sentiment, Network Position, and Recovery Communities on Reddit.	119
Kimberly Glasgow (Johns Hopkins University Applied Physics Laboratory).	119
2B.2 Information Wellbeing: Balancing News Knowledge and Mental Health.....	125
Inhwa Song, Manoel Horta Ribeiro, Marianne Aubin Le Quere (Princeton).	125
2B.3 Navigating Health Information on YouTube: Trust, Misinformation, and Health Decision Making During COVID-19 in India.	132

Nisha Rani (Manipal Institute of Communication), Usha Manchanda Rodrigues (School of Information and Communication Studies), Padma Rani (Manipal Institute of Communication)	132
2B.4 Beyond screen time: Differentiating specific digital activities as risks for adolescent mental health problems.	137
Peiyao Tang, Katarzyna Kostyrka-Allchorne (Queen Mary University of London), Aja Murray (University of Edinburgh), Mariya Stoilova (LSE), Marianne Etherson (University of Glasgow), Eliz Azeri (King's College London), Ellen Townsend, Chris Hollis (University of Nottingham), Sonia Livingstone (LSE), Edmund Sonuga-Barke (King's College London)	137
Session 2C - Social Media Cultures and Everyday Life (2)	143
2C.1 Fentanyl Tourism on YouTube: Serial Commodification of Urban Marginality.	143
Elsa Soro (European Business School), Lluís Garay Tamajón (Universitat Oberta de Catalunya [UOC])	143
2C.2 Constituting Affective Reality: Platformed Nostalgia and Digital Materiality in the Chinese Hanfu Revival Movement	150
Erika Ningxin Wang (The Chinese University of Hong Kong, Shenzhen)	150
2C.3 Out for the Algorithm: Queer Visibility and Self-Presentation on TikTok	155
Irina Bănoiu (Institute for Advanced Studies Vienna [IHS])	155
2C.4 Enhancing Micro-phenomenology with Qualitative, Semi-structured, Elicitation Interviews When Combined with Interpretative Phenomenological Analysis.	160
Joan Owen, Jaigris Hodson (Royal Roads University)	160
Session 2D - Discourse and Public Opinion (2)	165
2D.1 Networked Hatred: Co-evolution of Dehumanising Discourse and Channel Structure of Russian and Ukrainian Telegram During the 2022 Invasion	165
Elizaveta Chernenko (Oxford Internet Institute), Balazs Vedres (Department of Network and Data Science, Central European University)	165
2D.2 Speaking Back Between Misogyny and Censorship: Three Counterpublic Strategies by Chinese Women	170
Jiahao Sun (Hitotsubashi University)	170
2D.3 The Duality of Media Actors and Culture: The White Afrikaner Refugee Controversy on YouTube	174
Jihye Kim (Umass Amherst)	174
2D.4 How Topics Flow Across Reddit Communities Through Shared User Participation: Comparative Study of Pre and Post COVID-19 Vaccine Discourse	179
Joohee Kim (University of Massachusetts)	179
Session 3A - Algorithms and AI (3)	184
3A.1 Privacy Concerns and Self-Disclosure Motivation to AI: A Multidimensional Perspective	184
Kunqiang Wu, Wanting Huang (School of Communication, Hong Kong Baptist University)	184
3A.2 What Else Can We Do? Pessimistic Proactivity and Coerced Domestication in Chinese Gay Men's Negotiations with Algorithms	192
Longxuan Zhao, Yuyang Cao (Tongji University)	192

3A.3 “Cutting Ties” with AI? Perceptions and Management of AI Shaming among Young Art Creators in China	197
Yuyang Cao, Yixiang Zhang, Longxuan Zhao (Tongji University)	197
Session 3B - Digital Intimacies	202
3B.1 “Amazing body. Does the face match:” A Qualitative Study of how Queer Embodied Masculinities on Dating Apps Practice, Communicate and Build Trust with Anonymity	202
John Hodson (University of Salford)	202
3B.2 Narrative Singularity and Collective Resonance in TikTok Sexual Violence Disclosures and Networked Publics	207
Talia Fiester (University of Pennsylvania)	207
3B.3 Intimacy capital, recognition, and ordinary creators on TikTok	213
Irida Ntalla (University of the Arts London)	213
Session 3C - Social Media Cultures and Everyday Life (3)	218
3C.1 Status in the Palm of the Hand: Class-Based Patterns of Digital Leisure	218
Kata Számel (ELTE Centre for Social Sciences)	218
3C.2 Impatient feminism: Platform-mediated intimacy advice on Chinese social media	222
Jiayi Chen (University of Warwick)	222
3C.3 “My cat asks you to buy a drone”: How civilians in Ukraine strategically navigate social media logic to run small grassroots military fundraising campaigns	228
Kateryna Bystrytska (Rutgers University)	228
Session 3D - Discourse and Public Opinion (3)	233
3D.1 “Chat, Is This Real?”: Truth-Seeking Heuristics & Distributed Trust on Reddit's r/isthisAI .	233
Kaitlyn McNamee (Marist University)	233
3D.2 From ‘Diaploki’ to Pluralism - Does Social Media Still Drive Change in Greece’s Political and Media Spheres?	237
Michael Nevradakis (College Year in Athens)	237
3D.3 A Multimodal Discourse Analysis of the Iranian #MeToo movement on the Islamic Republic Media	241
Pardis Yarahmadi (QUT)	241
Session 3E - PANEL: Meta Content Library Community Roundtable: Where to for Social Media Research?	248
Session 4A - Platform Governance and Regulation (1)	250
4A.1 Decentered Platform Governance: Learning trust and Safety Practices with Communities on Decentralised Social Media Platforms	250
Ashwin Nagappa, Daniel Angus, Kateryna Kasianenko (Queensland University of Technology) 250	
4A.2 Trust Café: A Marxist-feminist approach to digital commons and content moderation	256
Dayei Oh (University of Strathclyde), Craig Ryder (Ludwig Maximilian University of Munich)	256
4A.3 Social media ‘back in the hands of the people’? Nuancing assumptions of legitimate content moderation on Mastodon	262

Klara Matusiewicz, João C. Magalhães (University of Manchester)	262
Session 4B - Digital and Data Methods.....	268
4B.1 Stance-taking as a technology of power and the self: A linguistic anthropological approach for analyzing social media interaction	268
Julian Canjura (The Pennsylvania State University)	268
4B.2 Gone But Not Missed? Understanding the Drivers and Implications of Post Deletions on Bluesky	273
Yining Wang, Johannes Gruber, Yannik Peters, Kartin Weller (GESIS - Leibniz Institute for the Social Sciences)	273
Session 4C - Social Media Culture and Everyday Life (4).....	278
4C.1 Navigating Social Media in the Remote High Mountains: Aging, Literacy, and Technological Obsolescence in Nubri Valley, Nepal	278
Kunsang Choden (University of Washington).....	278
4C.2 Distinguishing suicide attempts from suicidal ideation in young people: The role of addictive digital behaviours and negative life events	282
Marianne Etherson (University of Glasgow).....	282
4C.3 Who gets to speak the truth? Russian invasion of Ukraine and construction of authenticity in social media communication	287
Marie Heřmanová (University College London).....	287
Session 4D - PANEL: When the trolls find our research: A workshop in online harassment of scholars and how to move forward	292
Anatoliy Gruzdt (Ted Rogers School of Management at Toronto Metropolitan University), Esteban Morales (University of Groningen), Yimin Chen (RMIT University), Jaigris Hodson (Royal Roads University)	292
Session 4E - PANEL: Social Media in Flux: Embracing uncertainty and experimentation in research methods	295
Anna Feigenbaum (University of Glasgow), Ozlem Demirkol Tonnesen (University of Bergen) Doug Specht (University of Westminster) Sweta Baniya (Virginia Tech) Taylor Annabell (Cardiff University)	295
Session 5A - Platform Governance and Regulation (2).....	297
5A.1 Call the (Bot-)Police: User-Led Platform Governance of “(In)Authenticity” on Instagram ..	297
Nathalie Schäfer (Bauhaus-University Weimar)	297
5A.2 The Epistemic Burden of AI Labels: Interface Governance and User Responsibility on Social Media.....	302
Piotr Sochoń (Shanghai Jiaotong University), Nadia Urban (East China Normal University).....	302
5A.3 Governing Social Media Harms: A Critical Realist Extension of the Theory of The Firm	304
Robert Pavlovich, Bertrand Guillotin (Fox School of Business, Temple University)	304
Session 5B - Use and Users.....	312
5B.1 To understand and to be understood: Young people’s perceptions of peer support for mental health in social media.....	312

Essi Holopainen, Meri Kulmala, Anu Katainen, Anna-Maija Multas (University of Helsinki)	312
5B.2 The AI Citizenship Framework: Developing interactive AI literacies for critical and proactive citizens	318
Elinor Carmi (City, University of London), Anna Feigenbaum (Glasgow University), Photini Vrikki (UCL)	318
5B.3 The Dynamics of Use, Motivations, Mindsets on Well-Being for Young Adults on TikTok ...	322
Kate Mays, Ellen Kozelka (University of Vermont)	322
Session 5C - Discourse and Public Opinion (4)	330
5C.1 Hot Nerd, Femininity and Gender Relations in Post-socialist China: A Critical Discourse Analysis of Sapiosexuality on RedNote	330
Qianzao Yang (University of Cambridge)	330
5C.2 Not-So-Happy Ending: How failed series finales come to be and the reaction from online fan bases	336
Tess Arnold (Toronto Metropolitan University)	336
5C.3 Sexual violence (in)justice online: Exploring the influence of negative discourses on victim-survivor justice seeking	339
Ursula Shepherd (University of Glasgow)	339
Session 5D - PANEL: Towards a Future Research Agenda: TikTok Creators and Digital Economies	343
Zoetanya Sujon (University of the Arts London), Taylor Annabell (Cardiff University), Marie Heřmanová (University College London), Irida Ntalla (London College of Communication)	343
Session 5E - PANEL: From Red Dresses to Memory Stones: Multimedia Activism and Gender-Based Violence in Canada	348
Nicolette Little (University of Alberta), Esteban Morales (University of Groningen), Jaigris Hodson (Royal Roads University), Jessica Ringrose (UCL, Institute of Education)	348
Session 6A - Digital Education	352
6A.1 Where do UK universities represent themselves online? The institutional social media landscape post-Twitter	352
Katy Jordan (Lancaster University)	352
6A.2 Student social media etiquette: Negotiating teacher authority in a Chinese university	360
Linan Tang (King's College London)	360
6A.3 Bluesky as a Teaching and Learning Tool for International Students in Hybrid Communication Courses	364
Wahiba Chair, Elva Fisher (University Canada West)	364
6A.4 AI, Social Media, and Higher Education: A Comparative Policy Analysis across UK, India, and South Africa	371
Devina Sarwatay, Anubha Sarkar (City St George's, University of London), Swarantra A (Indian Institute of Management Indore), Kealeboga Aiseng (University of the Witwatersrand)	371
Session 6B - Promotional Social Media, Branding, & Marketing	377
6B.1 Children in the Creator Economy of Family Influencers: Visibility, Engagement, and Regulatory Gaps in Italy	377

Elisabetta Locatelli (Università Cattolica del Sacro Cuore)	377
6B.2 Performing ‘Mountains and Spices’ on Dianping: Cultural Politics and Platform Marketing of “Yunguichuan Bistros” in Urban China	383
Hao Xie (University of Warwick).....	383
6B.3 Virtual vs. Human Faces: A comparative study of influencer opinions and their effects on user perceptions and pro-environmental behavioral intentions	388
Son Tung Duong (Deakin University), Asma Husna, Nhung Nguyen (University of South Alabama)	388
6B.4 Enterprising Platform-Dependent Creators and the Rise of Creator-Dependent Labour in India	392
Srikanth Nayaka (Indian Institute of Management [IIM] Indore).....	392
Session 6C - Cyberbullying, Trolling and Antisocial Behavior (1)	397
6C.1 Using YouTube Data to Measure Online Harassment in Canadian Politics	397
Inessa De Angelis (University of Toronto)	397
6C.2 Differentiating Contact and Fantasy Online Grooming Offenders	403
Jake Street, Isibor Ihianle, Ahmad Lotfi (Nottingham Trent University).....	403
6C.3 Online Racism, Digital Mental Health Tools, and Online Mental Health Communication: A Cross Sectional Study Among Black Young Adults With and Without Depression or Anxiety	407
Melissa Holland, Kyle Walker, Oreoluwa Badejoh, Vanessa Volpe (North Carolina State University), Brian TaeKyuk Keum (University of California, Berkeley), Jimi Huh, Hans Oh (University of Southern California).....	407
6C.4 Social Media Use, Online Racism, and Mental Health among Black Emerging Adults	413
Oreoluwa Badejoh (North Carolina State University), Shedrick Garrett (University of North Carolina at Chapel Hill), Melissa Holland, Vanessa Volpe (North Carolina State University), Brian TaeHyuk Keum (University of California, Berkeley), Jimi Huh, Hans Oh (University of Southern California)	413
Session 6D - Social Media Cultures and Everyday Life (5)	418
6D.1 From Forums to Feeds: The Platformized Transformation of Chinese K-pop Fandom.....	418
Shangru Jiang (University of Bristol).....	418
6D.2 Chords of Small Media Through Time	422
Sahar Sagha (Durham University)	422
6D.3 Masculinity and Social Change: A dramaturgical explanation of online misogyny	426
Shona Rooney (Queen's University Belfast)	426
Session 6E - Affordances	431
6E.1 “What Should I Wear?” Facebook Groups as Sociotechnical Infrastructures for Embodied Rhetorical Agency.....	431
Najia Nazir (Texas Tech University)	431
6E.2 The logistics of visibility: Reconfiguring illicit drug trade on Telegram	435
Piotr Siuda (Kazimierz Wielki University in Bydgoszcz)	435

6E.3 Palestinian women's social media practices in London: platformed strategies of resistance, solidarity and political expression	441
Sevilay Cesurer (UAL)	441
6E.4 Undoing Dravidian Feminism: How Tamil Women Influencers Are Re-Traditionalizing Gender Through Digital Culture	445
Sinthu Bhairavi Govindasamy Vinoba (Oktopost Technologies) Neelambari Govindasamy Vinoba (Community College, Central University of Tamil Nadu)	445
Session 6F - Power, Platforms, and Politics	451
6F.1 MAGA on TikTok: Issue Mapping a Political Network	451
Gil Sharon (University of Leeds)	451
6F.2 From Encyclopedic Authority to Platformed Knowledge: Wikipedia, Social Media, and the Reconfiguration of Epistemic Power	455
Giota Alevizou (King's College London)	455
6F.3 Platform Nonmarket Strategies as Power Projection	458
Gordon Dimitrieff (Antwerp Management School)	458
6F.4 The multimodal use of Instagram in Italian politics: How do candidates use verbal and non-verbal strategies during regional elections?	463
Umberto Famulari (Colorado State University)	463
Session 7A - Platform Epistemologies and Imaginaries	471
7A.1 How is the internet imagined?: An analysis of drawings of the internet, 2014-2024	471
Nicholas John, Thomas Trillo (The University of Manchester)	471
7A.2 ZOOMWORK: Work imaginaries in platform culture. Content creators and Gen Z	476
Camilla Volpe, Luca Giuffè, Guilia Giorgi (University of Milan)	476
7A.3 Authenticity Effects of Synthetic Images: Interactions among Political Cues, Context, and Platform Environments	480
Yu-Chung Cheng (National Chengchi University)	480
7A.4 Hyper-datafication: How Platforms Make Everything Feel Like Social Media	485
Aparajita Bhandari (University of Waterloo), Sara Bimo (York University), Chelsea Butkowski (American University)	485
Session 7B - Health and Wellbeing (2)	489
7B.1 Navigating Mental Health Online: How Platform Affordances Shape Emotional Connection on TikTok and Threads	489
Adrienne Hall-Phillips, Khushali Shah (Worcester Polytechnic Institute)	489
7B.2 Designing for Attention vs. Designing for Wellbeing: The Impact of a Screentime Nudge on TikTok Flow	493
Amber Priem, Malaika Brengman, Nanouk Verhulst (Vrije Universiteit Brussel)	493
7B.3 Predictive Approaches to Suicidal Thoughts in Spanish Youth: The Role of Online Media Engagement Patterns	498
Daniel Barredo Ibáñez (Universidad de Sevilla), Lucía Caro Castaño (Universidad de Cádiz) ...	498

7B.4 The Many Faces of Stigma: Profiles of Adolescent Mental Illness Stigma Beliefs and Social Media Use Across Switzerland, Germany, and Austria	507
Dora Weubel, Sophie Mayen (USI Università della Svizzera Italiana), Kathrin Karsay (University of Vienna), Ruth Wendt (Ludwig Maximilian University), Anne-Linda Camerini (USI Università della Svizzera Italiana)	
Session 7C - Infrastructure, Platformisation and Datafication	512
7C.1 Platform culture and Traditional Performance: Algorithmic Mediation of Ghoomar Dance on Instagram	512
Bhavani Singh (Madan Mohan Malviya University of Technology), Rashmi Jha (Banaras Hindu University)	
7C.2 Platform culture and Traditional Performance: Algorithmic Mediation of Ghoomar Dance on Instagram	517
Jessica Witte, Ari Stillman (University of Edinburgh)	
7C.3 The Platform Turn in Online Games	521
Aditya Deshbandhu (University of Exeter), Devina Sarwatay (City St George's, University of London), Aparna Vincent (Indian Institute of Management - Indore)	
7C.4 Platformed Placemaking: Commemorative Stickers and the Production of Networked Places.....	525
Sarit Navon, Noam Gal (Open University), Ido Ramati (Hebrew University)	
Session 7D - Social Media Cultures and Everyday Life (6)	529
7D.1 Contemporary mediated civic becoming: Filipino American cultivation of civic agency through social media	529
Stephanie Thelwall (London School of Economics)	
7D.2 From street to screen: The visibility labor of Italian organized crime on TikTok	533
Vincenzo Luise (Università degli studi di Napoli Federico II)	
7D.3 New Cultural and Religious Practices: The Use of Facebook by Azorean Religious Communities	537
Pedro Lucas (Universidade Nova de Lisboa)	
7D.4 "I Don't Know Many Words, But I'm Smart": Platform Literacies and Older Content Creators on Douyin.....	540
Yingwen Wang, Zoetanya Sujon (University of the Arts, London)	
Session 7E - Online and Offline Communities	544
7E.1 Information Practices in Fragmented Online Communities: Taylor Swift Fans and Anti-Fans on Reddit	544
Samantha Vilkins, Katherine M. FitzGerald, Sebastian Svegaard, Axel Bruns (Queensland University of Technology).....	
7E.2 From Mate Selection to Testosterone Talk: Mapping Science Claims in the (Neo-) Manosphere.....	550
Catherine Baker (Dublin City University)	

7E.3 "Cuckoo Nut Lady C*nt Face McGee": Gendered Constructions of Mothers Who Kill in True Crime Content	555
Avery Fraser, Ella Duncan, Kathleen Rodgers (University of the Fraser Valley)	555
S7E.4 'Tapping into Your Feminine Energy' Down the Alt-Right Pipeline: Comparing Narratives and Community Bonds in the Comment Sections of Tradwife and Divine Feminine Content Creators on TikTok.....	558
Monica Vania Chavez, Anna Feigenbaum, Rinlapas Ketverapong (University of Glasgow)	558
Session 7F - Knowledge, Value, and Power in Platform Economies	563
7F.1 The Social Life of the MSI Metric: Rationalities and Metrologies around "Meaningful Social Interactions" at Meta	563
Matías Valderrama Barragán (Adolfo Ibáñez University)	563
7F.2 "My CPU Is Heating Up": Relational Labor and the Making of AI Companionship in the Creator Economy	569
Yena Kang (University of Massachusetts Amherst)	569
7F.3 AI and the Assault on Epistemic Structures in Russian Information Operations	573
Walker Winslow-Stephenson (University of North Carolina at Chapel Hill)	573
Session 8A - Social Media Cultures and Everyday Life (7)	578
8A.1 Authenticity Beyond Strategy: Granfluencers and the Reimagining of Ageing in Digital Culture	578
Rinlapas Ketverapong (University of Glasgow)	578
8A.2 Bartering networks: An ethnographic study on how Indian fashion micro-influencers engage in non-monetary exchanges.....	584
Neha Murthy (International Institute of Information Technology - Hyderabad), Brishti Sen Banerjee (Indian Institute of Technology, Kanpur), Rajorshi Ray (International Institute of Information Technology - Hyderabad)	584
8A.3 Reimagining the Metrosexual Moment for the 21st Century: The Plight of the Performative Male.....	588
Riley Wilson (Toronto Metropolitan University)	588
8A.4 Discourses of Anti-Trans Disinformation on the Joe Rogan Experience Podcast: A Corpus Assisted Critical Discourse Analysis	592
Jaigris Hodson (Royal Roads University)	592
Session 8B - Networked Ideologies	597
8B.1 Platformed wifetypes: Comparative global #tradwife recipes for feminine social media success	597
Zoe Hurley, Sreya Mitra (American University of Sharjah)	597
8B.2 From Post-Truth to Post-Objectivity: A Cross-Cultural Analysis of Newsfluencer Values ..	602
Tommaso Trillò, CJ Reynolds, Pyung Hwa Park (The Hebrew University of Jerusalem)	602
8B.3 Who watches whom: How YouTube channel networks organize migration discourse	607
William Hollingshead, Anatoliy Gruzd, Philip Mai (Toronto Metropolitan University)	607

8B.4 Down the Content Rabbit Hole: Making Algorithmic Content Shift on Short-Form Video Platforms Measurable	614
Tobias Kietreiber, Felix Krejca, Anna Beubacher, Sebastian Neumaier, Lukas Daniel Klausner (University of Applied Sciences St Pölten)	614
Session 8C - Politics, Policy and Regulation (1)	621
8C.1 Aesthetic Authoritarianism? Visual Propaganda and the Politics of Urban Renewal in Ethiopia	621
Amanuel Tesfaye Kebede (Helsinki University)	621
8C.2 "Flooding the Tropical Zone": Understanding Algorithmic Amplification on Instagram During Election Campaigns	625
Bruna Paroni (University of Urbino)	625
8C.3 Sowing Discord: Exploring the use of online chat platforms in mobilising worker resistance	632
Daniel Jordan (University of Glasgow)	632
8C.4 "Don't hurt me, no more": What can we learn from the pharmaceutical industry regulation about platform and AI regulation	637
Elinor Carmi (City, University of London)	637
Session 8D - Platformisation	641
8D.1 Creators, journalists, and relational precarity in India's beauty web	641
Anuja Premika (Mahindra University)	641
8D.2 Algorithmic Visibility as a Capital Conversion Mechanism: The Practice Logic of Content Creators in the Digital Public Sphere	646
Ben Xu (University of Leeds)	646
8D.3 Platform-Mediated Labour and Professional Precarity: Producing Preconditions for Union Organizing in Journalism	650
Errol Salamon (University of Stirling)	650
8D.4 Memes as Micro-Archives: Documenting Everyday Life in Digital Culture	654
Jyoti (IIT Jodhpur)	654
Session 8E - Social Media and Platform Infrastructure	657
8E.1 Misogynoir, Influencer Tactics, and Profit Maximization on YouTube	657
Areyana Proctor (University of Wisconsin - Madison)	657
8E.2 "TikTok... Are You Okay?!" Understanding the U.S. TikTok Ban and Forced Sale through Infrastructural Inversion	662
Carlos Entrena Serrano (Vrije Universiteit Brussel), Thomas A. Wright (University of Sheffield)	662
8E.3 The Shifting Frontier of Data Access: Research, Restriction, and Rigmarole in the Meta Content Library	668
Laura Vodden, Axel Bruns (Digital Media Research Centre, QUT)	668
8E.4 Learning the Algorithm: Algorithmic Grooming and Governance in Social Media Cultures	673
Nadia Urban (East China Normal University)	673

Session 8F - PANEL: Performing Platforms: Bodily Performances and Hashtag Cultures in/beyond Performance Studies	677
Miya Shaffer (York University, Toronto), Archer Porter (University of Texas), Ali Na, Eric Chalfant (Queens University)	677
Session 9A - Platform Governance and Regulation (3)	683
9A.1 The Day-Job Paradox: AI Vulnerabilities of Independent Musicians	683
Travis Lloyd (Cornell Tech)	683
9A.2 Governance over hateful ideology: A longitudinal analysis of TikTok's hate speech governance in 2018 - 2025	688
Yang Xu (University of Helsinki)	688
9A.3 Who Should Be in Charge? Safety and Agency in the Public Imagination of EU Platform Governance	692
Claire Stravato Emes (SciencesPo)	692
9A.4 "It is not about vaccines, it is about THEM!": Othering and Discursive Violence in Anti-Vaccine Telegram Groups	696
Raquel Recuero (Universidade Federal de Pelotas)	696
Session 9B - Uses & Users (2)	700
9B.1 Paths towards the far-right: the role of social media in young people's political socialization	700
Laia Corxet Solé (Scuola Normale Superiore)	700
9B.2 Feeling climate: how emotional responses to data visualisation shape political engagement	705
Monika Fratzczak (Manchester Metropolitan University)	705
9B.3 Playing with Fate: Play Framing, Context Collapse, and Personal Data Disclosure in AI-Mediated Bazi Divination	710
Hui Lin (King's College London), Yingwen Wang (University of The Arts London)	710
9B.4 "you are treshhhhh player" - Toxicity in Twitch Livestreams	714
Tamara Bodden (Universität Koblenz), Maite Taboada (Simon Fraser University)	714
Session 9C - Politics, Policy, & Regulation (2)	720
9C.1 Do Electoral Stages Shape Digital Campaigning? Meta Strategies in Uruguay's 2024 Elections	720
Juan Pérez Martirené, Rafael Piñeiro Rodríguez (Universidad Católica del Uruguay), Fernando Rosenblatt (University of Manchester)	720
9C.2 Enabling informed voter decisions via parsimonious factually falsifiable social media design	725
Patrick McHugh, Tommy Barton, Sami Peretz (Brown University)	725
9C.3 A wholesome insult, an uplifting threat: Making sense of toxicity amid 'good vibes' digital cultures	729
Esteban Morales (University of Groningen)	729

9C.4 “The culture is CHANGING!!!”: #MAHA, RFK, and the TikTok health influencer	735
Hannah Artman (University of Missouri), Rutendo Chimbaru (University of Miami), Bonnie Deal (Phoenix College)	735
Session 9D - Digital Selfhood and Social Experience	743
9D.1 Navigating Agency and Constraint: Racialised Emotions and Resistant Subjectivities in TikTok Anti-Asian Activism	743
Xin Zhao, Yunfei Li, Xiao Ma (Bournemouth University)	743
9D.2 Who owns #RealEnglish? Influencer Teachers’ Construction of Authenticity on Instagram	747
Ali Fuad Selvi (University of Alabama), Erhan Aslan (University of Reading)	747
9D.3 Seeing what your friends are up to: FOMO through group chats	751
Carolin Lehmann, Sonja Utz (Leibniz-Institut für Wissensmedien (IWM) Tübingen)	751
9D.4 For a Theory of Onlife Practices: Studying Media with Bourdieu to Interpret Practices in the Interaction between Autochthonous Contexts and Digital Platforms	757
Roberto Graziano (Università degli studi di Napoli Federico II)	757
Session 9E - Platform Design and Politics	762
9E.1 Decoding algorithmic suppression: How digital activists uncover and navigate platform censorship in Finland?	762
Sanna Malinen (Tampere University)	762
9E.2 Immigration-Related Community Notes on X: Topic Structure and Evaluation Dynamics ..	767
Junyuan Lin, Anatoliy Gruzd (Toronto Metropolitan University)	767
9E.3 Triadic polarisation in climate information sharing: Insights from Twitter during COP26 and COP27	772
Ümit Bedretdin (University of Helsinki), Dayei Oh (University of Strathclyde)	772
9E.4 X marks the end: Wartime digital activism, community-led moderation and decentralisation in the “ruins” of social media platforms	778
Kateryna Kasianenko (Queensland University of Technology), Ariadna Matamoros-Fernández (University College Dublin), Olga Boichak (The University of Sydney), Ashwin Nagappa (Queensland University of Technology)	778
Session 9F - PANEL: Conditional Belonging, Regulated Visibility: Women’s Platform Participation in China	782
Yue Qin (London College of Communication), Yijia (Yika) Gu (Deakin University), Xintong Jia (University of Essex), Ruoqian Li (University of Sheffield), Nan Ding (London School of Economics)	782
Session 10A - PANEL: Cloud-C: Studying climate communication on social media through a "many researchers, one dataset" approach	788
Dayei Oh (University of Strathclyde), Axel Bruns (Digital Media Research Centre at Queensland University of Technology), Barbara Pfetsch (Freie Universität Berlin), Risto Kunelius, Anton Berg (University of Helsinki)	788
Session 10B - PANEL: Algorithmic Childhood, Intimacies, and Literacies: Networked & synthetic selves and relationships	793

Jessica Ringrose, Chiara Fehr (UCL Institute of Education), Ayşe Aslı Bozdağ (Bahçeşehir University), Devina Sarwatay (City St George's, University of London), Sandra El Gemayel (Digital Futures for Children Centre, LSE)	793
Session 10C - PANEL: Algorithmic Sociality: Platforms, Connection, and the Reconfiguration of Digital Relations	800
Ellenrose Firth (Sapienza University of Rome), Andreas Schellewald (Goldsmiths University London), Hui Lin (King's College London), Yingwen Wang, Wenxin Guo (London College of Communication)	800
Session 10D - PANEL: From Creative Class to Popular Class: TikTokization, AI, and the Standardization of Art, Media, and Platform Culture	806
Adam Arvidsson, Roberto Graziano (University of Naples Federico II), Maitrayee Deka (University of Essex), Sabrina Bellafronte (University of Naples Federico II), Alessandro Gandini (University of Milan), Guo Jianpeng, Vincenzo Laezza (University of Naples Federico II)	806
Session 11A - Workshop: Would a Poem by Any Other Platform Sound As Sweet?: Examining Social Media Metrics through Sentiment Analysis	809
JuEunhae Knox (University of Stirling)	809
Session 11B - Workshop: Collaging the Datafied Self: Mapping Postdigital Subjectivity through Creative Methods	810
Liv Owens (City St George's, London), Chiara Fehr (UCL Institute of Education)	810
Session 11C - Workshop: Resurrecting platform pasts: A workshop for research in a changing platform landscape	811
Frances Corry (University of Pittsburgh), Chelsea Butkowski (American University's School of Communication)	811
Session 11D - Workshop: Beyond the Hairball: An Applied Approach to Social Network Analysis with Communalystic	812
Anatoliy Gruzd (Ted Rogers School of Management), Philip Mai (Toronto Metropolitan University)	812
Session 12 – Poster Presentations	813
12.1 Dynamics of Perceived Control in Finnish Young Adults' Social Media Use	813
Pinja Akkanen (University of Turku)	813
12.2 Perceived Influence of Health Misinformation on Social Media: A Qualitative Study through the Lens of Third-Person Effect Theory	817
Wed Alfotawi (Bournemouth University)	817
12.3 Sensation seeking, online risks and digital skills: Untangling the pathways in adolescence.	824
Hansika Ambahelagedara, Hansika Ambahelagedara, Hansika Ambahelagedara, Florencia Sorthaix (University of Helsinki)	824
12.4 Shooting the Messenger: Audience Responses to Travel Vloggers	828
Gracey Nana Amoakoa-Amoako (Indiana University Bloomington)	828
12.5 The impact of Verbal, Non-Verbal, and Vocal Cues in Political Posts on TikTok on Users' Verbal Responses	833

Gracey Nana Amoakoa-Amoako (Indiana University Bloomington)	833
12.6 Everyday Digilantism: Platformed Neighbourhoods and the Informal Governance of Urban Life	838
Sophie Anson (Birkbeck, University of London)	838
12.7 “They Can Say It, We Get Banned”: Folk Theories of Platform Governance in Conservative Reddit Communities	842
Morgan Badurak (Louisiana State University)	842
12.8 Building a National Archive of State Social Media	848
Jake Bickford, Claire Newing (The National Archives, UK)	848
12.9 Feeling Safe or Being Safe? Social Media Privacy Strategies Among Women of Color	852
Alexa Bolanos-Carpio (University of Costa Rica)Fanny Ramirez (Louisiana State University) ...	852
12.10 What is ‘acceptable’ versus ‘problematic’ smartphone use for young adults? A qualitative study	859
Magaly Brodeur, Natalia Muñoz Gómez, Andrée-Anne Légaré, Ariane Girard, Catherine Hudon, Adèle Morvannou (Université de Sherbrooke)	859
12.11 A Scoping Review of Empirical Research on Social Exclusion and Responses to Misinformation	864
Ayana Cameron, Vanessa Volpe, MarQia Allen, Abbey Collins, Melissa Holland, Ore Badejoh, Julia Ross (North Carolina State University)	864
12.12 Is Outsourced Moderation Like Call Centre Work? An Examination of Commercial Content Moderation Ecosystem in India	871
Tania Chatterjee (The University of Queensland - Indian Institute of Technology Delhi) Agam Gupta (Indian Institute of Technology, Delhi)	871
12.13 Mapping Canadian Immigration Discourse on YouTube: A Dashboard-Driven Analysis Using LLM-Assisted Multi-Label Sentiment Coding and Geo-Mention Extraction	876
Augustine Farinola (University of Alberta)	876
12.14 Under the Influence: Beauty and Fashion Professionals’ Understandings of Social Media Influencers in Promoting Diversity and Inclusion	882
Jordan Foster (MacEwan University)David Pettinicchio (University of Toronto)	882
12.15 ‘Preparatory Social Capital’ and the Gendered Platform Economy: Chinese Young Mothers Navigating WeChat-Based Multi-Level Marketing	887
Jingjing Fu (Queen Mary University of London)	887
12.16 Populist Posts vs. Community Notes : Strategies for Countering Election Misinformation	893
Hiroyuki Fujishiro (Department of Media and Communication Studies /Hosei University)Mone Saito (Uniquevision)	893
12.16 How TikTok influencers frame sustainability across cultures	898
Lina Gomez-Vasquez, Sebastian Madero (The University of Tampa)	898
12.17 The medium’s agenda or the audience’s feedback? Tensions between editorial lines, metrics and comments in social media in digital newspapers in Chile	905

Francisca Greene (Universidad de los Andes), Cristian Muñoz (Universidad de la Serena)	905
12.18 Short Videos under Conditions of Epistemic Uncertainty: How Young People Interpret Social Challenges	914
Elisa Haralampides (University of the Bundeswehr Munich)	914
12.19 Origins of City Reputation: Measuring Bias in Online Reviews Using Agentic AI and NLP..	918
Rafiazka Hilman, Júlia Koltai (ELTE Centre for Social Sciences)	918
12.20 Constructing Masculinity in Cyberspace: Male Sex Worker Usernames as Performative Identities	922
Kristopher Jackson, Martha Whitfield (University of Massachusetts Dartmouth)	922
12.21 I demand you cancel your cancellation! When politics meets entertainment: The effects of canceling a late night show host	926
Olivia Johnson (University of Houston)	926
12.22 Platform-Driven Right to Information in India: A Mixed-Methods Analysis of Politics and Participation	932
Shravani K, Neha Murthy, Maitreya Prafulla Chitale, Rajorshi Ray (International Institute of Information Technology Hyderabad)	932
12.23 Is “Chat” our friend? A study on conversational AI and parasocial relationships	937
Trevor Kartalija, Julia Belmiro (Drexel University)	937
12.24 Curating the Continent: Instagram and the Visual Promotion of African Tourism	943
Ammina Kothari (Simmons University) Sally Ann Cruikshank (Middle Tennessee State University)	943
12.25 Redesigning Transparency Cues on Instagram for Agency	948
Manish Kumar, Ryan Moore (University of Texas Austin)	948
12.26 Cultures of celebration around femicide perpetrators in Italy: a digital methods research	954
Valentina Lagati Sirino (Sapienza University of Rome)	954
12.27 Trading Grammar for Likes: Linguists and Language Influencers in Czech Language Popularization	961
Helena Lázníčka (Charles University)	961
12.28 Ethical Principles for Engineering Digital Services for Children	964
Beckett LeClair (5Rights Foundation) Maud Stiernet (A Little Lining Comes) Natalia Kucirkova (International Centre for EdTech Impact) Neil Gordon (University of Hull)	964
12.29 Repair in Online Discourse: A Social Listening and Machine Learning Approach to Understanding Circular Economy Conversations	968
Barna Lipics, Pálma Mogyorósi, Zoltán Kmetty, Ágnes Buvár (Eötvös Loránd University Human-Environment Interactions Program)	968
12.30 “Read Me! Conceptualizing Social Media Attention Accumulation as Discursive Work” .	973
Gaultier Lonqueu, Aliette Lambert (University of Bath)	973

12.31 Understanding Ordinary Users of Algorithmic Recommendation Systems: A Digital Ethnographic Approach	978
Michelle Gay Nidoy (Queensland University of Technology)	978
12.32 Digital Placemaking Between Strategy and Everyday Life: Negotiating Belonging in a Hybrid Urban Space	983
Giulia Nunes da Costa (Università della Svizzera Italiana)	983
12.33 Resisting medical misogynoir through digital counterpublics: Examining how Black Women use TikTok to support wellbeing	989
Chelsea Peterson-Salahuddin (University of Michigan Ann Arbor), Sadia Cheema (Seton Hall University), Ammina Kothari (Simmons University), Kelli Burns (University of South Florida), Kayla Hopper (Seton Hall University)	989
12.34 From Embodied Ritual to Platform Performance: The Digital Transformation of Nuosu-Yi Bimo Ethnic Culture	996
Aguo Qubi (Southwest Minzu University) Hao Chen (University of Glasgow) Zhiwei Wang (Southwest Minzu University)	996
12.35 Researching suicide and social media after the Online Safety Act	1001
Aguo Qubi (Southwest Minzu University) Hao Chen (University of Glasgow) Zhiwei Wang (Southwest Minzu University)	1001
12.36 Blondies, Botox, and Bird Legs: Analyzing Gender Humor in TikTok Videos Criticizing Republican Women	1006
Jerrica Rowlett (Bryant University)	1006
12.37 Vine Compilations on YouTube: Social Media Archiving, Digital Nostalgia, and the Preservation of User-Generated Content	1012
Joelle Saunders (Oxford Internet Institute)	1012
12.37 Who Is Afraid Of Fat Bodies? Fatphobic Representations On Social Media: An Analysis	1016
Lia Selvi (Sapienza Università di Roma)	1016
12.38 The Body in Algorithmic Colonialism: Artistic Inquiry into Emotion-AI and Embodied Affective Relations	1021
Chunyi Shen (UAL)	1021
12.39 Parents Under Pressure: Social Media's Impact on Parental Wellbeing	1025
Ali Shull (University of Nebraska-Lincoln)	1025
12.40 "Healing Your Heart": The Production and Commodification of Emotions on Chinese Social Media	1031
Congman Zhang (University of Macau)	1031
12.41 Whose Voice? Power, Geopolitics, and Narrative Dynamics in the #ChineseVirus Discourse	1037
Qian Zhao (University of Strathclyde)	1037
12.42 (Re)listening to a Better Day on YouTube: Memory, Affect, and Hope in Global Charity Singles from Record to AI	1040
Jingyu Zhu (Nanjing University)	1040

Session 1A - Algorithms and AI (1)

Day 1 · Mon 13 July · 9:30–11:00 · RD Suite, ARC

1A.1 Coded Biases and Demographic Stereotyping in Facebook's Curation of AI-Generated Images.

Aanila Kishwar (Indiana University Bloomington)

Introduction

This multi-method study investigates the proliferation of AI-generated images on Facebook, focusing on the coded bias in the curation of images and its cultural implications. From religious imagery and celebrity deepfakes (Bond, 2024) to bizarre DIY creations (Koebler, 2024) and hypersexualized depictions of cartoon characters, recent reports detail how Facebook's algorithm presents a wide array of AI-generated visual content to the average user. In July 2022, Meta's CEO Mark Zuckerberg announced that "social feeds are going from being driven primarily by the people and accounts you follow to increasingly also being driven by AI recommending content that you'll find interesting from across Facebook or Instagram" (Contreras, 2022). As planned, posts from unconnected sources went from occupying 8 percent of the "typical" news feed of a US-based user in the second quarter of 2021 to 29.6 percent in the second quarter of 2024 (Meta, 2024). Facebook has over 3 billion monthly active users worldwide, and their news feeds are filled with AI-generated content (Bond, 2024). Although Facebook and its parent company Meta have previously attempted to paint the news feed algorithm as "some sort of independent force of nature or something that is merely responsive to [the user], without any other value embedded in its design," the algorithm is a crucial part of the platform's business model (Tufekci, 2015, p. 10). For example, researchers have postulated that Facebook's news feed promotes AI-generated images due to the high levels of engagement they attract (DiResta & Goldstein, 2024). This may be attributed to the visually sensational nature of the images and the ease with which they can be generated (DiResta & Goldstein, 2024).

Even when Facebook has no knowledge of a user's interest besides the demographic data required to open the account, the platform continues to populate the user's news feed. This paper addresses a gap in the existing literature by providing a detailed account of what Facebook's algorithms recommend and curate to users who have only provided their age, gender and name to the platform, without signaling any specific

interests or adding any friends on the platform. Moreover, the bulk of the content being recommended to new, unconnected users are AI-generated. This adds another layer of opacity to people's experience of using the social media platform, considering that the inner workings of generative AI technologies like ChatGPT, Dall-E, and Midjourney are equally, if not more, opaque than Facebook's algorithm.

Objectives/Research Questions

This mixed-methods study tackles the following research questions in the context of the internet communicative landscape discussed above:

1. What categories AI-generated images are people being exposed to?
2. What does Facebook's algorithm curate to users based on their demographic data?
3. In the process, does the algorithm reinforce existent societal stereotypes and biases?

Methods

The interpretation and analysis of images is informed by visual semiotic theory, particularly drawing from Peirce's (1931) tri-level mode of visual analysis and Barthes' (1968) conceptualization of denotation-connotation. The inductive coding framework adopted in this study is limited to denotative elements at the iconic and indexical levels. However, the holistic interpretation of the data draws conclusions on the connotative meanings of AI-generated images. Adapted from Sandvig et al. (2014), a sock puppet auditing system was used to collect data. In this form of algorithm auditing, researchers examine platform processes by simulating user behavior through "sock puppet" accounts that collect large quantities of data using automated tools. This study employs a hybrid of classic and sock puppet auditing approaches since Facebook prohibits large-scale automated data scraping (Bruns, 2019).

Using a sock puppet auditing system and criterion sampling, 96 AI-generated images were collected from four demographic-specific accounts (Gen X woman, Gen X man, millennial woman, millennial man). The images were coded in an inductive and iterative process, noting the presence of human subjects and other denotative visual elements.

Results

The findings demonstrate how Facebook's algorithmically curated content varies significantly based on the demographic profiles of its users, reflecting the platform's use of age- and gender-based profiling to recommend visual media. The prevalence of

specific categories—such as sexually suggestive imagery in the Gen X man’s account, military and birthday imagery in the Gen X woman’s account, landscapes and religious imagery in the millennial man’s account, and domestic spaces in the millennial woman’s account—indicates that the algorithm caters content to stereotypical interests associated with these demographic groups.

A notable observation is the sexualized depiction of women in the data collected from the Gen X man’s account, containing a high number of such images, overlapping with the “celebrity” and “athletic” categories. This pattern underscores how AI-generated content often perpetuates reductive gender portrayals, emphasizing physicality and traditional notions of beauty. In contrast, the millennial man’s account exhibited no such depictions, focusing instead on landscapes and religious imagery, which may suggest an assumption by the algorithm about age-specific preferences for spiritual or escapist imagery.

The Gen X woman’s account, meanwhile, prominently featured military imagery, highlighting the valorization of service and the propagation of the military-industrial complex. The juxtaposition of such images with birthday celebrations, often involving veterans, reinforces a cultural narrative that links patriotism to militarism, subtly endorsing these values through algorithmic curation. Similarly, the overwhelming representation of kitchens and living spaces in the millennial woman’s account reflects a reinforcement of domestic stereotypes. The absence of human subjects in most of these images further abstracts these spaces into aspirational ideals rather than functional realities, aligning with broader societal trends in lifestyle marketing and interior design aesthetics.

The reliance on AI-generated content in Facebook’s visual recommendations raises critical questions about the biases embedded within generative algorithms. The aestheticization of domesticity, the fetishization of female bodies, the romanticization of religious or natural imagery, and the militarization of everyday life reveal how the algorithm reinforces cultural norms through visual media. It is not possible to trace who generated these images with what prompt using which genAI software (DiResta & Goldstein, 2024). The reasoning behind Facebook’s AI-powered algorithm curating these images cannot be determined with certainty; nor is it possible to know why the genAI software’s algorithm created images with the specific elements discussed above.

Future Work

The opacity of the algorithm, paired with platforms’ tendency to profile users according to age and gender, emphasizes a need for greater scrutiny of algorithmic practices and their implications for cultural representation in digital spaces. Future work can closely

analyze the images themselves within this data set, while another possibility would be to scale up this study by incorporating more sock puppet accounts of different generations and genders. Additionally, despite the increased use of AI-generated content for political means, including campaigns, elections, and even by ruling parties, there are few research studies that analyze the contents of such imagery or discuss their impact. All of these would be appropriate avenues to utilize the learnings from this study.

References

- Barthes, R. (1968). *Elements of Semiology*. Annette Lavers & Colin Smith (Trans.). New York: Hill and Wang.
- Bond, S. (2024, May 14). AI-generated spam is starting to fill social media. Here's why. NPR. <https://www.npr.org/2024/05/14/1251072726/ai-spam-images-facebook-linkedin-threads-meta>
- Bruns, A. (2019). After the 'APocalypse': Social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, 22(11), 1544–1566. <https://doi.org/10.1080/1369118X.2019.1637447>
- Contreras, B. (2022, July 28). Facebook is in trouble. Its escape plan: Turn into TikTok. *Los Angeles Times*. <https://www.latimes.com/business/story/2022-07-27/facebook-is-in-trouble-its-escape-plan-turn-into-tiktok>
- DiResta, R., & Goldstein, J. A. (2024). How spammers and scammers leverage AI-generated images on Facebook for audience growth. *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-151>
- Koebler, J. (2024, March 19). Facebook's Algorithm Is Boosting AI Spam That Links to AI-Generated, Ad-Laden Click Farms. *404 Media*. <https://www.404media.co/facebooks-algorithm-is-boosting-ai-spam-that-links-to-ai-generated-ad-laden-click-farms/>
- Meta. (2024). *Widely Viewed Content Report: What People See on Facebook*. Meta. <https://transparency.meta.com/en-gb/data/widely-viewed-content-report/>
- Peirce, C. S. (1931). *Semiotics* (Vol. 8). José Teixeira Coelho Netto (Trans.).
- Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014). Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms. *Data and Discrimination: Converting Critical Concerns into Productive Inquiry*, 22, 4349–4357.

Tufekci, Z. (2015). Facebook Said Its Algorithms Do Help Form Echo Chambers, and the Tech Press Missed It. *New Perspectives Quarterly*, 32, 9.

1A.2 Hybrid AI Visuals and Deepfakes as instruments of political deception in the pre-election period of Bangladesh.

Arafat Al Yeasin (University of Liverpool)

Introduction

AI-manipulated visuals have reshaped contemporary global politics and related scholarships and blurred well-defined distinctions and understanding between authenticity and falsehood (Bennett & Livingston, 2018). Because generative AI and synthetic media programmes are widely accessible and they have eased and enabled the production of highly realistic visual content that circulates widely across social media platforms.

Ongoing scholarship explored the risks of deepfakes and AI-mediated synthetic media in relation to misinformation, electoral interference, and democracy, but much of this work remains focused on technical detection, factual falsity, and explicit deception in the Western context (Chesney & Citron, 2019; Vaccari & Chadwick, 2020). Although these approaches are empirically valuable, they offer limited empirical knowledge on the context where political communication through deception, insinuation and moral framing.

Facebook already established not merely as a social networking platform but as a central point for political communication, news consumption and informal political mobilisation in Bangladesh (Bennett & Segerberg, 2012). Visual content shared on Facebook has shaped role in shaping political narratives and interpretations in recent events of political tension, protest, and electoral uncertainty in Bangladesh.

In this media ecology, AI-manipulated visual media should not be considered isolated or technically discrete elements, but they circulate through social media, particularly Facebook as a part of broader hybrid visual forms.

These hybrid AI visuals generate meaning not only in explicit way but also perform in a strategic, implicit way where they invite the users to deduce political meaning rather than evaluate factual claims (Messaris & Abraham, 2001; Couldry, 2012). In this way, users left with a confusion of facts and may impact their political decision making, especially in the pre-election context when the political situation is intense.

Current research on AI-mediated political communication is largely concentrated on technical aspects and identifying deepfakes, so the empirical knowledge on political deception and meaning making through deepfakes and hybrid AI visuals are understudied especially on Global South context. Furthermore, Deepfakes are conceptualised in this study as a part of broader ecology of hybrid AI visual rather than positioning it as a stand-alone and random method of visual deception.

This paper seeks to address the empirical gaps by exploring AI-manipulated deepfakes and hybrid visuals circulated on Facebook during pre-election period from November 2025 to February 2026 in a post-uprising politically volatile Bangladesh and examining their dominant political functions and meaning-making.

Objectives/Research Questions

The objective of this study is to examine how AI manipulated visual media, especially deepfake and hybrid AI visuals utilised for political deception on Facebook during Bangladesh's electoral period, and to analyse the political functions of such content and understand their meaning-making.

This study understands political deception as only the dissemination of explicitly false claims (Entman, 1993) but as a process where AI-manipulated visuals misled political perception by moral judgment, insinuation of falsehood and amplifying opposing interpretations of narratives and obscure contextual reality (Papacharissi, 2015; Messaris & Abraham, 2001).

Deepfake elements enable such deception not only through altered realism, but also through their integration into hybrid visual forms that operate strategically in ways that enable political deception.

Based on this conceptual understanding, the study addresses the following research questions:

1. How are AI-manipulated visuals, including deepfakes and hybrid AI content, used for political deception on Facebook during Bangladesh's pre-election period?
2. What dominant political functions, such as opponent de-legitimation, character assassination, insinuation visual falsification, and narrative polarisation, are performed through the visual and textual construction of these visuals?
3. How visual and textual features of AI-manipulated content contribute to deceptive political meaning-making during pre-election period?

These questions enable a function-centred analysis of AI-mediated political deception in a digitally constrained and polarised electoral context in Bangladesh.

Methods

A qualitative approach was adopted as methodology for this study which was grounded in political communication studies. The main study focus is on AI-manipulated visual media such as photos, Videos posted on social media platform Facebook from the period from November 2025 to February 2026, a timeline characterized by heightened political tension ahead of national election.

In this study, deepfakes are understood as a subset of AI-manipulated visuals that simulate realistic appearances of people or events through AI generation or alteration. But for the findings, this study extends beyond deepfakes and explore a broader realm of hybrid AI visuals that combines AI-generated or altered elements with real imagery, individuals.

Purposive sampling strategy was designed with specific criteria to capture the data on publicly accessible Facebook pages, groups, and media outlets associated with political actors, political commentators, activist networks, and highly engaged visual political content.

Furthermore, a limited snowballing strategy was taken to ensure variation in dataset in which way different edited variants of visuals were identified, contextualised versions were sampled within the set sampling strategy of this study.

Qualitative content analysis combined with multimodal discourse analysis approach was taken for the data analysis method. Each data from the dataset is coded for dominant political function based on the visual and textual cues. The coding framework is mainly consists of four primary functions: opponent de-legitimation, character assassination, insinuation visual falsification, and narrative polarisation.

Results

Although the analysis of the study is ongoing but preliminary findings on AI-manipulated media on Facebook ahead of election suggest patterns consistent with study's analytical framework on Facebook ahead of February 2026 election in Bangladesh. Primary findings also suggest that AI-manipulated media on Facebook during the study timeline mainly function as an instrument of political deception were delegitimising opponents and polarising the narratives beside factual falsity.

First, this study finds that AI-manipulated visuals that combine deepfakes and cheapfakes techniques is already widespread on Facebook and being shared by the political actors from several camps to advance their political aim as the election approaches. These include misrepresentations of political opponents, fabricated character depictions, and political endorsements by AI-generated fictional individuals.

Second, circulated AI-manipulated content is often visually imperfect (e.g., unnatural facial movements, mismatching background, distortion) but politically persuasive with high engagement in terms of user interaction. This study documented at nearly a hundred pieces of AI-manipulated content till Mid-January 2026 across pages, profiles and groups on Facebook. Combinedly, they draw over 1.6 million engagements within 24 hours of posting.

Third, manipulative narratives are found predominate deepfakes and hybrid AI media, followed by disinformation, hate speech, religious provocation and character

assassination. These patterns are aligned with outlined political functions and research questions of this study.

Fourth, these contents often try to undermine moral legitimacy of the opposing parties and individuals through personalised political character-focused attacks. Most of the visuals appeared as insinuating visual falsification on ideas of betrayal of the uprising, violence, corruption and immorality with both implicit and explicit assertion. For example, AI-generated visuals from Jamaat-aligned sources often portray their opponents as corrupt, while other political actors use AI-generated endorsements or nationalist, symbolic representations to boost their legitimacy.

Fifth, user engagement in these posts indicate that these patterns of AI-generated deception offer falsified interpretations and allowing misleading political meanings to influence and confuse users. Many of the users' comments show that they believe the content to be real and react in a way that are somewhat influenced by them. This observation supports this study's focus on how manipulated visual and textual features develop deceptive meaning-making.

Together, these ongoing findings suggest that AI-manipulated political visuals in the pre-election context in Bangladesh is proliferating and functioning in moral framing, distortion of credibility and insinuation of falsehood.

Future Work

Further research may extend this research by exploring deceptive political functions of AI-manipulated visuals from before and after the electoral events providing a comparative view in political narrative following the electoral outcome. This comparison may reveal how political deceptive strategies are set into motion, intensified, or diminished once the electoral outcomes are available.

Future research may provide a comparative analysis on different platforms such as YouTube, Instagram or TikTok to examine how deepfakes and AI-mediated media function in political agenda setting or narrative construction. Each platform offers different media ecology and deceptive visuals may function differently and produce several patterns or findings.

Moreover, audience oriented ethnographic approaches to examine and explore how users on these platforms received, interpret, and negotiate AI-mediated political deception on day-to-day basis may complement political communication practices. Such studies would extend function centred analysis of AI-manipulated visuals on different platforms within the broader areas of political meaning-making and deceptive narrative alignment.

References

- Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122–139. <https://doi.org/10.1177/0267323118760317>
- Bennett, W. L., & Segerberg, A. (2012). The logic of connective action: Digital media and the personalization of contentious politics. *Information, Communication & Society*, 15(5), 739–768. <https://doi.org/10.1080/1369118X.2012.670661>
- Chesney, R., & Citron, D. K. (2019). Deepfakes and the new disinformation war. *Foreign Affairs*, 98(1), 147–155
- Couldry, N. (2012). *Media, society, world: Social theory and digital media practice*. Polity Press.
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51–58
- Messararis, P., & Abraham, L. (2001). The role of images in framing news stories. In S. D. Reese, O. H. Gandy, & A. E. Grant (Eds.), *Framing public life: Perspectives on media and our understanding of the social world* (pp. 215–226). Lawrence Erlbaum Associates.
- Papacharissi, Z. (2015). *Affective publics: Sentiment, technology, and politics*. Oxford University Press
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1), 1–13. <https://doi.org/10.1177/2056305120903408>
-

1A.3 AI Fortune-Teller: Procedural Rhetoric and the Algorithmic Construction of Trust in AI Divination. *Erika*

Ningxin Wang (The Chinese University of Hong Kong, Shenzhen)

Introduction

In the quiet hum of university dormitories and the bustling solitude of internet cafés across China, a new form of digital spirituality is taking shape. A growing cohort of educated young women, navigating the intense pressures of modern Chinese society, are turning to Generative AI platforms - such as the domestic model DeepSeek and global counterparts like ChatGPT - for a practice as ancient as their civilization: Bazi (八字) divination. Instead of consulting roadside mystics or visiting temples, these users meticulously input the precise hour of their birth into a dialogue interface, asking the algorithm to decode their destiny, predict career trajectories, and assess romantic compatibility.

This emergent phenomenon, which we term "AI divination," represents a paradoxical intersection of hyper-modern computational power and traditional metaphysical belief. While digital fortune-telling apps have existed for decades, the integration of Large Language Models (LLMs) marks a distinct shift. Unlike the static responses of earlier software, LLMs offer a conversational, responsive, and eerily "human" interaction. This practice has exploded on social media platforms like Xiaohongshu (Rednote), where hashtags related to "AI fortune telling" garner millions of views. Here, users share complex "prompts" as if they were magical incantations, debating which model yields the most "accurate" reading of their fate.

However, this practice presents a profound epistemic puzzle. LLMs are probabilistic text generators designed for pattern matching, not causal prediction; they are objectively untrustworthy oracles capable of hallucination. Yet, users exhibit high levels of trust in these systems, often valuing them over human practitioners. This paper addresses a central question relevant to the study of social media and society: How is trust formed in an AI system for a purpose it was not designed for, and for which its outputs are inherently unfalsifiable?

To answer this, we move beyond traditional models of trust based on accuracy or institutional reputation. Drawing on Ian Bogost's concept of Procedural Rhetoric - persuasion through rules-based interactions - we argue that trust in AI divination is not a passive belief in the output, but an active state constructed through the process of interaction. This extended abstract details how users are persuaded to accept the AI as

a credible spiritual guide through specific rhetorical procedures, highlighting the shifting nature of authority in a datafied society.

Objectives/Research Questions

The primary objective of this study is to locate the emergent practice of AI divination within the specific socio-technical moment facing China's youth. This generation is coming of age in an era defined by *neijuan* (involution)—a state of intense, relentless competition with diminishing returns. In this high-pressure environment, divination functions less as superstition and more as a psychological coping mechanism for anxiety management. However, the migration of this practice from human practitioners to algorithmic agents presents a significant theoretical challenge for scholars of human-computer interaction (HCI) and Science and Technology Studies (STS).

This research aims to address the "calibration problem" inherent in this phenomenon: a dangerous misalignment where a user's high trust is placed in a technology with low actual capability for the assigned task. As established in trust literature (Gille et al., 2020), trust (a psychological state) and trustworthiness (an objective property) are disentangled concepts. An LLM is objectively untrustworthy as a metaphysical oracle; its outputs are statistical hallucinations rather than causal predictions.

Therefore, the central research question guiding this inquiry is:

How and why do users develop significant trust in an AI system for a task -divination -for which its outputs are inherently unfalsifiable and for which the AI is not an objectively trustworthy tool?

To answer this, this article seeks to:

Analyze the specific user interactions and interface designs that bridge the gap between the AI's lack of intent and the user's perception of benevolence.

Apply the theoretical framework of Procedural Rhetoric to explain how trust is manufactured through rule-based persuasion rather than technical verification.

Investigate the role of technonationalism in shaping trust, specifically regarding the preference for domestic AI models (DeepSeek) over foreign competitors.

Methods

This study adopts a qualitative, interpretive approach to explore the lived experiences and sense-making processes of AI divination users. Data was collected through in-depth, semi-structured interviews with 15 university-educated young women in China. Participants were recruited via purposive and snowball sampling on social media

platforms prominent in the digital divination community, specifically Xiaohongshu (Rednote) and Douban. Inclusion criteria required participants to have consulted an LLM for divination on multiple occasions, ensuring the data reflected established practice rather than mere novelty seeking.

Interviews, lasting 60 to 90 minutes, were conducted in Mandarin and focused on user motivations, prompting rituals, and the comparative perception of AI versus human practitioners. The analytical strategy employed a two-stage process: first, an inductive thematic analysis was used to identify key patterns in user behavior and sentiment. Second, these themes were re-interpreted through the deductive lens of Procedural Rhetoric to explain how specific interactional mechanics function as persuasive arguments for the AI's credibility.

Results

Our analysis reveals that trust in AI divination is not a passive acceptance of technology, but an active achievement constructed through four distinct "rhetorical procedures."

The Rhetoric of Objectivity: The Incorruptible Machine The foundational procedure constructs the AI as an "incorruptible" actor. Unlike human fortune-tellers, who are often viewed as commercially compromised and manipulative, the AI's interaction is free, instantaneous, and devoid of upselling. This transactional purity enacts a powerful argument: the lack of profit motive implies a lack of bias. The sterile, non-commercial interface paradoxically frames the machine as a spiritually superior, neutral conduit for truth.

The Rhetoric of Authority: Rituals of Validation Authority is manufactured through user-led rituals. Users first engage in a Prompting Ritual, sourcing complex "incantations" from social media to "train" the oracle. Crucially, they play a Validation Game, testing the AI on known past events. When the AI "guesses" correctly—a feat of statistical inference—it provides empirical "proof." This validation acts as a cognitive anchor: having verified a truth about the past, the user is procedurally persuaded to accept the AI's unfalsifiable predictions about the future.

The Rhetoric of Nativism: Technonationalist Alignment Trust is further shaped by geopolitical identity. Participants expressed a strong preference for the domestic model DeepSeek over foreign competitors like ChatGPT. This relies on a technonationalist argument: a Chinese model is viewed as culturally "native" (better equipped to interpret traditional Bazi) and ideologically "safe" (shielded from US

surveillance). Trust is thus generated by aligning the divination ritual with narratives of data sovereignty and national pride.

The Rhetoric of Personalization: The Intimate Mirror Finally, trust is solidified through intimacy. The procedure of inputting unique birth data creates a vulnerability that frames the output as a bespoke "mirror" to the self. For expert users, this becomes a dialogic co-construction where iterative prompting refines the reading. This interactive depth persuades the user that the system possesses profound understanding, effectively masking its actual nature as a probabilistic text generator.

Future Work

The emergence of AI divination signals a significant shift in algorithmic authority, opening urgent avenues for future inquiry. First, comparative research is needed to determine if the Rhetoric of Nativism is specific to the Chinese context or if similar ideological alignments shape AI trust globally. Second, future scholarship should complement these qualitative findings with experimental designs. By manipulating variables such as service cost, model origin, or response latency, researchers can quantitatively measure which procedural elements most effectively drive trust formation. Finally, this study advocates for an ethic of "Procedural Humility" in AI design. To address the dangerous "calibration problem" where users over-trust the system, future work must explore how interface friction and the visualization of uncertainty can disrupt the illusion of omniscience, ensuring user trust is earned rather than rhetorically manufactured.

References

- Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial intelligence and human trust in healthcare: A systematic review. *Journal of Medical Internet Research*.
- Bogost, I. (2007). *Persuasive Games: The Expressive Power of Videogames*. MIT Press.
- Chong, D., & Ross, A. D. (2022). A multilevel framework of trust in digital media. *New Media & Society*.
- Gille, F., Vranken, L., & D'hoore, W. (2020). Trust versus trustworthiness in explainable AI. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*.
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*.

Pasquale, F. (2015). *The Black Box Society*. Harvard University Press.

Xiang, B. (2021). The age of “involution” (neijuan). *Anthropology News*.

1A.4 Algorithm Adoration, Synergy, Discord, and Apathy: Categorizing Human–Algorithm Relationships in Social Media.

Biying Wu-Ouyang (Education University of Hong Kong)

Introduction

Algorithms play an important role in content personalization and recommendations on social media platforms, significantly shaping users' digital experiences. While much research has focused on users' rational and cognitive awareness of algorithms, limited attention has been paid to users' affective evaluations and coping behaviors toward algorithms, which are critical components of algorithmic literacy. This study addresses these gaps by adopting a folk theory perspective to examine how individuals understand, evaluate, and cope with algorithms within social media environments. Drawing on in-depth interviews with 26 Xiaohongshu users, we developed a four-quadrant model that integrates levels of human agency (high/low) and attitudes toward algorithms (appreciation/aversion) to classify types of human–algorithm relationships. These include algorithm adoration, synergy, discord, and apathy. Our findings advance the human–AI agency framework by integrating affective perspectives and offering new theoretical insights into the diverse ways individuals interact with algorithms on social media platforms.

Objectives/Research Questions

RQ1. How do users understand, evaluate, and cope with the algorithm in their daily use of Xiaohongshu?

RQ2. How does the affective perspective influence the human–algorithm relationship, and what implications do these dimensions have for developing algorithmic literacy?

Methods

To achieve the objectives of this study, we conducted in-depth interviews with 26 users of Xiaohongshu, a social media platform that has become one of the most downloaded apps in China and globally. Xiaohongshu's rise has been partly driven by the TikTok ban in early 2025, which led millions of self-identified American “TikTok refugees” to migrate to Xiaohongshu digitally (Baptista et al., 2025; Xiao & Zhang, 2025).

Results

Through in-depth interviews, we provide rich insights into individuals' perceptions of algorithms and introduce a four-quadrant model (Figure 1) that integrates levels of human agency (high/low) and attitudes toward algorithms (appreciation/aversion) to classify types of human–algorithm relationships.

Future Work

Admittedly, this study has several limitations that point to directions for future research. First, as the research findings and conceptualized classification of users' relational dynamics and negotiator types are situated within the context of the Xiaohongshu algorithm, future studies should examine our findings and conceptualizations in wider contexts across different platforms and sociocultural backgrounds, such as TikTok and Douyin. It is open to examination, improvement, or revision by future studies. Second, the implementation and practices of the proposed classification for algorithm literacy training and education are not yet sufficiently validated and examined. This study aims to primarily achieve theoretical universality not through statistical generalizability based on representative sampling but through analytical generalizability, which is the qualitative validation of a theory, argument, and concept through the repetitive examination, pattern-matching, and comparisons across multiple cases, echoing the logic of replication (Carminati, 2018; Yin, 2009).

Future studies should provide more empirical support based on multiple methods for assessing our proposed relational dynamics and personal types. Third, several informants had backgrounds in media or the social sciences, which may have been influenced by the researchers' positionality and Xiaohongshu's algorithms. While this may limit the generalizability of findings, their academic backgrounds provided high digital literacy and nuanced perspectives on interactions with platform algorithms. Finally, since our first practical implication is to locate users in the complex human–algorithm relationship, we do not intend to claim that any particular type of relational dynamics or personal type is the most ethical, socially beneficial, or advocated one. However, some types are likely to encounter more negative effects than others. For example, adolescents and older adults who are algorithmic followers may experience some negative effects, such as social media addiction. We thus advocate for additional research and discourse regarding the normative implications of various human–algorithm-related dynamics.

References

Baptista, E., Hu, K., Oladipo, D., & Hu, K. (2025, January 15). Over half a million 'TikTok refugees' flock to China's RedNote. Reuters. Retrieved from

<https://www.reuters.com/technology/over-half-million-tiktok-refugees-flock-chinas-RedNote-2025-01-14>.

Carminati, L. (2018). Generalizability in qualitative research: A tale of two traditions. *Qualitative Health Research*, 28(13), 2094-2101.

Xiao, Y., & Zhang, J. (2025). “No longer our place”: Tiktok refugees and the politics of digital migration to Xiaohongshu. *New Media & Society*, 14614448251368890. <https://doi.org/10.1177/14614448251368890>

Yin, R. K. (2009). *Case study research: Design and methods* (Vol. 5). Sage.

Session 1B - Misinformation and Disinformation

Day 1 · Mon 13 July · 9:30–11:00 · Seminar Suite A, ARC

1B.1 Seeing Is No Longer Believing: Audience Trust, Scepticism, and AI-Generated Visuals on Social Media.

Chisom Omeokachie (Independent Researcher)

Introduction

Visual content has long played a central role in how information circulates on social media. Images and videos are often perceived as direct evidence of events, people, or places, lending them a persuasive power that exceeds that of text alone (Grabe & Bucy, 2009; Messaris & Abraham, 2001). However, the rapid proliferation of generative artificial intelligence has disrupted this assumption. AI-generated images and videos now circulate alongside authentic content on platforms such as Instagram, X, and TikTok, frequently without clear indicators of origin or authenticity (Vaccari & Chadwick, 2020; Nightingale & Farid, 2022). As a result, users increasingly encounter visual material whose credibility is uncertain.

This shift has intensified concerns about misinformation, democratic participation, and public trust. Monitoring reports suggest that the volume of deepfakes circulating online increased by an order of magnitude between 2022 and 2023 (Sensity AI, 2024), raising alarms among policymakers and civil society organisations, particularly during a period marked by an unprecedented number of national elections worldwide in 2024 (World Economic Forum, 2024). While much scholarly and technical attention has focused on improving detection systems and verification tools, less attention has been paid to how ordinary users experience and respond to this uncertainty in everyday social media contexts (Eslami et al., 2018; Kizilcec, 2016).

This paper argues that the social consequences of AI-generated visual content cannot be understood solely through technical performance metrics. Instead, they must be examined through the perspectives of users who are required to interpret, judge, and act on visual information under conditions of ambiguity. Drawing on a mixed-methods study, this paper examines how social media users assess the authenticity of AI-generated visuals, how confident they feel in these assessments, and how this shapes trust, scepticism, and disengagement. In doing so, it contributes to ongoing discussions about how AI reshapes meaning-making, credibility, and responsibility on social media platforms.

Objectives/Research Questions

The aim of this study is to explore how the presence of AI-generated visual content affects audience trust and credibility judgments in social media environments. Rather than evaluating the technical accuracy of detection systems, the study focuses on users' subjective experiences of uncertainty, confidence, and responsibility when encountering potentially synthetic media.

The study addresses the following research questions:

RQ1: How do social media users assess the authenticity of AI-generated and authentic visual content?

RQ2: How does exposure to AI-generated visuals affect users' confidence in their own credibility judgments?

RQ3: How do users describe responsibility for verifying visual content on social media platforms?

These questions foreground audience interpretation and social experience, aligning with Social Media & Society's emphasis on the cultural and societal implications of digital technologies.

Methods

The study employed a mixed-methods design combining a visual authenticity assessment task with follow-up semi-structured interviews. This approach was chosen to capture both observable judgment patterns and users' reflective accounts of how they reason about visual credibility under conditions of uncertainty.

Seventy-six participants completed an online task in which they were shown a series of images and short videos and asked to classify each item as either "real" or "AI-generated." The stimuli included thirteen visuals spanning four broad categories commonly encountered on social media: neutral, political, social, and cultural content. Participants were not provided with external verification tools or contextual cues, reflecting typical everyday encounters with visual content on social media, where credibility judgments are often made quickly and without system support.

Following the task, ten participants were selected for semi-structured interviews based on variation in detection performance. The interviews explored participants' reasoning strategies, confidence trajectories, emotional responses, and perceptions of responsibility for verifying AI-generated media. Together, the quantitative and qualitative data provide insight into both how users make credibility judgments and how they interpret their own limitations in doing so.

Results

Analysis of the task and interview data revealed three interrelated patterns that illuminate how AI-generated visuals shape audience trust and engagement on social media.

Confidence Erosion and Guessing Behaviour

Participants achieved a reasonably high overall accuracy rate (71%) when distinguishing AI-generated from authentic visuals. However, this performance masked a significant decline in confidence over the course of the task. Many participants reported feeling increasingly uncertain as they encountered more images, even when their accuracy remained stable.

Rather than refining their evaluation strategies, participants often described resorting to guessing. Interviewees expressed frustration with their inability to articulate why an image appeared real or fake, noting that traditional cues such as lighting, facial detail, or image quality no longer felt reliable. This erosion of confidence suggests that repeated exposure to synthetic media may undermine users' willingness to engage thoughtfully with visual content, replacing deliberation with resignation.

Scepticism Bias and Over-Distrust

A second pattern was a systematic tendency to misclassify authentic images as AI-generated. Participants frequently described adopting a defensive mindset, assuming that "anything could be fake." This scepticism bias was especially pronounced for politically or socially salient images, which participants scrutinised more intensely due to perceived stakes.

While scepticism is often framed as a desirable response to misinformation, excessive distrust had unintended consequences. Several interviewees noted that they no longer trusted their own judgments or the authenticity of visual evidence more broadly. In this context, scepticism did not empower users but instead contributed to disengagement and confusion, raising questions about how persistent uncertainty reshapes public trust on social media.

Trust Displacement and Deferral of Responsibility

As confidence declined, participants increasingly framed credibility assessment as a responsibility that should lie with platforms, experts, or automated systems rather than individual users. Interviewees expressed expectations that social media companies or AI tools should "handle" verification, particularly given the perceived sophistication of generative technologies.

This trust displacement reflects a broader shift in how responsibility is understood in AI-mediated environments. Rather than positioning themselves as active evaluators, users described withdrawing from the task of verification altogether. While this deferral may be a rational response to cognitive overload, it also raises concerns about how

agency and accountability are redistributed in platform-governed information ecosystems.

Taken together, these findings indicate that AI-generated visuals affect not only accuracy but also how users experience credibility itself.

Future Work

Future research should examine how sustained exposure to AI-generated visual content shapes user confidence and engagement over time. Longitudinal studies could explore whether repeated encounters with ambiguous visuals lead to lasting scepticism, disengagement, or adaptive strategies for credibility assessment. Comparative research across cultural and political contexts would also help clarify how existing levels of media trust influence responses to synthetic media on social platforms.

Further work is needed to investigate how platform design mediates these dynamics. Rather than focusing solely on detection accuracy, future studies could examine how different credibility signals, such as uncertainty indicators, contextual explanations, or graduated authenticity cues, affect user confidence, agency, and trust. Understanding how users interpret and respond to such signals is crucial, particularly if persistent ambiguity produces a form of credibility fatigue that undermines participation and shared sense-making.

Finally, this work points to the need for more socially responsive approaches to media literacy and platform governance. If the cognitive and emotional costs of constant verification become unsustainable for users, responsibility for mitigating misinformation cannot rest solely on individuals. Future research should therefore explore how responsibility for credibility assessment is negotiated among users, platforms, and institutions in AI-mediated social media environments.

References

- Eslami, M., Krishna Kumaran, S. R., Sandvig, C., & Karahalios, K. (2018). Communicating algorithmic process in online behavioral advertising. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Article 432, 1–13. <https://doi.org/10.1145/3173574.3174006>
- Grabe, M. E., & Bucy, E. P. (2009). Image bite politics : news and the visual framing of elections (1st ed.). Oxford University Press.
- Kizilcec, R. F. (2016). How much information? Effects of transparency on trust in an algorithmic interface. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 2390–2395. <https://doi.org/10.1145/2858036.2858402>
- Messararis, P., & Abraham, L. (2001). The role of images in framing news stories. In Framing public life (pp. 231-242). Routledge.

Nightingale, S. J., & Farid, H. (2022). AI-synthesized faces are indistinguishable from real faces and more trustworthy. *Proceedings of the National Academy of Sciences - PNAS*, 119(8), Article e2120481119.

<https://doi.org/10.1073/pnas.2120481119>

Sensity AI. (2024). The state of deepfakes 2024. <https://sensity.ai/reports>

Vaccari, C., & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media + Society*, 6(1). <https://doi.org/10.1177/2056305120903408>

World Economic Forum. (2024). Global risks report 2024.

<https://www.weforum.org/publications/global-risks-report-2024>

1B.2 Infodemics and Routine Vaccine Hesitancy in Low- and Middle-Income Countries: A Systematic Review of Social Media and Communication Environments.

*Isabela Braga de Moura, Claudia Araujo (COPPEAD Graduate School of Business),
Marina Martins Siqueira (Hospital Israelita Albert Einstein)*

Introduction

Routine immunization is a cornerstone of global public health and a key mechanism for achieving the Sustainable Development Goals (Immunization Agenda 2030 Partners, 2024), particularly in low- and middle-income countries (LMICs). Despite the availability of effective vaccines, vaccine hesitancy—defined as delayed acceptance or refusal of vaccines despite access (MacDonald & SAGE Working Group on Vaccine Hesitancy, 2015)—remains a persistent and growing challenge. This phenomenon is especially pronounced in LMICs, where structural inequalities, trust deficits, and information asymmetries intersect.

In parallel, the rapid expansion of digital and social media platforms has intensified what the World Health Organization describes as an infodemic: an overabundance of information, including misinformation and disinformation, that complicates individuals' ability to identify trustworthy guidance (Pan American Health Organization, 2020). Social media platforms play a central role in this process by accelerating the circulation of vaccine-related narratives, amplifying emotionally charged content, and enabling the visibility of anti-vaccination discourses across national and cultural boundaries. While prior research has examined vaccine hesitancy and misinformation separately, the mechanisms through which infodemics—particularly those shaped by platform affordances and media environments—contribute to hesitancy toward routine vaccines remain underexplored.

This gap is especially evident in LMICs, where digital media ecosystems coexist with limited institutional trust, uneven access to authoritative information, and strong community-level influences (Acharya et al., 2024; Simas & Larson, 2021). Existing reviews have focused largely on COVID-19 vaccines (Gonçalves et al., 2023), high-income countries, or specific platforms, leaving a fragmented understanding of how infodemics shape vaccine decision-making in resource-constrained settings.

Responding to this gap, this study systematically reviews empirical evidence on the relationship between infodemics and hesitancy toward routine vaccines recommended by the World Health Organization (WHO, 2025) in low- and middle-income countries. The analysis is guided by two complementary frameworks: the Vaccine Hesitancy Determinants Matrix (MacDonald & SAGE Working Group on Vaccine Hesitancy, 2015), which captures contextual, individual, and vaccine-specific drivers of hesitancy, and a

multilevel health service ecosystem perspective (Beirão et al., 2017) that distinguishes micro-, meso-, and macro-level actors and interventions. Together, these frameworks enable an examination of how infodemics shape vaccination decisions across multiple levels of influence. By integrating insights from social media studies, public health, and policy research, the paper advances an interdisciplinary understanding of how communication environments, digital platforms, and information flows influence vaccination attitudes and behaviors.

Objectives/Research Questions

The study aims to examine how infodemics contribute to hesitancy toward routine vaccines in low- and middle-income countries and to identify policy-relevant strategies to counter these effects. Specifically, it addresses two research questions:

1. How do infodemics—particularly those mediated through social and digital media—contribute to hesitancy toward routine vaccines recommended by the WHO in LMICs?
2. What multilevel public policy and communication initiatives are proposed in the literature to mitigate vaccine hesitancy in infodemic contexts?

By answering these questions, the study seeks to clarify the social, communicative, and institutional mechanisms linking platform-based information environments to vaccine-related decision-making.

Methods

A systematic literature review was conducted following PRISMA 2020 guidelines and registered in PROSPERO (CRD42024611043). Five databases—Scopus, Web of Science, PubMed, EBSCO, and the Virtual Health Library—were searched in October 2024 for peer-reviewed empirical studies published in English, Spanish, or Portuguese. The search strategy combined terms related to vaccine hesitancy, infodemics (including misinformation and disinformation), and routine vaccines recommended by the WHO. Studies focusing exclusively on COVID-19 vaccines, non-empirical publications, and research without separable data for LMICs were excluded. Snowballing techniques were used to identify additional relevant studies.

Twenty-nine studies published between 2011 and 2024 met the inclusion criteria. Data extraction captured study characteristics, geographic focus, vaccine type, methodological approach, and key findings. Study quality was assessed using the Mixed Methods Appraisal Tool (MMAT 2018) (Nha HONG et al., 2018).

Data analysis employed Directed Qualitative Content Analysis (Assarroudi et al., 2018), guided by the WHO SAGE Vaccine Hesitancy Determinants Matrix (MacDonald & SAGE Working Group on Vaccine Hesitancy, 2015). Infodemic effects were mapped onto hesitancy determinants using thematic synthesis, enabling the identification of dominant communication-related pathways. Finally, proposed interventions were categorized using a multilevel service ecosystem framework (Beirão et al., 2017),

distinguishing actions at the micro (individual), meso (organizational/community), and macro (policy/government) levels.

Results

The systematic review included 29 empirical studies published between 2011 and 2024, with a noticeable increase after 2020. Most studies were conducted in Asia (70%), followed by Africa (17%) and Latin America (7%), revealing regional concentration and important geographic gaps. Many studies examined multiple antigens, indicating that infodemics often shape vaccination attitudes broadly rather than targeting a single vaccine.

Contextual determinants of vaccine hesitancy were dominant, appearing in 28 of the 29 studies. Within this category, the communication and media environment emerged as the most influential determinant. Studies consistently showed that social media platforms and digital news environments amplify misinformation, conspiracy narratives, and emotionally charged content, distorting perceptions of vaccine safety and effectiveness. Platform affordances such as rapid dissemination, peer-to-peer sharing, and algorithmic amplification intensified exposure to misleading information.

Individual and group determinants were identified in nearly half of the studies and were closely linked to online information exposure. Knowledge gaps, fear of adverse events, and declining trust in health systems were frequently reinforced by contradictory or sensationalized vaccine content encountered on social media. Vaccine decision-making was often shaped by heuristic reasoning and social norms emerging within online communities.

The most frequently reported infodemic effect was reduced willingness to vaccinate, identified in more than three-quarters of the studies. This was followed by the erosion of trust in governments, science, and public health institutions, adverse impacts on health-related decision-making, and heightened social fear and anxiety. Several studies connected misinformation exposure to tangible public health consequences, including delayed vaccination, disrupted outbreak control, and increased morbidity and mortality.

Importantly, the review identified twelve infodemic effects not previously emphasized in research, such as damage to patient–healthcare professional relationships, silencing of pro-vaccine voices, stigmatization, and the promotion of political or ideological agendas. Additionally, three vaccine hesitancy determinants emerged as distinct in LMIC contexts: adverse events following immunization as an information-driven trigger, persistent vaccination information gaps, and negative spillover from past vaccine controversies amplified through digital media.

Regarding interventions, most proposed strategies targeted the macro level, emphasizing government-led communication policies, social media monitoring, and

transparency initiatives. Meso-level actions focused on community engagement, partnerships with religious and local leaders, and organizational communication strategies. Micro-level interventions emphasized healthcare professionals' roles as trusted communicators and the importance of digital and health literacy.

Future Work

Future research should further investigate how social media platforms, algorithms, and data infrastructures shape the production, circulation, and visibility of vaccine-related information in low- and middle-income countries. While the reviewed studies highlight the importance of communication and media environments, more granular analyses are needed to understand how platform-specific affordances—such as recommender systems, content moderation practices, and engagement metrics—contribute to the amplification of misinformation and the persistence of vaccine hesitancy.

Comparative studies across platforms commonly used in LMICs, including Facebook, WhatsApp, TikTok, and regionally dominant applications, would help clarify how different sociotechnical systems influence vaccine discourse, trust formation, and health-related decision-making. Research using computational social science methods, such as social network analysis, social listening, and sentiment analysis, could complement qualitative approaches by identifying patterns of misinformation spread, influential actors, and discursive dynamics within online communities.

Future work should also address the geographic and cultural gaps identified in the literature, particularly the underrepresentation of Africa and Latin America. Context-sensitive qualitative research—such as digital ethnography and in-depth interviews—would provide deeper insights into how individuals and communities interpret vaccine information, negotiate uncertainty, and integrate online content with offline social, religious, and institutional influences.

Another important avenue for research concerns platform governance and public policy. Studies should examine how collaborations between governments, public health institutions, civil society, and social media companies can be designed to counter infodemics while safeguarding freedom of expression and public trust. Evaluating the effectiveness of platform-based interventions, fact-checking initiatives, and algorithmic adjustments in LMIC contexts remains a critical research need.

Finally, future research should explore strategies for rebuilding trust in science and public institutions after periods of intense misinformation. This includes examining participatory communication models, the role of healthcare professionals and community leaders as trusted digital actors, and the potential of digital literacy initiatives to foster long-term resilience against health misinformation.

References

- Acharya, S., Aechtner, T., Venaik, S., & Dhir, S. (2024). Bibliometric analysis of vaccine hesitancy research from behavioural perspectives (2015–2022). *Journal of Risk Research*, 27(2), 238–253. <https://doi.org/10.1080/13669877.2024.2317318>
- Assarroudi, A., Heshmati Nabavi, F., Armat, M. R., Ebadi, A., & Vaismoradi, M. (2018). Directed qualitative content analysis: the description and elaboration of its underpinning methods and data analysis process. *Journal of Research in Nursing*, 23(1), 42–55. <https://doi.org/10.1177/1744987117741667>
- Beirão, G., Patrício, L., & Fisk, R. P. (2017). Value cocreation in service ecosystems: Investigating health care at the micro, meso, and macro levels. *Journal of Service Management*, 28(2), 227–249. <https://doi.org/10.1108/JOSM-11-2015-0357>
- Gonçalves, B. A., de Souza Amorim Matos, C. C., dos Santos Ferreira, J. V., Itagyba, R. F., Moço, V. R., & Couto, M. T. (2023). COVID-19 vaccine hesitancy in Latin America and Africa: a scoping review. *Cadernos de Saude Publica*, 39(8). <https://doi.org/10.1590/0102-311XEN041423>
- Immunization Agenda 2030 Partners. (2024). Immunization agenda 2030: A global strategy to leave no one behind. *Vaccine*, 42, S5–S14. <https://doi.org/10.1016/j.vaccine.2022.11.042>
- MacDonald, N. E., & SAGE Working Group on Vaccine Hesitancy. (2015). Vaccine hesitancy: Definition, scope and determinants. *Vaccine*, 33(34), 4161–4164. <https://doi.org/10.1016/j.vaccine.2015.04.036>
- Nha HONG, Q., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., Rousseau, M.-C., & Vedel, I. (2018). Mixed Methods Appraisal Tool (MMAT), version 2018. <https://mixedmethodsappraisaltoolpublic.pbworks.com/>
- Pan American Health Organization. (2020). Understanding the infodemic and misinformation in the fight against COVID-19. <https://iris.paho.org/handle/10665.2/52052>
- Simas, C., & Larson, H. J. (2021). Overcoming vaccine hesitancy in low-income and middle-income regions. In *Nature Reviews Disease Primers* (Vol. 7, Issue 1). Nature Research. <https://doi.org/10.1038/s41572-021-00279-w>
- World Health Organization (WHO). (2025). Table 1: Summary of WHO Position Papers- Recommendations for Routine Immunization. <http://www.who.int/immunization/documents/positionpapers>
-

1B.3 “The solution to your pains!”: How scam ads exploit health disinformation on Meta platforms in Brazil.

Nicole Sanhotene, Débora Salles (Federal University of Rio de Janeiro), Daphne Silva (Federal University of Alagoas), João Haddad, Marie Santini (Federal University of Rio de Janeiro)

Introduction

Previous research has pointed to multiple forms of disinformation in science and health that undermine trust in science and public policies, promote distorted information, myths about vaccines and treatments without proven scientific basis (Rochel de Camargo, 2024; Oliveira, 2020). This discourse frequently promises access to alleged hidden secrets, drawing on the fascination with the forbidden, conspiracy theories, and so-called miracle solutions to mobilize emotions (Andrejevic et al., 2025).

Scientific disinformation finds fertile ground in the digital advertising market, exploiting the lack of oversight mechanisms and the inaction of big tech companies (Santini et al., 2025). This environment fuels an industry centered on scam ads, defined as deceptive online ads designed to mislead users into fraudulent transactions or interactions, promoting counterfeit products and deceptive commercial practices through social engineering strategies that exploit users' trust for profit and cause physical, financial, and psychological harm (Di Domenico et al., 2022, Button et al., 2014; Kikerpill & Siibak, 2021).

Within this ecosystem, digital platforms operate simultaneously as essential infrastructures for the promotion, curation, and distribution of content and as actors that directly profit from this industry (Kim, 2024). Meta documents revealed by Reuters indicate that the company is aware that its platforms are widely used to disseminate scam ads. These ads reportedly account for approximately 10% of the company's revenue and expose users to an estimated 15 billion scam impressions per day worldwide (Horwitz, 2025).

This model relies on microtargeting tools that enable the delivery of customized advertisements based on consumers' psychosocial profiles, constructed through algorithmic systems designed to optimize these companies' business interests (Carah et al., 2024). In the case of health-related scams, such ads may be deliberately targeted at vulnerable individuals, generating profit at the expense of public health and exploiting users' so-called “weak points” (O’Neil, 2020).

Additionally, this market is sustained by guarantees of anonymity and a low likelihood of accountability (Santini et al., 2025; Kim, 2024). Moreover, the social media scam market remains difficult to measure and monitor due to platforms' lack of transparency (Santini et al., 2025). This opacity hinders oversight and the effective enforcement of

regulations, enabling big tech companies to profit while users remain vulnerable to sophisticated scam schemes.

Objectives/Research Questions

Building on research on the role of advertising infrastructures in health-related disinformation (Zeng et al., 2020; Nelson et al., 2020), this study systematically maps health-related scam ads on Meta platforms to show how social media are used and actively facilitate the disinformation industry online. By assembling empirical evidence, we seek to demonstrate the mechanisms of this large-scale industry that generates profits for both big tech companies and scammers.

To this end, we draw on a case study of Brazil, considering that it offers a valuable research context given the existence of health advertising regulations and platform accountability frameworks. At the same time, like other countries in the Global South, Brazil faces significant challenges related to regulatory enforcement and researchers' access to platform data (Santini et al., 2025). This combination makes the country particularly suitable for examining how scam advertising operates under conditions of limited oversight, highlighting the role of platforms in enabling this market and the challenges it poses for accountability. This leads to the following question:

RQ1: What is the prevalence of scams among health-related ads on Meta platforms in Brazil?

In addition to examining the prevalence of scam ads, this study seeks to systematically characterize this type of content. To do so, we build on prior discussions of health disinformation and snake oil ads, which identify a range of recurrent strategies, including the use of urgency cues to pressure consumers, appeals to a rhetoric of “secrecy” to attract attention, practices of impersonation and the imitation of legitimate website layouts, as well as the use of clickbait headlines and low-quality imagery (Andrejevic et al., 2025; Zeng et al., 2020). Thus, we seek to assemble evidence that addresses the question:

RQ2: What deceptive strategies are used in online health scams to persuade consumers?

Previous evidence shows that Meta not only is aware of illegal advertising content circulating on its platforms but has also shown a high degree of permissiveness toward advertisers promoting illegal content (Horwitz & Au-Yeung, 2025; Horwitz, 2025). Reports suggest that the company has intentionally pursued increased profits by charging scammers higher fees, while downplaying the severity of advertisements that expose users to significant risks. These dynamics ultimately motivate our final question.

RQ3: How do Meta's platform governance allow the circulation of online health scams?

Methods

To collect advertising data from Meta's platform ecosystem, we used the company's Ad Library user interface. In Brazil, this process is limited, as systematic data collection is available only for advertisements related to social issues, elections, or politics, while other ads are displayed only while they are active. To address this issue, we developed a web scraper to extract data on non-political ads from the Ad Library. Between April and May 2025, we collected approximately 169,000 health-related ads, from which we sampled 1,000 unique ads across different advertisers, resulting in 6,579 ad instances due to multiple ad versions.

To conduct a systematic content analysis of the advertisements, we developed a protocol consisting of 20 criteria based on their textual and visual content. The analysis sought to identify issues related to the promoted product and/or treatment, violations of health and advertising laws and regulations in Brazil, and the persuasive strategies employed by advertisers. We also examined ads in which advertisers attempted to impersonate health professionals or government institutions. Finally, to identify ads employing voice and image manipulation techniques, we considered signs of visual alteration, abrupt cuts and overlapping speech.

Two researchers independently annotated and coded the sample of 1,000 unique ads. To ensure the reliability of the results, in cases of disagreement, the ads were analyzed by a third researcher. The final classification of the ads was determined based on majority agreement among the coders. Inter-coder reliability was assessed using Krippendorff's alpha (Krippendorff, 2019), with values ranging from 0.70 to 0.90, considered satisfactory to highly reliable in the academic literature (Sampaio & Lycarião, 2021).

Results

We identified 4,997 (76%) fraudulent health-related ads promoted by 505 advertisers between 2022 and 2025. Of these, 4,282 ads (85,7%) redirected users to WhatsApp, providing clear evidence that scammers exploit the platform's infrastructure to carry out fraudulent schemes.

These ads contained multiple violations of existing health regulations and codes in Brazil (Brasil, 2008; CFM, 2023), including exaggerated claims about product effectiveness and outcomes, failure to identify medical professionals, and the dissemination of false scientific information. We also identified the promotion of 165 potentially fraudulent products, including unregistered products and supplements.

The analysis also revealed that 99,5% of the scam ads relied on narratives that appeal to fear and urgency while promising rapid cures, reinforcing emotional manipulation by exploiting audience vulnerabilities (O'Neil, 2020). Beyond the inherent gravity of illegal content dissemination, these findings reflect a broader dynamic that exploits human

vulnerabilities while playing on people's anxieties and trust in health institutions to profit from promises of "miraculous" cures, leveraging Meta's sophisticated microtargeting systems to maximize persuasive reach.

These strategies are part of commercial populism, characterized by the promotion of magical and easy solutions to complex problems (Andrejevic et al., 2025). In the ads, this discourse manifests through promises of miraculous cures and infallible treatments that are allegedly being deliberately concealed by governments or health institutions — portrayed as immoral actors who supposedly maintain their power by withholding "forbidden" knowledge and therapies from the public.

Of the total number of scam ads, 3,738 (74.8%) contained some form of audiovisual manipulation, including the misuse of images and testimonials from institutions, health specialists, as well as prominent figures from Brazilian journalism, entertainment, and politics to confer legitimacy to the advertised products. The same strategy was employed to make advertisers appear trustworthy. The presence of this practice among 322 advertisers suggests that Meta exercises little to no oversight over advertisers operating on its platforms.

Scam ads remained online for an average of 30 days. The ads with the longest circulation periods remained active from July 16, 2022 to at least August 26, 2025, meaning that users were exposed to scams for more than three years. It provides evidence of systemic failures in the company's services, raising questions about its effective commitment to protecting users.

Future Work

Although our study examined only a subset of the ads collected over a single month, the findings reported here already reveal critical evidence regarding both the content of the strategies used to deceive users and their systemic nature.

Future work should extend the present analysis by examining larger dataset and leveraging automated classification methods to better characterize scam dynamics and assess the scale of health-related scam advertising ecosystems online.

In addition, our study shows that the existence of norms and regulatory frameworks governing advertising and the operations of social media platforms is essential, as they provide the legal basis for identifying corporate negligence and characterizing platform "failures" as abusive practices against users. However, such frameworks can only ensure a safer online environment when accompanied by robust oversight and enforcement mechanisms. Future research may therefore evaluate the quality of the data made available by platforms in their public ad repositories, proposing quality standards that ensure greater transparency and the availability of high-quality information.

Finally, as this study focuses exclusively on Meta platforms, future research should broaden the scope to include other online advertising and content distribution platforms. Expanding this debate through analyses of additional platforms would help provide a more comprehensive understanding of the digital advertising landscape, recognizing that the online disinformation ecosystem extends well beyond Meta.

References

- Andrejevic, M., O'Neill, C., & Mahoney, I. (2025). "The scandal that shocked the world": Conspiratorship and online scam ads. *Journal of Information Technology & Politics*, 1–15. <https://doi.org/10.1080/19331681.2025.2453920>
- Brasil. (2008). Resolução-RDC Nº 96, de 17 de dezembro de 2008. Ministério da Saúde. https://bvsms.saude.gov.br/bvs/saudelegis/anvisa/2008/rdc0096_17_12_2008.html
- Button, M., Nicholls, C. M., Kerr, J., & Owen, R. (2014). Online frauds: Learning from victims why they fall for these scams. *Australian & New Zealand Journal of Criminology*, 47(3), 391–408. <https://doi.org/10.1177/0004865814521224>
- Carah, N., Hayden, L., Brown, M.-G., Angus, D., Brownbill, A., Hawker, K., Xue Ying, T., Dobson, A., & Robards, B. (2024). Observing "tuned" advertising on digital platforms [Info:eu-repo/semantics/article]. Alexander von Humboldt Institute for Internet and Society gGmbH. <https://doi.org/10.14763/2024.2.1779>
- CFM - Conselho Federal de Medicina. Resolução CFM Nº 2.336, de 13 de julho de 2023, Manual de Publicidade Médica (2023). Conselho Federal de Medicina. <https://www.flip3d.com.br/pub/cfm/index9/?numero=45&edicao=5578>
- Di Domenico, G., Nunan, D., & Pitardi, V. (2022). Marketplaces of Misinformation: A Study of How Vaccine Misinformation Is Legitimized on Social Media. *Journal of Public Policy & Marketing*, 41(4), 319–335. <https://doi.org/10.1177/07439156221103860>
- Horwitz, J. (2025, November 6). Meta is earning a fortune on a deluge of fraudulent ads, documents show. Reuters. <https://www.reuters.com/investigations/meta-is-earning-fortune-deluge-fraudulent-ads-documents-show-2025-11-06/>
- Horwitz, J., & Au-Yeung, A. (2025, 15 de maio). Meta battles an 'epidemic of scams' as criminals flood Instagram and Facebook. *The Wall Street Journal*. <https://www.wsj.com/tech/meta-fraud-facebook-instagram-813363c8>
- Kikerpill, K., & Siibak, A. (2021). Mazephishing: The COVID-19 pandemic as credible social context for social engineering attacks. *Advance*, 1. <https://doi.org/10.31124/advance.14607894.v1>

- Kim, H. Y. (2024). What's wrong with relying on targeted advertising? Targeting the business model of social media platforms. *Critical Review of International Social and Political Philosophy*, 1–21. <https://doi.org/10.1080/13698230.2024.2309047>
- Krippendorff, K. (2019). *Content Analysis: An Introduction to Its Methodology*. SAGE Publications, Inc. <https://methods.sagepub.com/book/mono/content-analysis-4e/toc>
- Nelson, M. R., Das, S., & Ahn, R. J. (2020). A Prescription for Health: (Pseudo)Scientific Advertising of Fruits and Vegetables in the Early 20th Century. *Advertising & Society Quarterly*, 21(1). <https://muse.jhu.edu/pub/21/article/752637>
- Oliveira, T. (2020). Desinformação científica em tempos de crise epistêmica: Circulação de teorias da conspiração nas plataformas de mídias sociais. *Fronteiras - estudos midiáticos*, 22(1), 21–35. <https://doi.org/10.4013/fem.2020.221.03>
- O'Neil, C. (2020). *Algoritmos de destruição em massa: Como o Big Data aumenta a desigualdade e ameaça a democracia*. Editora Rua do Sabão.
- Rochel de Camargo, K. (2024). Disputed expertise and chaotic disinformation: COVID-19 and denialist physicians in Brazil. *Transcultural Psychiatry*, 61(5), 714–723. <https://doi.org/10.1177/13634615231213835>
- Sampaio, R. C. & Lycarião, D. (2021). *Análise de conteúdo categorial: manual de aplicação*. Enap. <http://repositorio.enap.gov.br/jspui/handle/1/6542>
- Santini, M. R., Salles, D. G., Martins, B. M. M., Sanchotene, N., Belin, L. L., Moreira, A., Mendes, A., Dias, B., Barros, C. E., Mello, D. P., Silva, D., Graef, F., Ferreira, F., Haddad, J. G., Murakami, L., Canavarro, M., Borges, M., Gomes, M., Medeiros, P. M., Ciodaro, T., & Scortegagna, V. B. (2025). Atingidos pelas redes sociais: Os impactos da indústria da desinformação nos consumidores brasileiros. <https://doi.org/10.29327/5566995>
- Santini, R. M., Salles, D., Martins, B., Barros, C. E., Moreira, A., Haddad, J. G., Silva, D., & Seade, R. (2023, November 17). 'O fim dos seus problemas': A permanência de anúncios que usam o programa 'Desenrola Brasil' para golpes e fraudes nas plataformas Meta. *NetLab UFRJ*. <https://netlab.eco.ufrj.br/post/o-fim-dos-seus-problemas-a-perman%C3%Aancia-de-an%C3%Bancias-que-usam-o-programa-desenrola-brasil-para-go>
- Zeng, E., Kohno, T., Roesner, F., & Allen, P. G. (2020). *Bad News: Clickbait and Deceptive Ads on News and Misinformation Websites*. <https://www.semanticscholar.org/paper/Bad-News%3A-Clickbait-and-Deceptive-Ads-on-News-and-Zeng-Kohno/a7d7a8e1eca1dee63e871751dad4e0481079acb4>

1B.4 Combating Climate Misinformation via AI chatbots: How Interactivity and Lateral Reading Support Shape Trust, Perceived Response Credibility, and Further Verification Intentions.

Shirley S. Ho (Nanyang Technological University), Junru Huang (CNRS@CREATE), Chang He (Nanyang Technological University)

Introduction

Climate change misinformation poses a significant threat to public understanding and policy support by eroding trust in scientific consensus and weakening motivation for pro-environmental behaviour (Cook et al., 2017). Although strategies such as expert rebuttals and fact-checking have been used to combat misinformation (Vraga & Bode, 2017; Walter et al., 2020), traditional correction methods often struggle to sustain users' attention or promote deeper cognitive engagement (Ecker, 2017; Bautista et al., 2024). Static rebuttals may be dismissed, ignored, or quickly forgotten, especially in polarized online environments where misinformation circulates within ideologically homogeneous networks (Del Vicario et al., 2016). These limitations heighten the need for more engaging and scalable correction approaches.

In recent years, Artificial Intelligence (AI) has been increasingly deployed to identify and screen suspicious content due to its considerable capability in classifying and retrieving information across large-scale textual and multimodal datasets (Hu et al., 2022; Khraisat et al., 2025; Thorne & Vlachos, 2018). As AI technologies have matured, their applications have expanded from general language processing tasks to more user-facing tools such as chatbots (Lee et al., 2024; Lee & Fussell, 2025). In this context, chatbots have begun to be applied to the task of correcting misinformation, such as the World Health Organization's chatbot used to fight COVID-19 misinformation (World Health Organisation, 2020) and Taiwan's fact-checking bot inside the LINE messaging app (Taiwan Insight, 2022). Unlike static rebuttals, AI chatbots offer a user-facing interface that enables real-time, interactive, and individualized delivery of corrective information, offering great potential to enhance the effectiveness of misinformation correction.

At the same time, research in media literacy identifies lateral reading - verifying information through cross-source consultation - as one of the most effective strategies for improving misinformation resilience (Brodsky et al., 2021; Wineburg & McGrew, 2019). While both chatbot interactivity and the media strategy of lateral reading have each demonstrated promise, little empirical research has examined how these features

can be integrated within AI-mediated systems to shape users' trust, credibility perceptions, and behavioral intentions in the context of climate misinformation. By experimentally testing the combined effects of chatbot interactivity and lateral reading support, this study aims to advance theoretical understanding of how technological affordances and epistemic scaffolding jointly influence AI-mediated misinformation correction.

Objectives/Research Questions

While prior research has demonstrated the technical capacity of AI systems in misinformation detection (Hu et al., 2022; Thorne & Vlachos, 2018), less attention has examined the effects of user-facing design affordances, particularly when integrated with media literacy strategies. To fill this gap, we study how two chatbot features—interactivity and lateral reading support—shape users' psychological responses and behavioral intentions in climate misinformation correction.

According to the MAIN model, interactivity cues signal technological agency and foster cognitive engagement and trust (Sundar, 2008). Conversational and responsiveness cues shape perceived humanness, trust, and credibility judgement (Ding & Najaf, 2024; Følstad et al., 2018; Go & Sundar, 2019; Metzger et al., 2010). In misinformation contexts, interactive features may also reduce verification effort and encourage deeper elaboration (Liu & Shrum, 2002). Accordingly, we hypothesize:

H1: Compared to low interactivity, high interactivity of a misinformation-correcting AI chatbot will be associated with higher (a) trust, (b) perceived message credibility, (c) intention to reuse, (d) intention to share corrected information, and (e) intention to conduct further verification.

Grounded in Epistemic Cognition Theory (Hofer & Pintrich, 1997), lateral reading is a media literacy strategy that encourage individuals to verify claims by consulting multiple credible sources (Wineburg & McGrew, 2019). Such interventions have been proven effective in improve misinformation resilience (Brodsky et al., 2021). Integrating lateral reading support into chatbots - through prompts or links to authoritative sources - may enhance transparency and thus trust (Xu et al., 2023). When users perceive effective verification support, perceived usefulness may increase, leading to greater reuse and information sharing (Bhattacharjee, 2001; Pennycook & Rand, 2022). Thus, we hypothesize:

H2: Compared to a no-support condition, a misinformation-correcting AI chatbot with lateral reading support will be associated with higher (a) trust, (b) perceived message credibility, (c) intention to reuse, (d) intention to share corrected information, and (e) intention to conduct further verification.

Finally, interactivity and lateral reading may operate synergistically. Interactivity sustains engagement (Sundar, 2008), while lateral reading support directs such

engagement toward cross-source verification (Wineburg & McGrew, 2019; Artmann et al., 2023). When combined, interactivity may amplify the impact of lateral reading by increasing users' trust and willingness to adopt verification practices. Therefore, we hypothesize:

H3: Chatbot interactivity and lateral reading will interact such that interactivity intensifies the effects of lateral reading on participants' (a) trust, (b) perceived message credibility, (c) intention to reuse, (d) intention to share corrected information, and (e) intention to conduct further verification.

Methods

A 2 (chatbot interactivity: high v.s. low) × 2 (lateral reading support: present v.s. absent) between-subjects online experiment will be conducted to test the hypotheses. Approximately 400 adult participants will be recruited, with sample size determined using G*Power based on a small-to-medium effect size.

Pilot Study

Before stimuli making, a pilot study (N=105) was conducted in Singapore. Participants aged 21 and above evaluated 24 climate-related statements, half accurate and half false, rating their familiarity and judging whether each statement was true or false. Finally, four statements (two false and two accurate) with the lowest familiarity and lowest correct response rates were selected to make the stimuli for the main experiment to ensure moderate difficulty and minimize the influence of prior knowledge.

Stimuli and procedure

Using the four selected statements, experimental stimuli were developed for all four conditions. The stimuli consist of a web-based AI chatbot interface designed to correct climate misinformation (see Figure 1). The left panel displays task instructions and the four statements to be evaluated. The right panel contains a conversational chatbot window for interaction. Participants are instructed to use the chatbot to find out whether each statement is true or false. A “proceed” button appears only after all four statements have been fact-checked, ensuring full task engagement before advancing to the survey.

In the high-interactivity condition, participants engage in a multi-turn conversation where the chatbot provides personalized responses, follow-up prompts, and clickable options for additional clarification. In the low-interactivity condition, participants receive static, one-time responses without opportunities for further dialogue or feedback. In the latera-reading-support-present condition, the chatbot provides suggestions for external verification, including links to credible fact-checking or scientific sources. In the support-absent condition, no such prompts or external

references is offered. Across all conditions, the factual content and tone of the chatbot's responses remain consistent to avoid confounds.

Measurement

Trust in chatbot is measured by 12 items adapted from Ho et al. (2024), which measured individuals' trust in AI systems from the dimensions of performance, reliability, and purpose. Perceived message credibility is measured by four items adapted from Besalú et al. (2021). Intention to use again will be measured by four items adapted from Song et al. (2021). Intention to share corrected information is measured by three items adapted from Song et al. (2021) and Majerczak & Strzelecki (2022). Intention to conduct further verification is measured by four self-developed items based on previous literature on misinformation correction (Majerczak & Strzelecki, 2022; Pennycook & Rand, 2021), such as "If I come across content the chatbot flagged, I will thoroughly research its accuracy."

Results

The study has received approval from the authors' university's institutional review board. A pilot study has been completed, and the four experimental stimuli are being finalized based on the pilot findings.

Data collection for the main experiment is currently in progress. Upon completion, ANOVA analyses will be conducted to examine the main and interaction effects proposed in the hypotheses. Based on prior theoretical and empirical work (Sundar, 2008; Wineburg & McGrew, 2019), we expected positive main effects of both interactivity and lateral reading support, as well as a significant interaction effect such that interactivity amplifies the impact of lateral reading support on the outcome variables.

Full statistical results will be reported upon completion of data collection, expected by February 2026.

Future Work

Future research will extend this work in two main directions. First, we plan to examine the generalizability of the findings across different types of misinformation. Beyond climate-related information, future studies will explore similar effects in domains such as political misinformation and other socially contentious issues. We also aim to expand the scope beyond text-based misinformation to include multimodal formats, such as image-based or video-based misinformation, which are increasingly prevalent on social media.

Second, subsequent studies may be designed based on the empirical results of the current experiment. Depending on the observed mechanisms, future work may explore additional design features, potential mediators, and boundary conditions to refine AI-mediated misinformation correction strategies.

References

- Bautista, J. R., Zhang, Y., & Gwizdka, J. (2024). Correcting vaccine misinformation on social media: Effect of social correction methods on vaccine skeptics' intention to take COVID-19 vaccine. *New Media & Society*, 26(12), 7412–7438.
<https://doi.org/10.1177/14614448231169697>
- Besalú, R., Pont-Sorribes, C., & Martí, A. (2021). Perceived credibility of tweets by opinion leaders during the COVID-19 pandemic in Spain. *International Journal of Communication*, 15(2021), 5158–5185.
- Bhattacharjee, A. (2001). Understanding Information Systems Continuance: An Expectation-Confirmation Model. *MIS Quarterly*, 25(3), 351.
<https://doi.org/10.2307/3250921>
- Brodsky, J. E., Brooks, P. J., Scimeca, D., Todorova, R., Galati, P., Batson, M., Grosso, R., Matthews, M., Miller, V., & Caulfield, M. (2021). Improving college students' fact-checking strategies through lateral reading instruction in a general education civics course. *Cognitive Research: Principles and Implications*, 6(1), 23.
<https://doi.org/10.1186/s41235-021-00291-4>
- Cook, J., Lewandowsky, S., & Ecker, U. K. H. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PLOS ONE*, 12(5), e0175799.
<https://doi.org/10.1371/journal.pone.0175799>
- Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3), 554–559.
<https://doi.org/10.1073/pnas.1517441113>
- Ding, Y., & Najaf, M. (2024). Interactivity, humanness, and trust: A psychological approach to AI chatbot adoption in e-commerce. *BMC Psychology*, 12(1), 595.
<https://doi.org/10.1186/s40359-024-02083-z>
- Ecker, U. K. H. (2017). Why rebuttals may not work: The psychology of misinformation. *Media Asia*, 44(2), 79–87. <https://doi.org/10.1080/01296612.2017.1384145>
- Følstad, A., Nordheim, C. B., & Bjørkli, C. A. (2018). What Makes Users Trust a Chatbot for Customer Service? An Exploratory Interview Study. 194-208.
https://doi.org/10.1007/978-3-030-01437-7_16
- Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, 97, 304–316. <https://doi.org/10.1016/j.chb.2019.01.020>

- Ho, S. S., & Cheung, J. C. (2024). Trust in artificial intelligence, trust in engineers, and news media: Factors shaping public perceptions of autonomous drones through UTAUT2. *Technology in Society*, 77, 102533.
- Hofer, B. K., & Pintrich, P. R. (1997). The Development of Epistemological Theories: Beliefs About Knowledge and Knowing and Their Relation to Learning. *Review of Educational Research*, 67(1), 88–140.
- Hu, L., Wei, S., Zhao, Z., & Wu, B. (2022). Deep learning for fake news detection: A comprehensive survey. *AI Open*, 3, 133–155.
<https://doi.org/10.1016/j.aiopen.2022.09.001>
- Khraisat, A., Manisha, Chang, L., & Abawajy, J. (2025). Survey on Deep Learning for Misinformation Detection: Adapting to Recent Events, Multilingual Challenges, and Future Visions. *Social Science Computer Review*, 08944393251315910.
<https://doi.org/10.1177/08944393251315910>
- Lee, S., Park, Y., & Park, G. (2024). Using AI chatbots in climate change mitigation: A moderated serial mediation model. *Behaviour & Information Technology*, 43(16), 4016–4032.
- Lee, T.-H., & Fussell, S. R. (2025). Countering Misinformation in Private Messaging Groups: Insights From a Fact-checking Chatbot. *Proceedings of the ACM on Human-Computer Interaction*, 9(1), 1–30. <https://doi.org/10.1145/3701189>
- Liu, Y., & Shrum, L. J. (2002). What is Interactivity and is it Always Such a Good Thing? Implications of Definition, Person, and Situation for the Influence of Interactivity on Advertising Effectiveness. *Journal of Advertising*, 31(4), 53–64.
<https://doi.org/10.1080/00913367.2002.10673685>
- Majerczak, P., & Strzelecki, A. (2022). Trust, media credibility, social ties, and the intention to share towards information verification in an age of fake news. *Behavioral Sciences*, 12(2), 51.
- Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, 60(3), 413–439.
<https://doi.org/10.1111/j.1460-2466.2010.01488.x>
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388–402.
- Pennycook, G., & Rand, D. G. (2022). Accuracy prompts are a replicable and generalizable approach for reducing the spread of misinformation. *Nature Communications*, 13(1), 2333. <https://doi.org/10.1038/s41467-022-30073-5>
- Song, S., Zhao, Y. C., Yao, X., Ba, Z., & Zhu, Q. (2021). Short video apps as a health information source: An investigation of affordances, user experience and users' intention to continue the use of TikTok. *Internet Research*, 31(6), 2120–2142.

- Sundar, S. S. (2008). The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility.
- Taiwan Insight. (2022, October 12). The Bot Fighting Disinformation: The Story of Cofacts. <https://taiwaninsight.org/2022/10/12/the-bot-fighting-disinformation-the-story-of-cofacts/>
- Thorne, J., & Vlachos, A. (2018). Automated Fact Checking: Task formulations, methods and future directions (No. arXiv:1806.07687).
- Vraga, E. K., & Bode, L. (2017). Using Expert Sources to Correct Health Misinformation in Social Media. *Science Communication*, 39(5), 621–645. <https://doi.org/10.1177/1075547017731776>
- Walter, N., Cohen, J., Holbert, R. L., & Morag, Y. (2020). Fact-Checking: A Meta-Analysis of What Works and for Whom. *Political Communication*, 37(3), 350–375. <https://doi.org/10.1080/10584609.2019.1668894>
- Wineburg, S., & McGrew, S. (2019). Lateral Reading and the Nature of Expertise: Reading Less and Learning More When Evaluating Digital Information. *Teachers College Record*, 121(11), 1–40. <https://doi.org/10.1177/016146811912101102>
- World Health Organization. (2020). WHO and Rakuten Viber fight COVID-19 misinformation with interactive chatbot. <https://www.who.int/news-room/feature-stories/detail/who-and-rakuten-viber-fight-covid-19-misinformation-with-interactive-chatbot>
- Xu, Y., Bradford, N., & Garg, R. (2023). Transparency Enhances Positive Perceptions of Social Artificial Intelligence. *Human Behavior and Emerging Technologies*, 2023(1), 5550418. <https://doi.org/10.1155/2023/5550418>
-

Session 1C - Social Media Cultures and Everyday Life (1)

Day 1 · Mon 13 July · 9:30–11:00 · Seminar Suite C, ARC

1C.1 “Meaningful” or “meaningless” social interactions? The automation of commenting in live streams on Douyin.

Andreas Schellewald (Goldsmiths, University of London), Hui Lin (King's College London)

Introduction

From a political economy perspective, the central function of media platforms lies in their capacity to attract and retain user attention in order to generate value (Smythe, 1981, Webster 2016). However, recent scholarship suggests that this logic has shifted from an attention economy toward an engagement economy dynamism, in which users are increasingly encouraged to interact, respond, and participate in more continuous and intimate ways (Ørmen and Gregersen 2022). Engagement becomes thus a key source of value for platforms, and communicative practices are increasingly shaped by platform design to fit commercial objectives, such as driving e-commerce related actions (Su, 2023). This paper investigates this shift through the case of Douyin and live stream commenting practices.

Douyin and its international version TikTok are amongst the most popular social media platforms globally. Live streaming is becoming an increasingly central feature of their platform ecosystems. The reason behind this is that live streaming has emerged as an important site for platform monetisation (Taylor, 2018), combining real-time interaction, social presence, and commercial exchange (Zhao et al., 2023) with the growing appeal of creator-driven content and promotion (Cunningham & Craig, 2019). Douyin in particular has been described as standing out through its intensified commercialisation, rapidly evolving its features to enable commerce and monetisation on the platform (Su, 2023).

Live streaming is a productive for examining these transformations, as it foregrounds real-time interaction and audience participation as a form of “labour” and “work” (Carter and Egliston, 2021; Meisner & Ledbetter, 2020; Taoyue, 2025). Within live streams, features such as live commenting play a central role in organising engagement

and structuring viewer interaction. By defining the very content and form of communal exchange, such commercially oriented platform features translate social interaction into promotional speech and monetisable activity. What is at stake, therefore, is not only the economic extraction of engagement, but the broader social consequences of how interpersonal communication contents and online socialities are reshaped within highly commercialised social media cultures.

Objectives/Research Questions

This study qualitatively examines why users engage with automatic live commenting on Douyin and explores its sociocultural impact on both users and live-streaming communities. Rather than focusing solely on the political economic dimensions herein, the article investigates how commercial-oriented live commenting features shape interpersonal communication and influence users' freedom to engage in open discussion on the platform. Taking a bottom-up point of view, the study examines how the commercialisation of live chats affects everyday social interaction, alters communicative norms, and reshapes users' experiences of social presence in live-streaming environments. By focusing on users' perspectives, the article seeks to understand how engagement-driven platform features are interpreted, negotiated, and enacted in practice. By doing so, the article moves beyond existing discussions of digital labour to examine users' real-time sociality and chatting behaviours as they unfold within a commercialised media environment. The central research concern is how live commenting, as a form of structured engagement, transforms the qualities of online social interaction from the perspective of participating users.

Methods

This paper follows a qualitative and ethnographic methodology to examine participation in automatic live commenting. Ethnographic observation and diary-interview methods with 21 Douyin users form the primary body of data, more specifically. Ethnographic observations were made over a four-month period. This allowed for sustained engagement with live-streaming practices on Douyin, features of the platform, and patterns of participation as they unfolded over time. Focusing on how users participated in live streams and how automatic comments were circulated in real time, live streams across different categories were observed. Field notes were taken during observation, and screen recordings were collected to document patterns of live commenting and audience interaction as they unfolded on the platform.

Alongside ethnographic observation, diary-interviews were used to capture participants' reflections on their experiences of live commenting. This included motivations, perceptions, and affective responses. Participants completed a seven-day diary by recording their live-stream viewing and commenting practices. This method enabled participants to document their interactions and feelings in situ, providing insight into how engagement practices were experienced in their respective contexts.

Thereafter, follow-up interviews were conducted. As the majority of participants preferred anonymity, all interviews were conducted one-on-one via telephone in informal settings and were audio-recorded and transcribed.

Following a grounded theory approach, ethnographic field notes, diary entries, and interview transcripts were analysed through an iterative coding process. This allowed themes to emerge inductively from the bottom-up perspective of participants' accounts and observed user practices. Attention was paid to how automatic live commenting was encountered, interpreted, and enacted within live-streaming communities, and how participation was shaped by platform affordances and engagement-oriented interactional norms. Analytical categories were developed through constant comparison between observational materials and participants' diary and interview accounts. This enabled the study to connect routine practices of live commenting with users' reported motivations and perceptions. The combined strategy supported here a contextualised understanding of live commenting as it unfolded in real time, and it captured retrospective reflections on the meanings and social consequences of their engagement.

Results

The paper will show that automatic live commenting streamlines participation by encouraging simplified and low-effort engagement. Participants described live commenting as requiring minimal cognitive or emotional investment, framing it, for example, as mere "clicking" or "tapping". Automation thus lowered barriers to entry for engagement, increasing the volume and speed of interaction within live stream chats. Data showed that most participants engaged with promotional live commenting even when they had limited interest in the promoted content. Engagement instead seemed shaped by interface design, algorithmic recommendation systems, and creator-led strategies that explicitly encouraged participation. Over time, promotional commenting became therefore normalised as a routine aspect of live-stream viewing amongst the participants.

Participants demonstrated awareness of the commercial logic underpinning live commenting. Many recognised that their interactions contributed to metrics of streamers and platforms. However, this awareness rarely led to resistance. Instead, participants expressed ambivalence, citing strategic participation to gain rewards, social pressure to conform to chat norms, and fear of missing out on potential benefits. The data further revealed that automatic live commenting blurred boundaries between social interaction and digital labour. While participants did not always describe their activity as "work," they framed engagement in instrumental terms. They emphasised efficiency, rewards, and outcomes rather than meaningful connection. As a result, participation became increasingly oriented toward platform-defined goals rather than communal meaning-making.

Participants repeatedly reported that live chats felt crowded, repetitive, and lacking meaningful exchange. The dominance of promotional comments limited opportunities for spontaneous conversation, making it difficult to sustain dialogue, or form social connections. Many described chat rooms as noisy and overwhelming, leading to detachment from other viewers. Overall, findings demonstrate that commercialisation of live commenting reshapes online sociality by structuring interaction around visibility, metrics, and incentives rather than reciprocity or dialogue. On Douyin, live commenting produces a form of sociality that is commercially oriented, prioritising engagement serving platform and creator interests. This transformation has broader implications for participation and agency. By predefining the forms and content of interaction, live commenting translates social communication into promotional speech, limiting open-ended expression and user-led meaning-making. Communicative agency becomes shaped by platform logics, highlighting how the engagement economy reshapes not only economic relations but also the content of online interaction (Ørmen and Gregersen 2022).

Future Work

By defining the forms of audience interaction and translating the content of social communication into promotional speech, this study highlights an ongoing shift toward an engagement economy within highly commercialised social media landscapes (Ørmen and Gregersen 2022). These findings suggest that greater attention is needed to understand how users negotiate, contest, and adapt to these engagement-driven conditions, including through creative and playful practices (Siles et al. 2023). Beyond existing discussions of digital labour, future research should further examine how the social dimensions of live streaming are affected by the increasing commercialisation of interaction itself. In particular, questions remain regarding the longer-term consequences of promotionally programmed communication shaping, and potentially dominating, users' online social experiences on platforms such as Douyin.

References

- Carter, M., & Egliston, B. (2021). The work of watching Twitch: Audience labour in livestreaming and esports. *Journal of Gaming & Virtual Worlds*, 13(1), 3–20. https://doi.org/10.1386/jgvw_00025_1
- Cunningham, S., & Craig, D. (2019). *Social media entertainment: The new intersection of Hollywood and Silicon Valley*. New York University Press.
- Meisner, C., & Ledbetter, A. M. (2022). Participatory branding on social media: The affordances of live streaming for creative labor. *New Media & Society*, 24(5), 1179–1195. <https://doi.org/10.1177/1461444820972392>

- Ørmen, J., & Gregersen, A. (2023). Towards the engagement economy: Interconnected processes of commodification on YouTube. *Media, Culture & Society*, 45(2), 225–245. <https://doi.org/10.1177/01634437221111951>
- Siles, I., Gómez-Cruz, E., & Ricaurte, P. (2023). Toward a popular theory of algorithms. *Popular Communication*, 21(1), 57–70. <https://doi.org/10.1080/15405702.2022.2103140>
- Su, C. (2024). *Douyin, TikTok, and China's online screen industry: The rise of short-video platforms*. Routledge, Taylor & Francis Group.
- Taoyue, W. (2025). Watching as working. In X. Xu, M. Zhao, & M. Y. Wang, *Chinese Social Media I* (1st ed., pp. 101–116). Routledge. <https://doi.org/10.4324/9781003596493-10>
- Taylor, T. L. (2018). *Watch me play: Twitch and the rise of game live streaming*. Princeton University Press.
- Smythe, D. (1981). On the audience commodity and its work. In M.G. Durham and D.M. Kellner (Eds.), *Media and cultural studies: Key works* (pp. 230-256). Blackwell.
- Webster, J. G. (2014). *The marketplace of attention: How audiences take shape in a digital age*. The MIT Press. <https://doi.org/10.7551/mitpress/9892.001.0001>
-

1C.2 Mediating ‘my London’ on social media: Knowledge migrants’ urban place experiences and perceptions through digital place memories.

Chenfei LI (King's College London)

Introduction

In the background of growing global student mobility, overseas Chinese students have become a significant population in the UK higher education, where they represent a major part of transnational knowledge migrants (Brooks and Waters, 2021; The Higher Education Statistics Agency, 2025). Despite their substantial academic and economic contributions to the host country, these students still often find themselves in a state of suspension – grappling with cultural barriers, psychological marginalisation, and a perceived lack of connection to their host cities. As a quintessential global metropolis and a world-known study destination, London provides a rich tapestry of urban experiences, yet it also amplifies the challenges these students face in cultivating a sense of place and belonging.

In order to relieve such hardships, existing studies typically put an emphasis on social media to ease such a cultural integration hardship merely as a functional tool or instant messaging for maintaining social networks (Forbush & Foucault-Welles, 2016; Guo, 2024), this study argues that its role is far more profound in self-driven memory- and meaning-making practices. Social media has evolved into a platform that mediates memories that are bricks composing knowledge of subsequent life-course for individuals. By capturing, editing, and selectively narrating urban encounters in London on social media, overseas students construct digital place memories. Through the memories, students negotiate their place perceptions, and ultimately are enabled to meaningful home-making with the host city in transnational contexts.

Objectives/Research Questions

Focusing on London as a primary site of inquiry, this study addresses both empirical and theoretical gaps by exploring how mainland Chinese international students perceive their urban experiences and dynamically construct place perceptions (specifically place identity and place attachment) through the creation of digital place memories. Grounded in a constructivist phenomenological lens (Harvey, 1996; Relph, 1976), this research facilitates interdisciplinary dialogues between urban communication, memory studies, platform culture, and migration studies. By employing an original ‘experience-memory-imagination’ cyclical framework, the study investigates:

How do overseas Chinese students in London construct and circulate their place memories through social media practices?

2. How does this memory-making contribute to negotiating their place identity and place attachment?

Ultimately, this work highlights the role of social media as a reflexive mediator of evolving person-place relationships, offering critical insights into how individuals creatively document, interpret, and share their urban encounters within transnational networks.

Methods

Digital photographs of London shared on social media are conceptualised in this study as digital place memories. Methodologically, the research employs a mixed-methods design comprising online questionnaires followed by in-depth, one-to-one photo-elicitation interviews. The target population consists of full-time mainland Chinese students in London (aged 18–34). From the questionnaire respondents who expressed interest, a purposive sample of 30 interviewees was recruited. To ensure a comprehensive representation, the sample covers a diverse range of demographics, including gender, institutional affiliation (e.g., KCL, UCL), academic discipline (ranging from arts and humanities to medical sciences), and level of study (undergraduate to PhD). Participants also vary in their residential distribution across London boroughs and their duration of stay (from under six months to over three years).

The questionnaire serves to gather data on demographics, urban experiences, and social media habits, utilising Likert-scale items to measure place identity and attachment from both offline and mediated perspectives. Subsequently, the photo-elicitation interviews facilitate a qualitative exploration of how participants' embodied place memories, socio-cultural meaning-making, and platform affordances intersect to shape the organisation of their digital memory content.

Results

1. Multi-platform orchestration and affordances of memories

The research identifies a multi-platform orchestration of memory. Interviewees navigate different digital affordances in social media to curate their urban experiences: WeChat functions as a 'digital archive' for strong-tie and mixed networks, maintaining a stable image for family, friends and acquaintances from versatile scenarios. Conversely, RedNote (XHS) serves an aesthetic stage where students perform an 'exquisite' London life to counteract feelings of daily lives or 'safety anxiety' (e.g. fear of crime) yet are more tolerant to accommodate negative experiences. Instagram acts as a 'social glue', where the '@' function in Stories validates friendships and anchors the self within the local social landscape.

Memory construction exhibits a significant temporal-spatial overlap. Students engage in this ‘creative writing’ process – thematically linking past unreleased photos of place with present ones and reflections to create a coherent narrative of their journey. This memory curation practice serves as not only a psychological reward to make sense of their life but also a subjective reclaim of urban places through digital traces. These curated place memories can bridge the gap between the ‘pre-conceived urban imaginary’ (built via social media before arrival) and the lived reality of the city. By ‘filling in’ the missing pieces of their experience with delayed posts, students actively construct a sense of ‘completeness’ in their digital life.

2. Digital rhythms of belonging: Identity validation and affective adaptation

Digital memory-sharing is a core process in forging place identity. It is perceived as making the city ‘part of the self’, where place identity is validated better through the ‘eyes of others’. Without these digital memory-sharing, interviewees expressed a lack of substantive connection to the sites. In the meanwhile, the sense of ‘home’ is a product of mobility and linguistic mediation. Stronger place attachment often crystallises only through mobility – returning to London after travels can trigger a transition from viewing a residence as a mere ‘apartment’ to a ‘home’, which also informs the Londoner identity. For those feeling culturally marginalised, place attachment is often anchored in institutional and interpersonal identification (e.g. societies and student union, and occasionally community networks) rather than the city at large. Visual representation of the body also serves as a mirror of affective adaptation: facial expressions and body language inform self-cognition. Finally, linguistic mediation plays a crucial role: even on Chinese platforms interviewees found that switching to English facilitates them to express certain topics such as love and gratitude more naturally, overcoming affective constraints often felt in the native Chinese social contexts.

Future Work

While this study offers insights into how overseas Chinese students in London negotiate their person-place relationships through memories on social media, several limitations might point to fruitful directions. The findings, drawn from a modest sample, may not be fully represent the diverse demographics and disciplinary backgrounds within the wider student population. In the meanwhile, there is a gendered usage gap between female and male students as social media users which may suggest further engagement with the phenomena of such.

Further research could be expanded to include longitudinal tracking studies. By comparing how students reinterpret their London memories across different life stages, we may better understand whether digital place perceptions dissolves with physical distance or becomes further romanticised through the lens of digital nostalgia.

There is also a need to deepen the focus on digital exclusion and inequality. Future studies should investigate the silent populations – those who lack a voice on social media or actively refuse to participate in digital display. By analysing these absences in digital memory construction, research may reveal how differences in digital literacy or activeness can impact affective belonging and identity negotiation. This will not only refine digital (connected) memory theory but also provide policy insights for enhancing the social well-being of the transnational mobile populations.

References

- Brooks, R., & Waters, J. (2021). International students and alternative visions of diaspora. *British Journal of Education Studies*, 69(5), 557-577. doi: 10.1080/00071005.2021.1948501.
- Forbush, E., & Foucault-Welles, B. (2016). Social media use and adaptation among Chinese students beginning to study in the United States. *International Journal of Intercultural Relations*, 50, 1-12. <https://doi.org/10.1016/j.ijintrel.2015.10.007>.
- Guo, W. (2024). Navigating cultural integration: The role of social media among Chinese students in the UK. *Journal of the Knowledge Economy*, 16, 2751-2782. <https://doi.org/10.1007/s13132-024-02111-8>.
- Harvey, D. (1996). *Justice, nature and the geography of difference*. Blackwell Publishers.
- Relph, E. (1976). *Place and placelessness*. SAGE.
- The Higher Education Statistics Agency (HESA). (03 April, 2025). Where do HE students come from? The Higher Education Statistics Agency. <https://www.hesa.ac.uk/data-and-analysis/students/where-from#main-content>
-

1C.3 From Feed to Floor: AI-Powered Social Media as a Sensory Buffer in Neurodivergent Retail Experiences.

Deep Sanghvi, Khushali Shah (Worcester Polytechnic Institute)

Introduction

Neurodivergent individuals, such as those with autism spectrum disorder (ASD), attention-deficit/hyperactivity disorder (ADHD), or sensory processing differences, often experience significant cognitive and sensory challenges in physical retail environments. Elements such as fluorescent lighting, crowded aisles, fabric textures, and unpredictable noise can trigger sensory overload and distress, limiting participation in routine shopping (Doucé & Adams, 2020; Finnigan, 2024; Gopan, 2025). Sensory discomfort is especially pronounced in apparel contexts. Kyriacou et al. (2023) found that autistic adults frequently experience fabrics, seams, and temperature as painful or distracting, elevating stress and reducing concentration. While some neurotypical consumers may view these sensations as minor inconveniences, other neurodivergent consumers often report heightened anxiety and cognitive fatigue.

Despite growing attention to accessibility, retail design still lacks comprehensive sensory-inclusive strategies. Mobility-related improvements have advanced, yet sensory needs remain overlooked (Bonetti et al., 2023; Gopan, 2025), reinforcing inequitable access and contributing to continued reliance on e-commerce among sensory-disabled consumers. Artificial intelligence (AI) offers emerging opportunities to ease these challenges. Retailers already deploy AI for personalization and experience enhancement (Yrjölä et al., 2024; Weber & Schütte, 2019), and tools such as adaptive recommendations, emotion-aware interfaces, and guided navigation may reduce sensory load. However, AI must evolve responsibly; poorly designed systems can amplify harm, as manipulative “dark patterns” disproportionately disadvantage consumers with attention or reading differences (Ocean et al., 2025).

While AI’s potential in sensory-inclusive retail is promising, it remains underexplored. Significant gaps persist in integrating sensory design, cognitive support, and adaptive digital tools for individuals. This research addresses these gaps by examining the sensory and cognitive challenges neurodivergent consumers face in physical retail environments and investigating how AI-driven interventions may reduce or exacerbate these burdens.

Objectives/Research Questions

Previous studies have demonstrated how a rich sensory retail environment can create sensory overload, resulting in increased cognitive load and emotional strain, and a reduction in customer engagement (Doucé & Adams, 2020; Sweller, 2011). When various stimuli are at play simultaneously, consumers must allocate cognitive

resources towards managing discomfort, maintaining concentration, monitoring emotional labor. Due to sensory threshold differences neurodivergent consumers may experience additional challenges to their shopping experience (Finnigan, 2024; Gopan, 2025; Kyriacou et al., 2023). Researchers are increasingly finding that AI systems affect the way information is processed and how emotions are managed. Personalization and decision-support technology driven by AI (like social media) creates less complex sets of information, filters out irrelevant information, and eases cognitive demands for users across many consumer applications (Bonetti et al., 2023; Weber & Schütte, 2019; Yrjölä et al., 2024).

For example, the algorithm-driven process of organizing social media content by user preferences and behaviours can shape an individual's emotional responses and engagement patterns, making social media a more complex set of algorithmically structured experiences rather than a straightforward or passive means of communication (Beyari & Hashem, 2025; Puntoni et al., 2021). Much research has been conducted on accessibility regarding in-store environment design involving AI through social networking sites. Few studies have examined how artificial intelligence mediates the relationship between consumers' physical retail environments and their online social media experiences. The present study addresses this gap by connecting cognitive load theory and sensory processing perspectives to understand the impact of AI-powered social media as an additional layer of algorithmic structure that impacts the cognitive assessment of sensory stimuli and subsequent emotional experience in brick-and-mortar retail spaces.

Hence, the purpose of this study is to (1) assess the impacts of sensory stimuli overload on cognitive effort and emotional strain on consumers within a physical retail environment; (2) determine if AI-powered social media tools are able to alleviate these effects of sensory overload by decreasing perceived uncertainty and cognitive effort; (3) determine how emotional strain or cognitive effort will impact satisfaction and engagement; and (4) determine how these experiences affect a consumer's likelihood to purchase. By conceptualizing AI-powered social media as an additional mediating system that structures the space between digital and analogue environments, this research will advance current scholarship on the relationship between algorithmic curatorial practices and accessibility, well-being, and opportunities for equitable experiences.

Methods

This study employs an experimental survey design to examine how sensory overload conditions in physical retail environments affect cognitive and emotional strain, and whether AI-driven retail features or AI powered social media tools can mitigate these effects among neurodivergent consumers. Participants will be randomly assigned to one of four scenarios that manipulate (1) control (no AI, no social media), (2) AI driven

retail features only, (3) AI-powered social media only, and (4) the presence of both types of AI. This 2x2 design allows for testing the main and interactive effects of sensory stimuli and AI support on cognitive outcomes and consumer responses. For example, the scenario for AI-powered will be the following:

Imagine that you use social media regularly and occasionally encounter shopping-related content while browsing. In this scenario, AI operates in the background of your social media apps after you give permission through standard app settings. The AI does not access private messages or change your social media feed. Instead, it observes general shopping-related activity, such as the types of products, clothing styles, brands, and price ranges you interact with over time.

Based on these patterns, the AI sends occasional shopping-related notifications directly to your phone. These notifications suggest relevant products or highlight options that align with your preferences and typical spending habits, including potential cost-saving alternatives. You decide whether to open, ignore, or dismiss each notification. The AI does not make purchases for you or affect your in-store shopping experience. You continue using social media as usual.

Participants will be consumers who self-identify as neurodivergent, who are heavy social media users, and those who shop in physical retail environments. These participants will be ages 18-35 years old and English speaking.

Results

This study anticipates that higher levels of sensory stimulus overload will be associated with increased cognitive effort and emotional strain, with these relationships expected to be particularly pronounced among neurodivergent participants. Drawing on sensory processing and cognitive load perspectives, it is expected that individuals facing elevated sensory demands will expend greater mental and emotional resources to navigate physical retail environments, thereby diminishing experiential quality.

Further, the study anticipates that exposure to AI-powered social media content prior to entering physical retail settings will be associated with lower anticipated sensory overload and reduced perceived cognitive burden. Rather than functioning solely as a persuasive or promotional tool, AI-aided social media systems are expected to operate as a form of cognitive infrastructure that shapes sensory expectations. Specifically, such systems may either alleviate or exacerbate sensory overload by mediating perceived uncertainty, cognitive effort, and emotional regulation before in-store exposure.

Participants exposed to AI-driven social media scenarios are expected to report feeling more prepared, less anxious, and more in control of their anticipated shopping experience compared to those in non-AI conditions. In addition, AI-powered social media is anticipated to moderate the negative effects of cognitive effort and emotional

strain on satisfaction and engagement, highlighting its role in shaping consumer experiences beyond the physical retail environment itself.

Finally, when AI-powered social media tools are combined with AI-driven in-store features, the study expects to observe the highest levels of satisfaction and purchase intention. This pattern would suggest that digital and physical AI systems function in a complementary manner, jointly supporting neurodivergent consumers across pre-visit and in-store stages of the retail journey.

Taken together, these anticipated findings suggest that social media functions not merely as a marketing or communication channel, but as a cognitive structuring mechanism that influences accessibility, emotional regulation, and experiential quality both before and during in-person consumption.

Future Work

Future research will further examine how different forms of AI-driven social media content, such as short-form video, influencer demonstrations, and algorithmic recommendations, differentially impact sensory preparedness and trust. Longitudinal studies may explore how repeated exposure to sensory-aware social media influences long-term shopping behaviour and platform reliance among neurodivergent consumers. Additionally, future work should address ethical considerations surrounding algorithmic personalization, including the balance between supportive curation and manipulative design practices (Ocean et al., 2025). Qualitative approaches could provide deeper insight into how neurodivergent individuals interpret and adapt to AI-curated social media environments over time.

For marketers, these findings might be useful for developing advertising on social media in a more inclusive way, paying special attention to how these stimuli impact all consumers. We also expect findings to provide insight on how store personnel can interact with all customers, paying special attention to in-store areas that have traditionally been challenging for neurodivergent customers. This work advances consumer behavior and social media scholarship by reframing accessibility as a dynamic, sensory, and emotionally mediated process rather than a static physical condition. By conceptualizing AI-powered social media as an infrastructural mechanism that shapes anticipatory regulation and psychological safety, the study opens new avenues for theorizing consumer emotional labor, platform-enabled inclusion, and harm reduction in AI-mediated environments.

References

Beyari, H., & Hashem, T. (2025). The role of artificial intelligence in personalizing social media marketing strategies for enhanced customer experience. *Behavioral Sciences*, 15(5), 700. <https://doi.org/10.3390/bs15050700>

- Bonetti, F., Montecchi, M., Plangger, K., & Schau, H. J. (2023). Practice co-evolution: Collaboratively embedding artificial intelligence in retail practices. *Journal of the Academy of Marketing Science*, 51(3), 867–888.
- Douc  , L., & Adams, C. (2020). Sensory overload in retail environments: How the presence of multiple in-store stimuli affects consumers. *Journal of Retailing and Consumer Services*, 55, Article 102110.
<https://doi.org/10.1016/j.jretconser.2020.102110>
- Finnigan, K. A. (2024). Sensory responsive environments: A qualitative study on perceived relationships between outdoor built environments and sensory sensitivities. *Land*, 13(5), Article 636. <https://doi.org/10.3390/land13050636>
- Gopan, G. (2025). Sensory inclusivity in retail environments: A design-oriented approach. *Journal of Interior Design*, 50(2), 118–140.
<https://doi.org/10.1177/10717641251327928>
- Grewal, D., Roggeveen, A. L., & Nordf  lt, J. (2017). The future of retailing. *Journal of Retailing*, 93(1), 1–6. <https://doi.org/10.1016/j.jretai.2016.12.008>
- Gursoy, D., Chi, O. H., Lu, L., & Nunkoo, R. (2019). Consumers acceptance of artificially intelligent (AI) device use in service delivery. *International Journal of Information Management*, 49, 157–169. <https://doi.org/10.1016/j.ijinfomgt.2019.03.008>
- Kyriacou, C., Forrester-Jones, R., & Triantafyllopoulou, P. (2023). Clothes, sensory experiences and autism: Is wearing the right fabric important? *Journal of Autism and Developmental Disorders*, 53(4), 1495–1508.
<https://doi.org/10.1007/s10803-021-05140-3>
- Ocean, L., Lang, L., & Pearson, E. (2025). Dark patterns and neurodiversity: How online retail design disadvantages neurodivergent consumers. *Journal of Consumer Policy*, 48(1), 55–78. <https://doi.org/10.1007/s10603-024-09582-2>
- Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and artificial intelligence: An experiential perspective. *Journal of Marketing*, 85(1), 131–151.
<https://doi.org/10.1177/0022242920953847>
- Sweller, J. (2011). Cognitive load theory. In *Psychology of learning and motivation* (Vol. 55, pp. 37–76). Academic Press. <https://doi.org/10.1016/B978-0-12-387691-1.00002-8>
- Weber, F. D., & Sch  tte, R. (2019). State-of-the-art and adoption of artificial intelligence in retailing. *Digital Policy, Regulation and Governance*, 21(3), 264–279.
<https://doi.org/10.1108/DPRG-09-2018-0050>
- Xiao, L., & Kumar, V. (2021). Robotics for customer service: A useful complement or an ultimate substitute? *Journal of Service Research*, 24(1), 9–29.
<https://doi.org/10.1177/1094670520910542>

Yrjölä, M., Hautamäki, P., & Ranta, R. (2024). Value propositions of artificial intelligence in retailing. *The International Review of Retail, Distribution and Consumer Research*, 34(5), 421–447. <https://doi.org/10.1080/09593969.2024.2411209>

1C.4 Mainstreaming the Red Pill: The Popularization of Red Pill Ideology on TikTok, Instagram, and YouTube.

Duygu Karatas (University of Siegen)

Introduction

A gruesome act of gender-based violence shook Türkiye on 4 October 2024, horrifying the nation and reigniting discussions surrounding femicide, misogyny, and incel culture. Semih Çelik, a 19-year-old man, brutally murdered Ayşenur Halil and İkbâl Uzuner in Istanbul before taking his own life (“İkbâl Uzuner ve Ayşenur Halil’in öldürülmesi”, 2024). The violent incident not only spotlighted the escalating crisis of gender-based violence in Türkiye but also introduced the public to the alarming influence of misogynistic subcultures, which play a growing role in fostering extremist attitudes and behaviours in society (Türkiye'nin İncel gerçeği, 2024).

Researchers widely agree that the emergence and proliferation of the Internet have significantly contributed to the spread of misogyny, gender-based discrimination, and extremist ideologies across a variety of online platforms (Marwick & Caplan, 2018; The CCDH Quant Lab. 2022). Initially confined to niche anti-feminist male groups, these communities have emerged as a reactionary force against contemporary feminist advancements, fostering online hate and promoting gender-based violence against women. The transnationalization of these anti-feminist ideologies has been further facilitated by online anonymity, which enables participants to craft idealized personas and engage in personal attacks without the constraints of physical identity and distance (Ging, 2017). Multimodal communication forms, such as memetic imagery, videos, trolling, and vernacular internet humor, have played a crucial role in disseminating and popularizing these ideologies (Marwick & Lewis, 2017).

Baele, Brace, & Ging (2023) highlight that current research frequently overlooks the dynamic nature and heterogeneity of incel communities, which are influenced by the specific affordances of different platforms. Solea and Sugiura (2023) highlight the need for a more nuanced, dynamic approach to studying incel communities, as the beliefs, levels of toxicity, and harms they propagate vary significantly depending on the platforms. While mainstream platforms enforce stricter moderation policies, their increased visibility allows for broader dissemination of incel ideologies. Conversely, fringe platforms, which often lack adequate moderation, harbor more extreme and unregulated content (Mamié et al., 2021; Ribeiro et al., 2020; Wright et al., 2020). Despite moderation efforts, misogynist narratives continue to infiltrate mainstream spaces, adapting to evade platform restrictions and reach wider audiences (Sugiura, 2021).

Objectives/Research Questions

Building on Solea and Sugiura's call for a dynamic approach and their findings on the 'normification' and normalization of incel ideology on mainstream social media platforms (2023), this cross-platform research aims to investigate the effect of mainstream visual-based social media platforms (TikTok, Instagram, and YouTube) on the popularization and normalization of red pill ideology. The research addresses the following questions: What popularization techniques and multimodal communication forms (e.g., visual aesthetics, memes, short-form videos, vernacular humor, emotional appeals) are employed to disseminate red pill narratives across platforms? What similarities and differences emerge in communication forms, discursive strategies, and narrative framings of red pill ideology across TikTok, Instagram, and YouTube? Which foreign influencers and content creators contribute to the transnational circulation and local adaptation of red pill discourse across these platforms?

Methods

To examine popularization techniques on major mainstream platforms such as YouTube, TikTok, and Instagram, the study focuses on high-reach accounts (with more than 10,000 followers) that utilize image- and video-based content to disseminate red pill narratives. The research employs "Multimodal Discourse Analysis" to explore how red pill ideology is visually, verbally, and affectively constructed across platforms. The analysis systematically examines visual elements (imagery, aesthetics, body display, and symbols), audio elements (narration, tone, music, and sound effects), textual elements (captions, overlay text, and hashtags), and narrative strategies, including themes of victimhood, masculinity, hierarchy, and anti-feminism.

In addition, a thematic analysis following Braun and Clarke (2021) will be conducted to identify recurring patterns, themes, and discursive tropes across platforms. This process involves the initial coding of visual, textual, and audio elements, the clustering of codes into themes and sub-themes, and a cross-platform comparison to identify both convergences and divergences in the popularization and normalization of red pill ideology.

Results

Preliminary findings indicate that red pill discourse on mainstream visual-based platforms is predominantly framed through power relations and relationship advice, while simultaneously mobilizing explicitly misogynistic and anti-feminist narratives. Although often presented as guidance on masculinity, self-improvement, and romantic success, the content relies heavily on antagonistic "us versus them" framings, reinforcing in-group and out-group polarization. This structure mirrors broader dynamics of contemporary political and cultural polarization, positioning red pill ideology as a gendered reflection of polarizing politics.

Across platforms, red pill narratives primarily target young men who are highly engaged with social media, particularly in the Turkish context. Preliminary analysis points to the emergence of a marketized masculinity centered on bodily discipline, fitness, aesthetics, and consumption practices. Masculinity is framed as something to be optimized, displayed, and consumed, with strong emphasis on physical appearance, luxury brands, and financial success. These values are frequently legitimized through claims that women prioritize money and status, reinforcing misogynistic interpretations of gender relations.

Short-form videos play a key role in normalizing these narratives through repetition, emotional intensity, and algorithmic visibility, while longer thematic videos enable more sustained ideological elaboration of gender hierarchies and male dominance. The discourse is characterized by a notably harsh and aggressive tone, targeting not only women and feminists but also men themselves through internal hierarchies such as “alpha” and “beta” classifications. Narratives of victimhood thus coexist with internal stratification and mutual devaluation among men.

Feminist women are frequently mocked, stereotyped, and caricatured, while women engaging in aesthetic practices are simultaneously labeled as materialistic or “gold diggers,” contributing to the homogenization of women as a collective out-group. Notably, the discourse appears largely detached from explicit religious or secular frameworks opposed to the society in Turkey, instead operating as a unifying rhetorical umbrella. Transnational influencers play a central role in this process, with international figures frequently translated, subtitled, or visually adapted for Turkish audiences. The growing use of stylized audiovisual formats, including AI-generated voices and depersonalized imagery, further facilitates the normalization and circulation of red pill ideology across platforms.

Future Work

YouTube to capture newly emerging content, accounts, and formats related to red pill discourse. While an initial corpus of data has already been collected, the dataset will be continuously refined to reflect platform dynamics and shifts in visibility over time. The collected material will be systematically analyzed using multimodal discourse analysis and thematic analysis to identify recurring patterns, narrative strategies, and platform-specific differences. Based on this analysis, the final findings will be articulated to demonstrate how red pill ideology is popularized and normalized across mainstream visual social media platforms.

References

Baele, S. J., Brace, L., & Coan, T. G. (2019): From “Incel” to “Saint”: Analyzing the violent worldview behind the 2018 Toronto attack. *Terrorism and Political Violence*, 33(8), 1667–1691. <https://doi.org/10.1080/09546553.2019.1638256>

- Baele, S. J., Brace, L., & Ging, D. (2023). A diachronic cross-platforms analysis of violent extremist language in the incel online ecosystem. *Terrorism and Political Violence*. <https://doi.org/10.1080/09546553.2022.2161373>
- Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical Guide*. SAGE.
- Ging, D. (2017): Alphas, betas, and incels: Theorizing the masculinities of the manosphere. *Men and Masculinities*, 22(4), 638–657.
- Han, X., & Yin, C. (2022). Mapping the manosphere. Categorization of reactionary masculinity discourses in digital environment. *Feminist Media Studies*, 1–18.
- İkbal Uzuner ve Ayşenur Halil'in öldürülmesi üniversitelerde protesto edildi (The murders of İkbal Uzuner and Ayşenur Halil were protested at universities..) (2024, October 5). BBC Türkçe. Retrieved from <https://www.bbc.com/turkce/articles/cp8m4en3mzvo>
- Mamié, R., Ribeiro, H., & West, R. (2021). Are anti-feminist communities gateways to the far right? Evidence from Reddit and YouTube. <https://arxiv.org/abs/2102.12837>
- Marwick, A., & Lewis, R. (2017). Media manipulation and disinformation online. *Data and Society*. <https://datasociety.net/library/media-manipulation-and-disinfo-online>
- Ribeiro, M. H., Jhaver, S., Zannettou, S., Blackburn, J., De Cristofaro, E., Stringhini, G., & West, R. (2020). Does platform migration compromise content moderation? Evidence from r/The_Donald and r/Incels. <http://arxiv.org/abs/2010.10397>
- Rothermel, A. (2023). The role of evidence-based misogyny in antifeminist online communities of the 'manosphere'. *Big Data & Society*, 10(1), 20539517221145671. <https://doi.org/10.1177/20539517221145671>
- Solea, A. I., & Sugiura, L. (2023). Mainstreaming the Blackpill: Understanding the Incel Community on TikTok. *European Journal on Criminal Policy and Research*, 29(3), 311–336. <https://doi.org/10.1007/s10610-023-09559-5#>
- Sugiura, L. (2021). *The Incel Rebellion: The Rise of Online Misogyny*. Emerald
- The CCDH Quant Lab - Centre For Countering Digital Hate (2022). *The Incelosphere*. Retrieved from <https://counterhate.com/wp-content/uploads/2023/08/CCDH-The-Incelosphere-FINAL.pdf>
- Türkiye'nin İncel gerçeği: Nereden geldi, nasıl yayıldı? (Turkey's Incel reality: Where did it come from, and how did it spread?) (2024, October 7). DW. Retrieved from <https://www.dw.com/tr/t%C3%BCrkiyenin-i%C3%87ncel-ger%C3%A7e%C4%9Fi-nereden-geldi-nas%C4%B1l-yay%C4%B1ld%C4%B1/a-70428938>

Wright, S., Trott, C., & Jones, C. (2020). 'The pussy ain't worth it, bro': assessing the discourse and structure of MGTOW. *Information, Communication & Society*, 23(6), 908–925. <https://doi.org/10.1080/1369118X.2020.1751867>

Session 1D - Discourse and Public Opinion (1)

Day 1 · Mon 13 July · 9:30–11:00 · Studio 2, ARC

1D.1 Exploring the temporal constructions of Facebook/Meta: presentist speed, progress, and exponentiality.

Asher Kessler (*London School of Economics*)

Introduction

At the turn of the millennium, many social theorists argued that Western society was experiencing a crisis in historical time. François Hartog (2015), for example, proposed that the West had entered a new “regime of presentism,” while scholars from Francis Fukuyama (1989) to Mark Fisher (2009) suggested that people could no longer imagine a future meaningfully different from the present. During this period, academic attention increasingly turned to the cultures and discourses of the ‘digital revolution.’ In this context, Manuel Castells (2010: 433-434) argued that the “linear, irreversible, measurable, predictable time” of progress was “being shattered in the network society,” replaced by a new temporal order called “timeless time” in which the future and past disappeared into “the ever-present.”

However, in the following decades, American computer culture has underwent significant transformation, especially as a result of the twin processes of platformisation and the entrenchment of ‘Big Tech’. As an exemplary case of both these processes, the evolution of Facebook/Meta’s discourse reveals more than the story of one company. It offers insights more generally into the development of American computer culture, and specifically Big Tech, in the decades following 2004.

While early theories suggested that digital culture was inherently presentist, the evolution of Facebook/Meta’s discourse over two decades suggests a more complex reality. Rather than a single overarching temporal order, the company’s internal and external communications reveal a “polyphony” of multiple temporal rhythms (Tsing, 2015).

This research draws on the theoretical metaphor of layers’ from Reinhart Koselleck (2018) to examine the “simultaneity of the non-simultaneous”, the different historical tempos that exist at any one time. It also utilises Sarah Sharma’s (2014) framework to understand how temporalities intersect with power relations, specifically how tech actors exert power by normalising certain experiences of time and demanding that

others recalibrate to them. By stripping apart these layers, the study reframes Facebook's discourse as strategic interventions in the broader human experience of time.

Objectives/Research Questions

The primary objective of this research is to question whether American computer culture should still be understood as existing within a purely presentist historical time. By analysing the discursive output of Facebook actors from 2004 to 2024, the study seeks to disaggregate the distinct articulations of historical time embedded in Big Tech's narratives. Further empirical questions included:

- How did Facebook/Meta's temporal discourse evolve from its inception to its current focus on the metaverse and AI?
- What were the specific layers of historical time present in the company's language?
- How were these different temporal constructions used to legitimate company actions, such as global expansion and massive technological investments?
- What were the distinct affects and demands for recalibration produced by these competing historical times?

Ultimately, the study aims to move beyond the totalising rhythm of 'timeless time' to reveal how Facebook/Meta wields a 'polyphony' of historical rhythms to achieve its corporate and world-historical ambitions.

Methods

The research utilises a qualitative thematic analysis of a custom-built digital archive, collected during and after my doctoral studies, covering the years 2004-2024. The archive consists of texts and utterances from pivotal Facebook/Meta figures, including Mark Zuckerberg, Sheryl Sandberg, Andrew Bosworth, Peter Thiel, and others.

Data Collection:

To analyse the evolution of Facebook/Meta's discourse, I analysed the texts and utterances Mark Zuckerberg, Andrew Bosworth, Sheryl Sandberg, Peter Thiel, Chamath Palihapitiya, Alex Schultz, Michael Abrash, and Nick Clegg, as well as documents and blogs published by Facebook/Meta the company from 2004-2024. For this research, which builds upon my doctoral thesis, over 2000 documents were ultimately collected and analysed. Key sources included the 'Zuckerberg Files,' official Facebook/Meta blogs, personal blogs, and transcripts of lectures given by company leaders.

Corpus Construction:

To ensure a manageable and relevant dataset, I employed a process of "corpus construction," selecting 371 documents based on systematic criteria (Bauer and Gaskell, 2000:20). Irrelevant documents (e.g., personal anecdotes) and repetitive

documents were discarded. Documents exceeding 75 pages were also excluded for practical reasons.

Thematic Network Analysis:

The study applied Thematic Network Analysis (Attride-Stirling, 2001), which distinguishes between basic codes, organising themes, and global themes. The analysis focused on temporalised language to identify the underlying tempos, rhythms, and cadences of the discourse. In the initial coding stage, I produced 50 based codes through an iterative process using Nvivo software. These codes were clustered into 'organising themes' (middle-order clusters) and finally into 'global themes' (high-level abstractions that present a core argument).

Results

The analysis reveals that Facebook's discourse is not a monolith of presentism but a polyphony of three distinct layers of historical time: presentist speed, progress, and exponentiality. In its formative years (2004–2008), Facebook was fixated on speed as a competitive necessity. This layer matches Castells' notion of 'timeless time', focusing on the ever-present movement of speed without concern for past precedents or future consequences.

However, as Facebook expanded globally, it adopted the language of linear, chronological progress. Mark Zuckerberg began framing the company's evolution within a long arc of human history - moving from tribes to cities to nations - legitimising Facebook's 'global community' as the inevitable next step in universal history. Unlike speed, the time of progress incorporates morality, depicting the company's actions as breeding tolerance and understanding.

As Facebook engaged with and legitimated their own global expansion, they not only turned to the time of progress, but also to the language and logic of exponentiality. Figures, such as Peter Thiel and Zuckerberg use exponential logic to justify losing money in the present to capture massive future value. Exponentiality produces a sense of disorientation and bewilderment whilst normalising vast inequalities between those who harness the curve and those who do not.

Three findings emerge from this historical analysis of Facebook/Meta's discourse. Firstly, Facebook moved away from a purely presentist fixation on speed toward the more complex narratives of progress and exponentiality. Secondly, actors in Facebook/Meta could strategically wield these different historical times to achieve different goals. Progressive time provided a calm and moral legitimacy and sense of order, whilst exponential time created an urgent, anxious demand for immediate recalibration to new technologies. Lastly, these layers existed simultaneously, sometimes harmonising to justify Facebook's global expansion, and sometimes clashing in their visions of the future.

Ultimately, this article concludes that Big Tech's power is derived not only from its manipulation of big data human sociality, nor its financial resources alone, but its capacity to define the tempo of history.

References

- Attride-Stirling J (2001) Thematic Networks: An Analytical Tool for Qualitative Research. *Qualitative Research* 1(3):385-405.
- Bauer ML and Gaskell G (2000) *Qualitative researching with Text, Image and Sound: A Practical Handbook*. SAGE.
- Castells M (2010) *The Rise of the Network Society: The Information Age – Economy, Society and Culture*, Oxford: John Wiley & Sons.
- Fisher M (2009). *Capitalist Realism*. London: Zero Books
- Fukuyama F. (1989). *The End of History?*. National Interest.
- Koselleck R (2018). *Sediments of Time: On Possible Histories*. Stanford: Stanford University Press.
- Sharma S (2014). *In the Meantime: Temporality and Cultural Politics*, North Carolina: Duke University Press
- Tsing AL (2015) *The Mushroom at the End of the World: On the Possibility of Life in Capitalist Ruins*. New Jersey: Princeton University Press.
-

1D.2 Children in Political Imagery: A Visual Values Analysis of Turkish Leaders' Instagram Posts.

Ayşenur Benevento (Bahcesehir University)

Introduction

This project examines how children are visually represented and politically mobilized in the Instagram communication of four central figures in contemporary Turkish politics: Recep Tayyip Erdoğan, Kemal Kılıçdaroğlu, Ekrem İmamoğlu, and Devlet Bahçeli. In Turkey, where political identity is intensely mediated through visual symbolism and emotional rhetoric, Instagram has become a key site for the construction of affectively resonant public personas and national narratives. Images of children occupy a particularly powerful place within these communicative ecologies: they evoke trust, warmth, hope, and continuity; they signal moral character; and they function as symbols of imagined collective futures. Despite their pervasiveness in political communication, children's images in social media have rarely been subjected to systematic empirical or interpretive analysis.

The study positions Instagram images not as neutral representations but as strategic narrative acts that encode cultural values, emotional orientations, and ideological commitments. While existing research on political communication has examined visual rhetoric, personalization, and affective appeals, there remains a significant gap in understanding how childhood itself is constructed as a political and emotional resource on social media platforms. Addressing this gap is especially important in the Turkish context, where leaders frequently feature children in posts related to national holidays, education, religious ceremonies, charity events, rallies, and everyday encounters staged for public visibility. These representations are not merely decorative; they participate in broader struggles over nationalism, morality, civic identity, and the future of the nation.

By focusing on a single country while comparing ideologically distinct leaders, this project offers a contextually grounded analysis of how childhood is symbolically shaped in digital political culture. It contributes to social media research by foregrounding visual data, platform-specific practices, and affective communication, and it introduces a systematic method for analyzing the values embedded in political imagery.

Objectives/Research Questions

The primary objective of this study is to identify and interpret the cultural values and emotional logics that structure political leaders' visual representations of children on Instagram. The project asks:

- 1) What values are enacted in Turkish political leaders' Instagram posts that include children?
- 2) How do these values differ across leaders representing distinct ideological positions within the same national and cultural context?
- 3) How are children positioned visually in relation to leaders, audiences, and national narratives (e.g., as symbols of the future, moral exemplars, objects of protection, or citizens-in-the-making)?
- 4) How do visual elements such as activity, setting, gesture, clothing, props, facial expression, and composition contribute to the affective framing of childhood in political communication?

By addressing these questions, the study aims to clarify how childhood is used as a symbolic and emotional resource in digital political discourse, and how social media platforms facilitate the circulation of such values in everyday public life.

Methods

The study employs a qualitative visual content analysis of Instagram posts using values analysis, a method developed to identify culturally specific goals, norms, and meaning systems embedded in visual narratives. Originally applied to everyday parenting cultures on Instagram, values analysis is extended here to the domain of political communication, representing a methodological innovation in the study of social media imagery.

The dataset consists of publicly available Instagram posts from Recep Tayyip Erdoğan, Kemal Kılıçdaroğlu, Ekrem İmamoğlu, and Devlet Bahçeli collected over a defined multi-year period. Posts featuring children are identified and compiled into a corpus of images accompanied by captions and hashtags. Each image is treated as a unit of analysis and coded for semiotic features that signal underlying values, including: the type of activity depicted (structured events versus informal interactions), environmental setting (institutional, domestic, public), social configuration (child alone, with peers, with leader), acknowledgment of the camera, facial visibility and expression, bodily proximity between child and leader, clothing and adornments, presence of symbolic props (e.g., flags, books, religious or national objects), and visual composition or editing.

Coding categories are iteratively developed and applied to capture recurring value dimensions such as protection, discipline, intimacy, collectivism, religiosity, nationalism, modernity, inclusion, and futurity. The analysis focuses on how these visual elements combine to construct coherent emotional and moral narratives about childhood. By comparing patterns across leaders within the same platform and national context, the method enables a fine-grained examination of how ideological positions are enacted through platform-based visual communication.

Results

Preliminary analysis indicates that children's images are central to the affective and symbolic economies of Turkish political Instagram communication. Across all four leaders, children function as emotionally charged figures through which moral authority, national continuity, and leadership character are visually articulated. However, the values enacted through these representations vary systematically by political orientation.

Erdoğan's posts tend to position children within narratives of protection, religious-moral continuity, and paternal authority, often framing the leader as a guardian figure responsible for the nation's youth. Kılıçdaroğlu's imagery more frequently emphasizes equality, compassion, and civic responsibility, presenting children as deserving subjects within a moral economy of social justice. İmamoğlu's posts commonly enact values of warmth, accessibility, urban modernity, and democratic optimism, positioning children as symbols of renewal and inclusive futures. Bahçeli's representations foreground discipline, loyalty, and nationalist solemnity, frequently situating children within ceremonial or patriotic contexts that stress collective duty.

Across cases, visual strategies such as bodily proximity, children's gaze, emotional expression, and the use of symbolic objects contribute to the construction of affective publics. Repeated exposure to such imagery naturalizes particular understandings of childhood, leadership, and national values. Instagram thus operates not only as a channel for political messaging but as a site where emotional and moral meanings are curated, archived, and circulated.

Future Work

Future phases of this research will expand the dataset temporally to examine how representations of children shift in response to major political events, electoral cycles, or national crises. Additional analysis will integrate audience-facing elements such as captions, hashtags, and user engagement metrics to explore how visual values interact with textual framing and public reception. Methodologically, the project will further refine the values analysis framework for application to other domains of political imagery and comparative platform studies.

The study also opens avenues for cross-national comparison, enabling examination of how childhood is symbolically constructed in different political cultures. Beyond political communication, the framework can be applied to branding, activism, and institutional social media use, contributing to broader debates about affect, representation, and power in digital societies.

References

Benevento, A. (2022). Studying values about childhood using networked photography on Instagram. *Semiotica*. <https://doi.org/10.1515/sem-2021-0056>

Daiute, C. (2014). Narrative inquiry: A dynamic approach. Sage.

Rose, G. (2016). Visual methodologies: An introduction to researching with visual materials (4th ed.). Sage.

1D.3 Mongolian Stand-Up Comedy on YouTube as Public Discourse.

*Borchuluun Yadamsuren (LUC), Undrah Baasanjav (Southern Illinois University),
Ideljargal Dashdondog (National University of Mongolia)*

Introduction

Since Plato and Aristotle's time people have used comedy and humor in challenging dominant norms and power structures. In the 17th century, when laughter was regarded as a not good thing socially, Thomas Hobbes wrote how those in power became the targets of public laughter for their behaviors. Immanuel Kant's writing explains the laughter resulting from a heightened expectation into nothing, leading to the incongruity theory of humor. Zygmund Freud wrote the subversive nature of humor and its nature of eliding civilization in "convulsive acts." These ideas have been at the heart of the humor theory, leading to the superiority, the incongruence and the relief theories of humor (Quirk, 2016).

In Mongolia, as a historically nomadic civilization, oral satire has long appeared through humorous narratives, folktales, folk songs, and proverbs to critique social norms and power structures. During the socialist time, there was limited criticism of the society in the form of light satire in films and a few television programs targeting bureaucrats and incompetent workers (Baasanjav, 2025). Comedians like Sembee and Latiff on television or comic films like Aman Khuur, Zereg Nemekhyn Omno preceded comedy groups (Khoshin Shogiinhon) and television shows in the 1990s and 2000s. Yet, in a broader sense critical and subversive texts were politically censored.

After the new Constitution of 1992, comedic expression expanded. Mongolia's contemporary stand-up comedy scene can be traced to 2016, when UB Comedy Club organized the country's first stand-up performances. Since then, the stand-up comedy scene has diversified with the establishment of additional clubs, including Comedy Lab, the female-led Big Sista Comedy Crew, and Clap Comedy. These organizations now maintain regular weekly performances, collaborate on large-scale joint shows. Comedians like Bataa, Amarjin, Beegii, Ider-Od, Hanu and Khosoo often blend English and Russian words into Mongolian, and mix culture references of Korean soap operas,

American pop, inner Mongolian or Russian accents when discussing modern city life and the nomadic lifestyle in countryside.

Prior research shows that stand-up comedy can challenge dominant cultural norms and identity discourses (Filani, 2020) and can remix political discourse through embodied performance and parody (Higgins & Smith, 2025). Comedy often tells “the particular with the implied general” (Corner, 2014), giving it discursive power and controversy. Yet despite the growing popularity of stand-up comedy in Mongolia, there is little to no scholarly research examining it as a form of public discourse. The contemporary stand-up scene therefore remains underexplored as a public arena where comedians negotiate politics, identity, urban life, and cultural change through humor.

Objectives/Research Questions

The goal of this study is to examine how Mongolian stand-up comedy functions as a form of public discourse by analyzing both what social, cultural, and political issues are foregrounded in live performances and how these performances circulate on YouTube to reach wider publics.

We explore whether or not Mongolian comedians influence public discourse, if they do, the extent to which they influence and how they contest public discourse.

This study combines framing analysis and discourse analysis to examine both what issues are made salient in comedians’ performances and how these issues are discursively constructed through humor, language use, and performance style.

RQ1. How do Mongolian stand-up comedians frame social, cultural, and political issues in their performances?

RQ2. How do comedians discursively perform these frames through humor and language, and how do these performances reproduce, contest, or renegotiate power relations and sociocultural tensions in contemporary Mongolia?

Methods

This study adopts a qualitative content analysis and discourse analysis approach to examine how Mongolian stand-up comedians frame social, cultural, and political issues. Framing theory (Entman, 1993) guides the identification of frame components at the level of jokes or story beats, while discourse analysis (Fairclough, 1995) supports interpretation of how frames are articulated through humor devices, language choice, and performance style. Our analysis was also informed by the theories of humor: superiority, the incongruence and the relief (Quirk, 2016).

The study applies a theory-guided coding framework (Hsieh & Shannon, 2005) to identify prevalent themes (e.g., corruption, gender relations, and generational differences) and rhetorical strategies (e.g., satire, parody, irony, exaggeration, role reversal). Discourse analysis enables researchers to connect sociocultural meanings

to broader sociopolitical structures, interactional hierarchies, deliberative processes, and everyday social practices (van Dijk, 1997). Together, these approaches conceptualize comedy as a discursive practice through which sensitive issues become speakable in playful yet politically meaningful ways.

A purposive sampling strategy was used to select widely viewed performances from the official YouTube channel of UB Comedy. Videos were selected from the “Popular” tab to prioritize high-visibility content likely to shape public talk. The unit of analysis was a joke or story beat. Each selected video was watched multiple times and segmented into discrete units based on topical coherence and punchline structure (Elo et al., 2014). For each unit, coders recorded timestamps, performer/channel, key quotes for reporting, and initial interpretation notes.

The codebook operationalized four framing functions: problem definition, causal attribution, moral evaluation, and remedy/response, alongside humor devices as framing strategies. A pilot test was conducted by two coders who independently watched and coded seven full videos. Coders applied the codebook and checklist to the same subset, compared results, and discussed disagreements to refine decision rules (e.g., satire vs. parody; implicit framing). The codebook was revised accordingly before full coding. Rather than emphasizing statistical intercoder reliability, coder discussion and documentation were used to strengthen dependability and transparency (Elo et al., 2014).

Full verbatim transcripts were not produced. Instead videos were repeatedly viewed to code frame elements and to extract representative quotations. Analysis focused on verbal content, excluding nonverbal cues, audience interaction, and platform engagement metrics. Throughout analysis, coders moved iteratively between the coding spreadsheet and the original videos to refine interpretations and to extract representative quotations with precise wording, context, and translation into English.

Results

An unprecedented phenomenon of stand-up comedians gaining influence in public discourse in Mongolia is evidenced by the millions of views on YouTube. The following themes were identified.

Gender Differences and the Reversal of Traditional Gender Roles.

Communicative differences and the reversal of gender expectations are depicted in these young male comedians’ performances as they joke about growing up, dating, going out, marrying, divorcing, and becoming parents. They joke about bros’ communicative incompetence, their short life expectancy, and their wives’ control.

When comedian Hanu jokes about the correlation between the attractiveness of female and the financial prowess of male in a relationship, he also brings up the possibility of

cosmetic surgery afforded by a rich guy, and the unfairness of the retirement system that does not take into account the short life expectancy of males.

He says “when his wife goes out with her girlfriends, the witches go to Shangri-La and spend much on expensive beers and better ambience.” He subverts the expectation of male knowing better by following his wife’s acts and blames the unfairness of the retirement system that allegedly benefits women more. The selfcare, communication competence, and physical prowess of Mongolian women are frequently compared to the “damsels in distress” in Korean soap operas as capable and even cruel.

Cultural Juxtaposition: Countryside and Abroad.

Cultural clash between city life and countryside is depicted in Ider-Od’s show describing a foreigner in a countryside being “traumatized” by a herder bro, who has good intention to offer hospitality. He chooses a situation of preparing sheep meat for feast and drinking together. The host tells Ider-Od that “he likes the guest as he is a nice guy, and he would prepare for the feast.”

The comedian feels for the foreigner, while the country bro continues without reading the situation, and even invites the guest's family next time. Drinking vodka together here appears to be a cross cultural comic relief. Ider-Od exaggerates the situation by saying if the Korean visitor is still alive after witnessing the “horror movie.” The comedian does not offer remedy for the situation, but sighs by saying “the bro has a beautiful heart that kills.”

The Normalization of Alcoholism

Bros drinking, as well women’s drinking, are commonly described in stand-up comedies, especially during new year and International Women’s Day celebration. Ider-Od depicts how women whistle at guys or brush them off in bars. On one hand, he appears to be admiring strong Mongolian women who can carry their drunken boyfriends, unlike women depicted in Korean television. On the other hand, he claims that Mongolian women are aggressive, free, aggressive and independent.

In these comedies, alcoholism is criticized, and at the same time is normalized. In that sense, comedies do not offer solutions, but lightly touch on social issues (Quirk, 2016).

Future Work

This study has several limitations that suggest directions for future research. First, because the sample was drawn from the most popular YouTube videos from UB Comedy and Comedy Lab, women comedians were less represented, partly due to their relatively later entrance into Mongolia’s stand-up scene and lower visibility. Future studies should examine women comedians’ performances to better capture gendered perspectives and framing practices. Second, while comedians frequently raised socially salient issues, they rarely offered explicit remedies. Future work should examine whether Mongolian stand-up comedy functions primarily as critique and

agenda-setting, and analyze audience responses to assess how these frames circulate and shape public discussion. Third, assessing “influence” of comedians to public discourse was difficult using performance content alone in this study. Future studies should incorporate audience-centered methods (e.g., interviews, focus groups, surveys) to evaluate perceived impact. Finally, the analysis focused on verbal content rather than nonverbal cues, audience interaction, or platform engagement metrics. Therefore, future research could strengthen rigor through full transcription and multimodal and engagement-based analysis.

References

- Baasanjav, U. (2025). *Information Society and Media Development in Modern Mongolia*. Routledge.
- Douglas, M. (1991). Jokes. In Mukerji, C. and Schudson M. (Eds.) *Rethinking Popular Culture: Contemporary Perspectives in Cultural Studies*, 291–310.
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 62(1), 107–115.
- Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, H. (2014). Qualitative content analysis: A focus on trustworthiness. *SAGE Open*, 4(1).
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51–58.
- Fairclough, N. (1995). *Media discourse*. London: Edward Arnold.
- Filani, I. (2020). A discourse analysis of national identity in Nigerian stand-up humour. *Discourse Studies*, 22(3), 319–338.
- Higgins, M., & Smith, A. (2025). Not playing the game: Political talk, hybridity and performance. *Journalism*, 26(3), 618–636.
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288.
- Quirk, S. (2016). Preaching to the converted? How political comedy matters. *International Journal of Humor Research*, 29(2), 243–260.
- Van Dijk, T.A. (1997). Discourse as interaction in society. In: T. A. van Dijk (Ed.), *Discourse as social interaction* (pp. 1–37). London: Sage.

1D.4 A Deficit of Shared Reality? Youth, Fragmented Information Environments, Social Media, and Political Participation.

Catarina Feio, Lúcia Oliveira (Universidade de Aveiro).

Introduction

New media have changed the way individuals consume information, entertainment, and communicate, particularly among young people, also affecting news consumption and its perception (Bengtsson & Johansson, 2021). In this media environment, the key question is no longer whether young people consume news, but how they do so, as they are widely exposed to informational content through social media, often incidentally (Ohme et al., 2022; Russmann & Hess, 2020).

Understanding this phenomenon is particularly relevant, as news plays a fundamental role in fostering political participation and shaping public opinion (Klopfenstein Frei et al., 2022). This concern highlights the importance of media literacy within educational systems, not only to support the development of informed citizens (Tamboer et al., 2022), but also to reduce inequalities and promote the participation of more vulnerable young people (Geers et al., 2020). In this context, digital media and the online sphere are increasingly viewed as essential tools for communication and mobilisation among young people (Pawluczuk et al., 2020). Digital devices, such as smartphones, and social media have reshaped the landscape of political participation among younger generations (Kaskazi & Kitzie, 2023)

Participation in online political debates influences young people, particularly first-time voters, as online information and opinions may shape voting decisions and electoral engagement (Intyaswati et al., 2021). However, reliance on social media as a primary news source entails risks, including political polarisation and exposure to misinformation, often amplified by emotional and moralising persuasive strategies (Igwebuike & Chimuanya, 2021). These dynamics are further complicated by artificial intelligence, which facilitates the rapid dissemination of deceptive content. Such engagement also affects peers within online networks (Marquart et al., 2020), while young people who support extremist views and distrust journalistic institutions remain especially vulnerable to anti-system narratives and fake news (Deinla et al., 2022; Juvalta et al., 2023).

Objectives/Research Questions

This work examines a segment of a broader study, which addresses the following research question: What is the impact of new media, particularly social media, on news

consumption habits and political participation among adolescents and young people in Portugal?

The study formulates four research hypotheses grounded in the literature, which are subsequently tested through statistical analysis. Regarding the preferred medium for news consumption among young people, social media emerge as the platforms on which they access news most frequently (Feio et al., 2023; Pérez-Escoda et al., 2021; Russmann & Hess, 2020), leading to the following hypothesis: H1 - Young people primarily consume news via social media. News consumption and political participation are mutually reinforcing: being informed about current political events facilitates participation, while keeping up-to-date with the news can increase interest in civic and political engagement (Matthes et al., 2023). This relationship gives rise to the following hypotheses: H2 - Young people with high levels of political participation also exhibit frequent news consumption, and vice versa. and H3 - Young people who do not consume news are politically disengaged and display low levels of political participation. The rise of political extremism and the mobilisation of citizens towards extreme positions is an increasing concern. Social media may provide a space in which such movements grow, and when combined with declining trust in traditional news agencies and institutions, this can expose young people to extremist ideologies (Cherney et al., 2022; Deinla et al., 2022). Based on this issue, the following hypothesis was formulated: H4 - Young people with a greater propensity to sympathise with extremist parties or ideologies tend to exhibit lower trust in traditional news agencies, while maintaining frequent news consumption and active engagement on social media.

Methods

The population of this study comprises young people residing in Portugal, aged between 15 and 29 years. Youth is a life stage characterised by individual development (Aroldi, 2011) and the process of emancipation, during which young people become aware of current issues and critically engage with them (Bourdieu, 1978). It is also during adolescence and youth that political attitudes are established and mature (Rekker et al., 2015). The study explores three age groups: adolescents (15-19 years), emerging adults (20-24 years; adapted from Arnett, 2000), and young adults (25-29 years).

The research adopts a mixed-methods approach, integrating both qualitative and quantitative methods (Bryman, 2012). In the first phase, focus groups were conducted to engage with the population and stimulate discussion around the topic (Morgan, 1996). A total of 18 focus group sessions were held, six per age group, with an average of 12 participants per session. A semi-structured interview guide was used to direct the discussions. The findings from these focus groups informed the design of a subsequent questionnaire survey, which forms the basis of the data analysed in this study. This

method is particularly advantageous for the statistical treatment and analysis of data (Quivy & Campenhoudt, 1995).

The questionnaire included items on sociodemographic characteristics, news consumption habits, and political participation. Existing validated scales relevant to the study's focus were adapted (Enchikova et al., 2019, 2019; Gopal & Verma, 2017; Ku et al., 2023; Waeterloos et al., 2021), and the instrument was reviewed by subject-matter experts.

The final sample comprised 1,547 participants, of whom the majority were female (65.7%), with 33% male and 1.2% preferring not to disclose their gender. The sample is representative of each age group according to the 2021 Portuguese Census.

Results

Statistical analysis indicates that social media is the preferred news source among young participants. Over half report using social media for news daily ($N = 53.1\%$), while only 4.4% never do so. Spearman correlations reveal significant, though weak, associations between social media use and other news formats: newspapers ($p < .001$; $r_s = 0.275$), blogs ($p < .001$; $r_s = 0.243$), podcasts ($p < .001$; $r_s = 0.180$), television news ($p < .001$; $r_s = 0.179$), and radio ($p < .001$; $r_s = 0.107$), suggesting that news consumption is self-reinforcing, supporting Hypothesis 1.

By quantifying news consumption habits and the relationship with news (from active and conscious engagement to avoiding news) alongside levels of political participation, a Mann-Whitney test revealed statistically significant differences in news consumption frequency between politically engaged and more apathetic individuals ($U = 63,946.5$, $Z = -7.647$, $p < .001$). These results support Hypotheses 2 and 3 and highlight the reciprocal relationship between news consumption and political participation, which can mutually reinforce or diminish each other.

Young people were analysed according to their partisan political preferences. Participants who are members of political parties, consistently support, or always vote for the same party were grouped accordingly (Bloco de Esquerda, Livre, Partido Socialista, Partido Social Democrata, Iniciativa Liberal, and Chega), while the remaining participants were classified as non-partisan ($N = 653$). Kruskal-Wallis tests compared these groups in terms of trust in media, institutions, and political participation practices. Chega supporters, a radical far-right party, reported the lowest trust in newspapers and television ($p < .001$), whereas supporters of moderate parties exhibited higher trust. Conversely, Chega supporters displayed the highest trust in social media ($p = .050$) and influencers ($p = .005$), reflecting reliance on digital platforms as news sources.

Similar patterns emerged regarding trust in institutions. Chega supporters reported the lowest trust in courts, government, parliament, and the presidency ($p < .001$). They also led online engagement, including commenting and replying to political posts ($p < .001$),

and in following politicians on social media ($p < .001$), dominating reactionary debates. Meanwhile, Chega and Socialist Party supporters jointly led in creating political content online ($p < .001$). These findings illustrate a phenomenon in which far-right-leaning youth combine low trust in institutions and traditional media with predominant social media use and high online mobilisation, reinforcing their political participation within digital spheres, validating hypotheses 4.

Future Work

This study shows that social media are the primary news source for Portuguese youth, with self-reinforcing consumption patterns closely linked to political engagement, supporting. Partisan orientation strongly influences trust: far-right supporters, particularly Chega, display low confidence in traditional media and institutions but high online engagement, whereas moderate party supporters maintain higher trust. These patterns exemplify a shared reality deficit, in which divergent media consumption and trust create fragmented perceptions of reality. Social media thus simultaneously fosters political participation and amplifies exposure to polarised content, highlighting the critical role of media literacy in promoting informed, inclusive civic engagement. This study validates a broader understanding of radicalised Portuguese youth and academically demonstrates the differences in their news consumption habits and engagement with the world around them. Future research should focus on exploring the motivations and contexts of young people, considering their informational and partisan preferences and practices. Additionally, it is recommended to reintroduce qualitative methodologies to delve deeper into the political motivations of individual youths, particularly those who are radicalised. The ultimate aim is to understand how the Portuguese educational system can serve as a foundation for citizenship, promoting media literacy and critical thinking to prevent the emergence of anti-democratic behaviours among future generations.

References

- Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, 55(5), 469–480.
<https://doi.org/10.1037/0003-066X.55.5.469>
- Aroldi, P. (2011). Generational belonging between media audiences and ICT users. June. <https://www.researchgate.net/publication/228365425>
- Bourdieu, P. (1978). Sport and social class. *Social Science Information*, 17(6), 819–840.
<https://doi.org/10.1177/053901847801700603>
- Bryman, A. (2012). *Social research methods* (4. ed). Oxford Univ. Press.
- Cherney, A., Belton, E., Norham, S. A. B., & Milts, J. (2022). Understanding youth radicalisation: An analysis of Australian data. *Behavioral Sciences of Terrorism and Political Aggression*, 14(2), 97–119.

- Deinla, I. B., Mendoza, G. A. S., Ballar, K. J., & Yap, J. K. (2022). The link between fake news susceptibility and political polarization of the youth in the Philippines. *Asian Journal of Political Science*, 30(2), 160–181.
- Enchikova, E., Neves, T., Mejias, S., Kalmus, V., Cicognani, E., & Ferreira, P. D. (2019). Civic and Political Participation of European Youth: Fair Measurement in Different Cultural and Social Contexts. *Frontiers in Education*, 4.
<https://doi.org/10.3389/feduc.2019.00010>
- Feio, C., Oliveira, L., & Martins, J. M. (2023). As Redes Sociais como meio de expressão política dos jovens em Portugal. *Observatorio (OBS*)*, 17(4), 205–223.
<https://doi.org/10.15847/obsOBS17420232325>
- Geers, S., Boukes, M., & Moeller, J. (2020). Bridging the gap? The impact of a media literacy educational intervention on news media literacy, political knowledge, political efficacy among lower-educated youth. *Journal of Media Literacy Education*, 12(2), 41–53. <https://doi.org/10.23860/JMLE-2020-12-2-4>
- Gopal, K. & Verma. (2017). Political Participation: Scale Development and Validation. *International Journal of Applied Business and Economic Research*, 15(21), 391–404.
- Igwebuike, E. E., & Chimunya, L. (2021). Legitimizing falsehood in social media: A discourse analysis of political fake news. *Discourse and Communication*, 15(1), 42–58. <https://doi.org/10.1177/1750481320961659>
- Intyaswati, D., Maryani, E., Sugiana, D., & Venus, A. (2021). Using media for voting decision among first-time voter college students in West Java, Indonesia. *Academic Journal of Interdisciplinary Studies*, 10(1), 327–339.
<https://doi.org/10.36941/ajis-2021-0028>
- Juvalta, S., Speranza, C., Robin, D., El Maohub, Y., Krasselt, J., Dreesen, P., Dratva, J., & Suggs, L. S. (2023). Young people’s media use and adherence to preventive measures in the “infodemic”: Is it masked by political ideology? *Social Science and Medicine*, 317(May 2022), 115596.
- Kaskazi, A., & Kitzie, V. (2023). Engagement at the margins: Investigating how marginalized teens use digital media for political participation. *New Media and Society*, 25(1), 72–94. <https://doi.org/10.1177/14614448211009460>
- Klopfenstein Frei, N., Wyss, V., Gnach, A., & Weber, W. (2022). “It’s a matter of age”: Four dimensions of youths’ news consumption. *Journalism*, 25(1), 100–121.
<https://doi.org/10.1177/14648849221123385>
- Ku, Z., Bertani, M., Ivan, L., & Mert, H. (2023). Developing a Digital and Traditional Political Participation (DTPP) Scale for Youth: A Validity and Reliability Study.

- Romanian Journal of Communication and Public Relations, 25(1), 7–21.
<https://doi.org/10.21018/rjcpr.2023.1.351>
- Marquart, F., Ohme, J., & Möller, J. (2020). Following Politicians on Social Media: Effects for Political Information, Peer Communication, and Youth Engagement. *Media and Communication*, 8(2), 197–207. <https://doi.org/10.17645/mac.v8i2.2764>
- Matthes, J., Heiss, R., & van Scharrel, H. (2023). The distraction effect. Political and entertainment-oriented content on social media, political participation, interest, and knowledge. *Computers in Human Behavior*, 142(January).
<https://doi.org/10.1016/j.chb.2022.107644>
- Morgan, D. L. (1996). Focus Groups. *Annual Review of Sociology*, 22(1), 129–152.
<https://doi.org/10.1146/annurev.soc.22.1.129>
- Ohme, J., Andersen, K., Albæk, E., & de Vreese, C. H. (2022). Anything Goes? Youth, News, and Democratic Engagement in the Roaring 2020s. *International Journal of Press/Politics*, 27(3), 557–568. <https://doi.org/10.1177/19401612221093008>
- Pawluczuk, A., Hall, H., Webster, G., & Smith, C. (2020). Youth digital participation: Measuring social impact. *Journal of Librarianship and Information Science*, 52(1), 3–15. <https://doi.org/10.1177/0961000618769975>
- Pérez-Escoda, A., Pedrero-Esteban, L. M., Rubio-Romero, J., & Jiménez-Narros, C. (2021). Fake news reaching young people on social networks: Distrust challenging media literacy. *Publications*, 9(2). <https://doi.org/10.3390/publications9020024>
- Quivy, R., & Campenhoudt, L. V. (1995). *Manual de Investigação em Ciências Sociais*. Gravitá.
- Rekker, R., Keijsers, L., Branje, S., & Meeus, W. (2015). Political attitudes in adolescence and emerging adulthood: Developmental changes in mean level, polarization, rank-order stability, and correlates. *Journal of Adolescence*, 41, 136–147. <https://doi.org/10.1016/j.adolescence.2015.03.011>
- Russmann, U., & Hess, A. (2020). News Consumption and Trust in Online and Social Media: An In-depth Qualitative Study of Young Adults in Austria. *International Journal of Communication*, 14, 3184–3201.
- Tamboer, S. L., Kleemans, M., Molenaar, I., & Bosse, T. (2022). Developing A Model of News Literacy in Early Adolescents: A Survey Study. *Mass Communication and Society*, 00(00), 1–25. <https://doi.org/10.1080/15205436.2022.2048027>
- Waeterloos, C., Walrave, M., & Ponnet, K. (2021). Designing and validating the Social Media Political Participation Scale: An instrument to measure political participation on social media. *Technology in Society*, 64(January), 101493.

Session 1E - **WORKSHOP: Meta Content Library as a Research Tool**

Day 1 · Mon 13 July · 9:30–13:00 · Seminar Suite B, ARC (Please note – there is a coffee and comfort break at 11:00)

Yair Rubinstein (Research Partnerships - Meta), Evgenii Strepetov (Software Engineer - Meta)

This session will introduce researchers to Meta Content Library and demonstrate the latest features available in its user interface (UI) and API. The Content Library provides researchers comprehensive access to public content from Facebook, Instagram, and Threads. Users can surface billions of data points from public posts, videos, photos, reels, story highlights, and more. The in-depth demonstration led by our partnerships team will show how the API and the UI can surface this content and illuminate topics relevant to the Social Media and Society community. We will also provide an overview for how individuals and research teams can apply for access to these tools, as well as provide an opportunity for attendees to share feedback with us about our products and services.

As a senior manager on Meta's Research Partnerships team, Yair has years of experience educating the global research community about Meta's transparency datasets. Meta Content Library is our flagship research tool for social media researchers and this tutorial will give attendees a hands-on look at how its latest features can enhance their analysis of Meta's platforms and contribute to advancing their social media research. Since Meta's last presentation at SMSociety in 2024, we've incorporated significant changes to Meta Content Library to improve the overall user experience. For instance, we've made the application and account management process more seamless efficient with the launch of our Meta-hosted portal Research Tools Manager. We've also expanded the set of options for researchers to access MCL's API with the launch of the first-party clean room, Secure Research Environment. We are excited to introduce these new features to the SMSociety community at this year's conference.

Session 2A - Algorithms and AI (2)

Day 1 · Mon 13 July · 11:30–13:00 · RD Suite, ARC

2A.1 Algorithmic Femininity on Instagram: A Case Study of the VI @fit_aitana.

Anna Pujadas-Gómez, Clara Virós-Martín (Universitat Pompeu Fabra).

Introduction

Over the last decade, digital platforms have transformed how identity is performed and monetized, notably through the emergence of virtual influencers (VIs)—digitally created personas that accumulate followers and endorse brands similarly to human social media influencers (SMI) (Sands et al., 2022; Thomas and Fowler, 2021). While early research emphasized SMIs as authentic and relatable figures (Freberg et al., 2011), VIs represent a further stage of performance, in which authenticity is scripted rather than lived (Drenten and Brooks, 2020). Unlike SMIs, whose lives build their narratives, VIs are carefully constructed to embody desirable aesthetics and strategically aligned brand values (Thomas and Fowler, 2021). Cases such as Lil Miquela, Shudu, and Lu do Magalu illustrate how VIs have moved from novelty to profitable marketing tools (Carrillo-Durán et al., 2024; Drenten and Brooks, 2020; Gambetti and Kozinets, 2024).

Yet VIs raise questions that exceed their commercial appeal. Scholars have noted that audiences frequently respond to them as social actors, generating parasocial relationships despite their acknowledged artificiality (Liu and Wang, 2025). Others emphasize how VIs' hyperreal imagery reproduces gendered and racialized ideals, aligning their bodies with conventional standards of beauty and desirability (Fowler and Thomas, 2025; Lee and Yuan, 2023). This duality positions VIs at the intersection of technological innovation and cultural critique: they embody posthuman possibilities of identity free from biological constraints (Conti et al., 2022), yet risk amplifying stereotypes of sexualization and objectification (Oktan and Yavas, 2024).

This tension is pronounced in the case of @fit_aitana, branded as Spain's first AI influencer. Unlike some VIs that obscure their artificial nature, @fit_aitana openly presents herself as digital, yet this disclosure does little to prevent audiences from engaging with her as if she were human. Her growing follower base, engagement levels, and curated presence make her a valuable case for understanding how virtual femininity is constructed, consumed, and debated within social media. Rather than an isolated example, @fit_aitana reflects patterns that are increasingly normalized in how virtual femininity is produced and engaged with on social media.

This study explores how audiences respond to @fit_aitana, paying particular attention to the dynamics of engagement, sexualization, authenticity, and emotional expression. These concerns connect to broader debates on empowerment and objectification in the digital age. Is a VI like @fit_aitana an empowering simulation that enables new forms of identity free from physical constraints? Or is she an optimized product that reproduces established ideals of beauty and desirability with even greater efficiency? This study investigates how @fit_aitana's digital embodiment navigates the tension between empowerment and objectification on social media.

Objectives/Research Questions

This study aims to examine how a highly visible virtual influencer constructs femininity on Instagram and how audiences respond to this representation. Specifically, it addresses the following research questions:

1. How do visual characteristics such as clothing, setting, and facial expression relate to engagement metrics (likes and comments)?
2. How is sexualization embedded in the visual presentation of a virtual influencer, and how is it rewarded by audience interaction?
3. How do audiences interpret and respond to the influencer's artificiality, body representation, and emotional expression in public comments?

Methods

A mixed-methods case study was conducted on the Instagram account @fit_aitana. The dataset includes all publicly available static image posts published between September 2024 and June 2025. Of 78 total posts, 69 static images were retained after excluding videos. Engagement data comprised 249,802 likes and 13,829 comments.

Quantitative analysis involved manual coding of visual variables, including clothing type, setting, and context, alongside engagement metrics. Facial expression analysis was conducted using the DeepFace Python library, classifying dominant emotions into seven basic categories. Statistical and visual analyses were performed to examine relationships between visual features, emotional expression, and engagement.

Qualitative analysis focused on public comments associated with each post. An inductive thematic coding approach was applied to identify recurring patterns in audience discourse. This analysis resulted in 14 codes that were later classified into three themes: (1) perceptions of artificial nature, (2) engagement with persona & community, and (3) value judgements & social reactions. Comments in all languages were included, and user identities were anonymized. The analysis prioritized interpretive depth and transparency over statistical generalization.

Results

Results show that posts featuring body-revealing clothing, particularly bikinis, consistently achieved the highest levels of engagement in terms of both likes and comments. Sportswear and underwear also performed strongly, while more covered or elegant outfits generated lower interaction. Similarly, leisure-oriented settings, especially beach environments, produced significantly higher engagement than everyday, urban, or sporty contexts. These findings indicate that sexualized visual cues and aspirational leisure imagery are systematically rewarded within Instagram's engagement economy, even when the body presented is explicitly artificial.

Facial expression analysis revealed a strong predominance of neutral expressions, with limited emotional variation. Neutral expressions were associated with higher likes but fewer comments, while more emotionally charged expressions were linked to increased discussion. Audience comments frequently criticized the lack of emotional realism, describing the influencer as emotionally "empty" or "unfocused".

Qualitative comment analysis revealed ambivalent and often contradictory audience responses. Many users engaged with the influencer as a social actor, expressing admiration, desire, and parasocial attachment, despite acknowledging her artificial nature. At the same time, users frequently highlighted anatomical inconsistencies and technical flaws, using them to question realism and authenticity. Sexualized and sexist remarks constituted a substantial portion of comments, mirroring patterns commonly observed with human female influencers. Alongside these responses, other users raised concerns about objectification, unrealistic beauty standards, and potential negative implications for mental health, particularly for young audiences.

Taken together, the results suggest that @fit_aitana operates less as a disruptive posthuman figure and more as an optimized digital product. Her representation reproduces conventional, gendered ideals of beauty and desirability, while audience engagement reflects a tension between attraction, critical awareness, and ethical unease. These findings highlight how algorithmic visibility mechanisms reward sexualized femininity and shape how virtual bodies are interpreted and consumed by online audiences.

Future Work

Future research should expand beyond single-case studies to include comparative analyses of multiple virtual influencers across gender, cultural contexts, platforms, and degrees of human realism. In particular, comparing female and male virtual influencers would allow researchers to examine whether sexualization, emotional expression, and visibility strategies are gendered, and whether similar visual cues produce different patterns of audience engagement and interpretation.

Such comparative work could help clarify whether female virtual influencers are disproportionately subjected to sexualized representation and scrutiny, and whether audiences respond differently to virtual bodies depending on gender. Experimental

designs could further isolate how exposure to gendered virtual influencer imagery affects viewers' attitudes, social comparison processes, and well-being, including differences in perceived authenticity, attractiveness, and emotional resonance.

Longitudinal studies may also examine how repeated exposure to idealized virtual bodies shapes audience perceptions over time, particularly among young or vulnerable users. Finally, future work could integrate audience surveys or interviews to better understand how users interpret gendered virtual representations and how these images influence expectations around beauty, identity, and embodiment in increasingly AI-mediated social environments.

References

- Carrillo-Durán, M. V., García García, M., & Corzo Cortés, L. (2024). Virtual influencers of human appearance as a form of online communication: the case of Lil Miquela and Lu do Magalu on Instagram. *Revista de Comunicación*, 23(1), 119–140. <https://doi.org/10.26441/RC23.1-2024-3453>
- Conti, M., Gathani, J., & Tricomi, P. P. (2022). Virtual Influencers in Online Social Media. *IEEE Communications Magazine*, 60(8), 86–91. <https://doi.org/10.1109/MCOM.001.2100786>
- Drenten, J., & Brooks, G. (2020). Celebrity 2.0: Lil Miquela and the rise of a virtual star system. *Feminist Media Studies*, 20(8), 1319–1323. <https://doi.org/10.1080/14680777.2020.1830927>
- Freberg, K., Graham, K., McGaughey, K., & Freberg, L. A. (2011). Who are the social media influencers? A study of public perceptions of personality. *Public Relations Review*, 37(1), 90–92. <https://doi.org/10.1016/j.pubrev.2010.11.001>
- Fowler, K., & Thomas, V. L. (2025). An investigation into the sexualization of human and virtual social media influencers. *European Journal of Marketing*. <https://doi.org/10.1108/EJM-04-2024-0259>
- Gambetti, R. C., & Kozinets, R. V. (2024). From killer bunnies to talking cupcakes: theorizing the diverse universe of virtual influencers. *European Journal of Marketing*, 58(13), 205–251. <https://doi.org/10.1108/EJM-12-2022-0948>
- Lee, Y. H., & Yuan, C. W. (2023). I'm not a puppet, I'm a real boy! Gender presentations by virtual influencers and how they are received. *Computers in Human Behavior*, 149. <https://doi.org/10.1016/j.chb.2023.107927>
- Liu, F., & Wang, R. (2025). Fostering Parasocial Relationships with Virtual Influencers in the Uncanny Valley: Anthropomorphism, Autonomy, and a Multigroup Comparison. *Journal of Business Research*, 186. <https://doi.org/10.1016/j.jbusres.2024.115024>

- Sands, S., Campbell, C. L., Plangger, K., & Ferraro, C. (2022). Unreal influence: leveraging AI in influencer marketing. *European Journal of Marketing*, 56(6), 1721–1747. <https://doi.org/10.1108/EJM-12-2019-0949>
- Oktan, A., & Yavas, B. (2024). Intertextuality in the Body Designs of AI-Based Virtual Influencers. *Rupkatha Journal on Interdisciplinary Studies in Humanities*, 16(03). <https://doi.org/10.21659/rupkatha.v16n3.02>
- Thomas, V. L., & Fowler, K. (2021). Close Encounters of the AI Kind: Use of AI Influencers As Brand Endorsers. *Journal of Advertising*, 50(1), 11–25. <https://doi.org/10.1080/00913367.2020.1810595>
-

2A.2 Algorithms for Academia: Mapping Users' Imaginaries of Academic Social Networks.

Guillaume Latzko-Toth (Laval University), Clémentine Fruchard (Université du Québec à Montréal), Elsa Fortant (INRS), Claudine Bonneau, Florence Millerand (Université du Québec à Montréal).

Introduction

Computational algorithms are ubiquitous, both in everyday life and in professional and academic contexts. Academic social network sites (ASNs)—such as Academia.edu, ResearchGate, and Mendeley, as well as platforms used for similar purposes (Google Scholar, LinkedIn, etc.)—are developing as modes of dissemination and recommendation of scientific knowledge (Gruzd et al, 2012; Hailu & Wu, 2021). Yet the algorithms that sort, recommend, and make visible scientific contributions and their authors remain surrounded by a certain “fuzziness,” on the one hand because their functioning is opaque and a source of “folk theories” (Siles et al., 2020), and on the other hand because they are the product of complex sociotechnical assemblages whose contours remain invisible and abstract. As they increasingly shape the ways in which research is conducted, it is important to uncover and question their growing mediation of scientific communication. We propose to study the “algorithmic imaginaries” that develop around these tools in order to better understand how these imaginaries influence the ways algorithms are used. Bucher (2017) introduced the concept of the “algorithmic imaginary” to theorize individual and collective relationships to algorithms. She defines it as “the way of thinking about what algorithms are, what they should be, how they work, and what these imaginations in turn make possible” (ibid: 40). Whether or not users understand how algorithms function, they experience their effects and develop imaginaries that shape their everyday practices and experiences.

Several bodies of work have examined the imaginaries developing around algorithmic systems in order to understand their influence on practices, including across different cultural contexts (Bucher, 2018; Siles, 2023). They describe how awareness and the perceived presence of algorithms affect, for example, listening and viewing behaviours (Zurovac et al, 2024; Prey et Estève-Del-Valle, 2024), or self-perceptions on dating apps (Parisi et Comunello, 2020). These studies also reveal the prevalence of dominant imaginaries, to which counter-imaginaries seek to respond (Kazansky et Milan, 2021). Research conducted from this perspective considers algorithmic power as constantly challenged by user practices, as users appropriate, circumvent, or even resist algorithmic prescriptions (Velkova et Kaun 2021; Siles, 2023). It also considers

algorithms as cultural entities, in the sense that they result from practices of design and computational development embedded in cultural, economic, and political contexts.

Objectives/Research Questions

The objective of the ongoing research presented in this paper is to describe and analyse the algorithmic imaginaries of ASN users in order to understand how these imaginaries shape their uses of these tools and the social practices in which such uses are embedded. In doing so, it aims to generate new knowledge about “ordinary” relationships to algorithms (Bucher, 2018), about the contexts in which users become aware of their existence, and about the meanings they attribute to them. This knowledge also seeks to contribute to the field of Critical Algorithm Studies, an emerging interdisciplinary field in the social sciences and humanities devoted to analysing the cultural, social, ethical, and political challenges raised by algorithmic systems. Work in this field rejects reducing algorithms to their purely technical dimension and instead seeks to understand them through their social embeddedness (Bucher, 2018; Gillepsie, 2014; Kitchin, 2017), as cultural objects situated within practices (Seaver, 2017).

By studying algorithmic imaginaries associated with ASNs, we seek to question, from a critical and interpretive perspective, their relationship to the usage frameworks of these devices, whether in terms of compliance with them or, potentially, their contestation.

We aim to identify dominant imaginaries, the nature of the relationships to algorithms that they engender, and, ultimately, to grasp their role in enabling—or limiting—the possibility of questioning these tools. The originality of this project lies in its analysis of the relationship between imaginaries and practices among ASN users (faculty members and students). In light of the issues raised by GenAI and recommendation algorithms, this research sheds new light on academic integrity and knowledge evaluation in relation to the impact of algorithms within the academic world.

Methods

In order to describe and analyse the algorithmic imaginaries of ASN users, we are conducting two types of interviews:

1) Individual interviews (N=30) with Canadian and French faculty members and graduate students who have developed different types of ASN usage (intensive, regular, occasional, non-use) and who come from various of disciplinary backgrounds. The interviews include a component involving discussion of texts and/or visuals produced by the participants (“rich images,” Bell et al, 2019; “short narratives,” Schellewald, 2022; Siles, 2023), as well as—for users—a “guided tour” component of their ASN

accounts inspired by “go-along” and “scroll-back” methods used in digital ethnography (Møller & Robards, 2019), during which participants verbalize their understanding of the algorithm.

2) Focus groups (N=2)—one composed of professors and the other of students—in which participants comment on preselected texts and visuals, making it possible to access collective imaginaries through the production and analysis of shared narratives.

These rich data are subjected to a thematic analysis informed by categorizations of algorithmic imaginaries proposed in the literature, including one put forward by Siles (2020).

Results

The preliminary analysis of the collected materials has already made it possible to support two hypotheses already outlined in the literature:

- 1) The experience of algorithmic systems takes shape through the development of dominant imaginaries;
- 2) These imaginaries give rise to certain types of relationships to algorithms that contribute to their naturalization and legitimization, thereby de facto limiting the possibilities for questioning or challenging the tools to which they relate.

The same analysis has also yielded a typology of individual and collective algorithmic imaginaries of ASNs, that we will be refining in the coming months.

Future Work

This research is part of a broader project in which we will also map the imaginaries of designers of recommendation algorithms intended for the academic world, and examine the relationships between the respective algorithmic imaginaries of users and designers of these systems. The second phase will therefore include interviews with developers conducted according to the same methodological protocol as that used with professors and students, supplemented by the analysis of three types of corpora:

-
- Online conversations on blogs, forums (e.g., Reddit, Stack Overflow), and software hosting and development platforms (e.g., Github), in order to access collective narratives (Bishop, 2019);
 - Discursive and content analysis of technical and promotional documents related to ASNs;

- Analysis of the tools themselves, which carry traces of designers' imaginaries (Schulz, 2023; Christin, 2020).
-

In order to grasp how imaginaries of design articulate with imaginaries of use, we will conduct a comparative analysis of individual and collective imaginaries among designers and users, with the aim of identifying shared features, points of distinction, and dominant imaginaries. To what extent do they converge or diverge? To what extent do these points of convergence and divergence help account for the uses and experiences of professional social network sites in the academic world? In doing so, the project seeks to address a significant gap in social studies of algorithms, which tend to focus on users' imaginaries to the detriment of those of designers.

It also aims to contribute to the development of methods for capturing algorithmic imaginaries. These methods may in turn make it possible to analyse the imaginaries associated with GenAI tools, which are also transforming academic work and the circulation of scholarly publications.

References

- Bell, S., Berg, T., Morse, S. (2019). Towards an Understanding of Rich Picture Interpretation. *Systemic Practice and Action Research*, 32, 601-614.
- Bishop, S. (2019). Managing visibility on YouTube through algorithmic gossip. *New Media & Society*, 21(11-12), 2589-2606.
- Bucher, T. (2017). The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. *Information, communication & society*, 20(1), 30-44.
- Bucher, T. (2018). *If... then: Algorithmic power and politics*. Oxford University Press.
- Christin, A. (2020). The ethnographer and the algorithm : Beyond the black box. *Theory and Society*, 49(5 6), 897 918. <https://doi.org/10.1007/s11186-020-09411-3>
- Gillespie, T. (2013). *The Relevance of Algorithms*. <https://doi.org/10.7551/mitpress/9780262525374.003.0009>
- Gruzd, A., Staves, K., & Wilk, A. (2012). Connected scholars: Examining the role of social media in research practices of faculty using the UTAUT model. *Computers in Human Behavior*, 28(6), 2340–2350. <https://doi.org/10.1016/j.chb.2012.07.004>

- Hailu, M., & Wu, J. (2021). The Use of Academic Social Networking Sites in Scholarly Communication: Scoping Review. *Data and Information Management*, 5(2), 277–298.
- Kazansky, B., & Milan, S. (2021). “Bodies not templates”: Contesting dominant algorithmic imaginaries. *New Media & Society*, 23(2), 363–381.
<https://doi.org/10.1177/1461444820929316>
- Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society*, 20(1), 14–29.
- Møller, K. & Robards, B. (2019). Walking through, going along and scrolling back: Ephemeral mobilities in digital ethnography. *Nordicom Review*, 40 (Special Issue 1): 95–109.
- Parisi, L., Comunello, F. (2020). Dating in the time of “relational filter bubbles”: Exploring imaginaries, perceptions and tactics of Italian dating app users. *The Communication Review*, 23(1), 66–89.
<https://doi.org/10.1080/10714421.2019.1704111>
- Prey, R., Esteve-Del-Valle, M. (2024). The algorithmic network imaginary: How music artists understand and experience their algorithmically constructed networks. *The Information Society*, 40(1), 18–31.
<https://doi.org/10.1080/01972243.2023.2274070>
- Schellewald, A. (2022). Theorizing “Stories About Algorithms” as a Mechanism in the Formation and Maintenance of Algorithmic Imaginaries. *Social Media + Society*, 8(1).
- Schulz, C. (2023). A new algorithmic imaginary. *Media, Culture & Society*, 45(3), 646–655. <https://doi.org/10.1177/01634437221136014>
- Seaver, N. (2017). Algorithms as culture : Some tactics for the ethnography of algorithmic systems. *Big Data & Society*, 4(2), 2053951717738104.
<https://doi.org/10.1177/2053951717738104>
- Siles, I., Segura-Castillo, A., Solís, R., & Sancho, M. (2020). Folk theories of algorithmic recommendations on Spotify : Enacting data assemblages in the global South. *Big Data & Society*, 7(1), <https://doi.org/10.1177/2053951720923377>
- Siles, I. (2023). *Living with Algorithms. Agency and User Culture in Costa Rica*. MIT Press.

Velkova, J., and Kaun, A. (2021). Algorithmic Resistance: Media Practices and the Politics of Repair. *Information, Communication & Society*, 24(4), 523-540.

Zurovac, E., Boccia Artieri, G., Donato, V. (2024). Tales of Visibility in TikTok: The Algorithmic Imaginary and Digital Skills in Young Users. In D. Radovanović (Éd.), *Digital Literacy and Inclusion: Stories, Platforms, Communities* (p. 113 126). Springer International Publishing. https://doi.org/10.1007/978-3-031-30808-6_8

2A.3 When Precarity Goes Viral: Entrepreneurial Ideology, Algorithmic Visibility and the Cross-Platform Reframing of Formal Employment.

Jessica Souza, Rosana Pinheiro-Machado (University College Dublin).

Introduction

Narratives that emphasize entrepreneurial autonomy, individual freedom, and self-responsibility have been exponentially growing in social media platforms, part of a broader political transformation which includes the rising appeal of conservative and illiberal worldviews. Within this context, Brazil's digital scene reveals a striking reframing of formal employment as it is increasingly portrayed as a symbol of hardship, stagnation, and precarity amid a broader labor crisis. Viral memes about everyday work life have gained traction on Instagram, TikTok, and YouTube. The acronym CLT (Consolidation of Labour Laws), used as equivalent to formal employment, has become less likely to be invoked in its original legal sense, instead coming to symbolize exhaustion, poor working conditions, and frustration.

The legitimate online complaints about working conditions under formal employment intersects with narratives pushed by influencers and digital coaches, who profit from growing dissatisfaction while presenting entrepreneurship as a path to financial freedom. This article investigates how formal labor is represented in highly visible social media content and analyzes how social media algorithms, influencer entrepreneurial culture, and political subjectivity converge in shaping contemporary labor discourse. It examines how digital content reconfigures the meaning of formal employment, presenting it as synonymous with precarity and failure, while entrepreneurial narratives gain prominence, reinforcing hyper-individualistic worldviews and solutions. The analysis focuses on two main datasets: the most viewed posts across Instagram, TikTok, and YouTube, and content produced by a mapped network of digital marketing coaches known for promoting entrepreneurial content. The analysis identifies a rhetorical displacement of collective struggle and systemic critique, whereby hope is rechanneled away from institutional and labor-based solutions and toward individualized narratives of resilience and self-enterprise.

Objectives/Research Questions

This article investigates how formal labor is represented in highly visible social media content. Specifically, it seeks to: a) Identify the narratives associated with CLT across Instagram, TikTok, and YouTube; b) compare discourses produced by general users with those amplified by digital marketing influencers promoting entrepreneurial and hyper-individualistic worldviews; and c) examine how algorithmic amplification and

platform-specific affordances shape the visibility and framing of content about formal employment.

Methods

Data was collected across the three major social media platforms Instagram, TikTok and YouTube using the acronym CLT to search for content related to formal employment in Brazil. Data access and collection were carried out through the official research tools provided by each platform: Meta Content Library, YouTube Data API, and TikTok Researcher API.

The methodological approach employs a mixed-methods design of qualitative and quantitative steps. In the first stage, a qualitative analysis of the most viewed Instagram content in the last five years identified visual patterns, recurring themes, and dominant narrative clusters. The second stage used a semi-inductive, data-driven approach on large datasets extracted from TikTok and Youtube , employing word frequency and n-gram analyses to identify semantic patterns, which informed the construction of thematic categories. This hybrid approach enabled cross-platform analysis, improving interpretability over fully automated models, and allowed integration of inductive discovery with theoretical coherence, aligning computational analysis with sociological insights.

The thematic categories were used to guide a large quantitative analysis, in which large language model (LLM) techniques were applied to assess the prevalence of these narratives and quantify their diffusion across platforms. Collecting data in real time and then extending the analysis retrospectively (2021 - 2025) made it possible to identify the point at which CLT-related content began to circulate at scale.

Finally, by applying this two-stage analytical framework within influencer clusters previously identified through digital ethnography, the study compares the discourses produced by general users with those promoted by a network of digital marketing influencers. This approach allows us to trace the storylines amplified by entrepreneurial actors and to examine how discontent with formal employment is strategically leveraged to promote individualistic and anti-collective interpretations of work.

Results

Preliminary results indicate a consistent pattern in CLT-related content across the three platforms, despite differences in content formats and distribution dynamics. By combining real-time data collection with a retrospective approach based on diminishing query returns, the analysis made it possible to identify that, in earlier periods, CLT was predominantly associated with informational content about workers' rights, which appeared in relatively limited volume. However, during the pandemic on Instagram and in the post-pandemic period on TikTok, both the volume of posts and the

thematic scope expanded substantially, with CLT increasingly associated with humorous content and narratives contrasting formal employment with self-employment (PJ).

In this later phase, the overwhelming majority of highly visible content consists of memes satirizing poor working conditions, establishing a recurring imagetic pattern in which formally employed workers are portrayed in overcrowded public transportation, exposed to heat and physical exhaustion, or experiencing mental health deterioration due to toxic employers, work overload, low wages, and insufficient time for rest and recovery. This pattern is followed by interview-style videos that explicitly debate the comparative advantages of CLT and temporary contract arrangements. Some of the most viewed videos highlight the income of interviewees engaged in informal activities, such as children selling snacks after school or women working in adult content production, resulting in viral reach accompanied by numerous perplexed comments from users on how such forms of work could offer greater economic advantages than formal employment. A substantial portion of the content also consists of personal testimonies from individuals who report having left CLT employment and express satisfaction with transitioning to entrepreneurship, citing higher earnings and autonomy, alongside influencer-produced content offering guidance on how to pursue similar paths. In contrast, only a marginal number of posts provide informational guidance on labor rights, alongside a small but recent set of news articles highlighting that CLT employment seems to no longer be perceived as desirable among younger generations, who increasingly favor independent work.

Future Work

Future research could investigate how younger generations are responding to this reframing of formal labor, in order to understand which segments of the population align with these narratives. A second avenue would explore how these discourses intersect with perceptions of higher education, particularly regarding its perceived value relative to entrepreneurial opportunities. Insights from the influencer content analyzed in the digital ethnography phase of this study suggest that, within these narratives, entrepreneurship is positioned as the ultimate goal, and formal education is increasingly reframed as less central to financial security than previous generations assumed. This raises questions about how such narratives may shape decisions regarding degree pursuit, career planning, and the broader social valuation of formal qualifications.

This article builds on previous literature on the political and economic logics of platforms to examine how CLT becomes a symbolic battleground across diverse platforms, highlighting how social media infrastructures and influencer rhetorics converge to undermine collective conceptions of labor in favor of entrepreneurial hyper-individualism. Research demonstrates that platforms shape user experience and

political discourse through algorithms that govern visibility and engagement (Soriano & Gaw, 2022; Bhandari & Bimo, 2022), influencing users' preferences and perceptions (Lewandowsky et al., 2020). These algorithms create feedback loops that amplify dominant narratives while marginalizing dissenting ones (Soriano & Gaw, 2021). As a result, the online debate is influenced and constrained by power structures that prioritize commercial interests (Lewandowsky et al., 2020).

In this context, digital influencers play a central role in reshaping labor discourse. In Brazil, the rapid migration of labor-related activity to social media, accelerated by the pandemic, has resulted in platforms such as Instagram being dominated by a concentrated network of digital marketing influencers who actively promote entrepreneurial narratives while undermining collective interpretations of labor and discouraging collective efforts to improve working conditions (Pinheiro-Machado et al., 2024). Influencers visibility in these platforms is enhanced by large-scale funding for advertisements, reinforcing their persuasive power. At the same time, influencers are perceived 'authentic' opinion leaders, subtly shaping beliefs on science, health, and politics (Hasell & Chinn, 2023), while everyday content embeds ideological values, reinforcing dominant discourses under the guise of neutrality (Bouvier & Way, 2021). These dynamics help explain how commercial incentives, platform visibility, and influencer credibility converge to normalize hyper-individualistic narratives of labor and success.

References

- Bhandari, A., & Bimo, S. (2022). Why's everyone on TikTok now? The algorithmized self and the future of self-making on social media. *Social media+ society*, 8(1), 20563051221086241.
- Bouvier, G., & Way, L. C. (2021). Revealing the politics in "soft", everyday uses of social media: the challenge for critical discourse studies. *Social Semiotics*, 31(3), 345-364.
- Hasell, A., & Chinn, S. (2023). The political influence of lifestyle influencers? Examining the relationship between aspirational social media use and anti-expert attitudes and beliefs. *Social Media+ Society*, 9(4), 20563051231211945.
- Lewandowsky, S., Smillie, L., Garcia, D., Hertwig, R., Weatherall, J., Egidy, S., ... & Leiser, M. (2020). Technology and democracy: Understanding the influence of online technologies on political behaviour and decision-making.
- Pinheiro-Machado, R., Matheus, J., & Frid, M. (2024). Social media as a digital labour platform: assessing the social, cultural, and political impacts of labour market migration to Instagram. DeepLab.
- Soriano, C. R. R., & Gaw, F. (2022). Platforms, alternative influence, and networked political brokerage on YouTube. *Convergence*, 28(3), 781-803.

2A.4 How do Terms of Service Influence Social Media Users Dynamics - A Privacy Anxiety Perspective.

Jingyuan Liu (Boston University).

Introduction

In October 2024, the social media platform X (formerly Twitter) announced an upcoming update of Terms of Service (ToS), giving permission to authorize the use of user-generated content for machine learning training and generative AI models. The terms explicitly state, 'You agree that this license includes the right for us to analyze text and other information you provide and to otherwise provide, promote, and improve the Services, including for use with and training of our machine learning and artificial intelligence models, whether generative or another type'.

Even though AI related clauses are not a new issue nowadays, this update still triggered heated public discussion and widespread opposition. The core issue is not only the controversial update to Platform X's terms of service, but also the systemic anxiety exacerbated by the platform's own architecture. Artists, writers, scholars, and developers expressed critical concerns about user privacy and data governance. Many interpreted the change more than a contractual revision but a severe violation of consent, ownership, and creative autonomy. Also, by structurally prioritizing highly stimulating content, the recommendation algorithm effectively commodifies users' anxieties, amplifying signals of distress while filtering out rational, technical analysis, thus hindering community connections and communication. Consequently, the controversy quickly escalated to behavioral shifts, including decreased engagement, post deletions, and migration to alternative platforms like BlueSky.

Objectives/Research Questions

To investigate the systemic collapse of trust on X following the ToS update, this thesis addresses three interconnected questions:

RQ1: The Affective Formation: How does privacy anxiety emerge following the new update, and how do users collectively interpret this policy change?

RQ2: Diffusion and Amplification Mechanism: How does privacy anxiety diffuse within and across user communities, and how do platform algorithms amplify these pre-existing affective emotions?

RQ3: Behavioral Consequence: How does this amplified anxiety shape community engagement, interaction dynamics, and user retention?

Methods

This study adopts a mixed-methods research design grounded in methodological triangulation. Privacy anxiety acts as an intermediate among value interpretation, affective response, and structural amplification. It is neither a purely psychological state nor a directly observable behavioral outcome in this study. No single method is sufficient to capture these strangled processes, thus, requiring both qualitative and quantitative methods.

Qualitative netnography examines how the policy is decoded and how privacy-related concerns are narratively framed. Construction of a community-based taxonomy and netnography observation are utilized to understand privacy concerns and value among user groups. A specialized anxiety lexicon and a value-based codebook are also provided.

Quantitative analysis then operationalizes these insights into measurable indicators, examining large-scale patterns in anxiety expression and engagement intensity. Data was collected via the X API through a three-stage temporal window. It covers from September 15 to December 16, 2024, including the month preceding the announcement, the immediate reaction period following the announcement, and the post-deployment phase. Using a set of quantitative metrics, including anxiety score, topic value, engagement, Risk Attention Index (RAI), Risk–Nonrisk Engagement Ratio (RER) and BERT-based classifier and GPT-assisted labeling with scoring, it captures broader patterns across communities and over time.

Results

The study finds that the ToS update acted as a risk signal that was transformed into a collectively reinforced phenomenon through social and digital amplification.

Risk Activation by Value Misalignment (RQ1): A 20% value shift towards IP-negative framing from the platform ToS was discovered while privacy-sensitive users kept placing great emphasis on authorship and privacy-oriented value. Based on an overall Social Amplification of Risk Framework, privacy anxiety is conceptualized as a value-induced risk signal, whose activation is explained through a Value Sensitive Design perspective. For creators, the new ToS is not merely a policy update but a symbolic violation of creative dignity. Within this value misalignment gap, privacy anxiety translates this simple value conflict into a socially affective form that can be communicated, shared and amplified.

Amplification Mechanism (RQ2): Our analysis revealed an non-linear, self-reinforcing mechanism of risk amplification, characterized by significant community-based sensitivity and algorithmic mediation.

A clear non-linear, inverted U-shaped pattern is discovered between anxiety score and initiative engagement. Engagement peaks at moderate anxiety levels (0.3-0.6),

increasing nearly fourfold compared to the baseline level, indicating anxiety functioning as a conditional driver of amplification. Risk-related posts also receive 3.77 times as much engagement as non-risk posts.

Regression analysis providing preliminary evidence of possible structural ‘algorithmic subsidy’ for privacy-risk content. While engagement remains largely insensitive to anxiety in neutral or non-risk contexts, there is a strong, significant positive association emerging within privacy-risk posts. While other factors such as user preference and network density contribute to engagement, this differential distribution points toward an algorithmic logic that treats anxiety as a high-value signal for distribution when associated with platform risk.

Network analysis also reveals critical structural isolation and amplification is also unevenly distributed across communities. Interaction graph shows an overwhelming dominance of self-loops within communities. While Creators occupied a ‘High-Attention, High-Amplification’ zone, Scholars and Developers exhibited high attention but limited reach. The amplification exhibits a relatively parallel and internally reinforced amplification mode, rather than chain-like transmission.

Engagement and Migration Outcome (RQ3): In the short term, there was a measurable decline in engagement as users adopted defensive actions. More significantly, half of the identified users in the privacy-sensitive clusters signaled a clear intent to leave and shared handles for decentralized alternatives like BlueSky. However, within the study’s recording timeline, only 20% of those who declared their intent actually completed the transition.

Future Work

Longitudinal Analysis: This study primarily captured the immediate impact of the policy updates in 2024. However, with the announcement of X platform's new ToS effective on January 15, 2026, further study whether there is a cumulative or fatigue effect on privacy anxiety could be done. Future work will adopt a longitudinal perspective, comparing the similarities and differences in user responses during the two updates in 2024 and 2026, and exploring whether the willingness to migrate and algorithmic risks have become normalized over time.

Causal Inference Analysis to Verify Algorithmic Mediation: The algorithmic tax found in current regression models is primarily based on correlation analysis. To more rigorously disentangle algorithmic intervention from spontaneous user behavior, future research will introduce causal inference techniques. We will further adopt Regression

Discontinuity Design or Synthetic Control Methods to quantitatively assess changes in algorithmic distribution logic. Furthermore, by including covariates such as historical interaction rates, causal contribution of algorithms will be more precisely calculated

Post-Migration Dynamics: Given that this study observed a 20% actual migration conversion rate, future research should track the post-X experiences of these migrating users. By analyzing how these creators rebuild communities on decentralized platforms such as BlueSky, we can further explore whether they successfully sustain the creative autonomy and how the community develops.

References

- Friedman, B., Kahn, P.H., Borning, A., Huldtgren, A. (2013). Value Sensitive Design and Information Systems. In: Doorn, N., Schuurbiers, D., van de Poel, I., Gorman, M. (eds) Early engagement and new technologies: Opening up the laboratory. Philosophy of Engineering and Technology, vol 16. Springer, Dordrecht.
https://doi.org/10.1007/978-94-007-7844-3_4
- Kasperson, R.E., Renn, O., Slovic, P., Brown, H.S., Emel, J., Goble, R., Kasperson, J.X. and Ratick, S. (1988), The Social Amplification of Risk: A Conceptual Framework. Risk Analysis, 8: 177-187. <https://doi.org/10.1111/j.1539-6924.1988.tb01168.x>
- Kozinets, R.V. (2010) Netnography: Doing Ethnographic Research Online. Sage Publications, London.
- Fiesler, C., & Hallinan, B. (2018). “We are the product”: Public reactions to online data sharing and privacy controversies in the media. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18) (Paper 53, pp. 1–13). Association for Computing Machinery.
<https://doi.org/10.1145/3173574.3173627>
-

Session 2B - Health and Wellbeing (1)

Day 1 · Mon 13 July · 11:30–13:00 · Seminar Suite A, ARC

2B.1 “I’m Crushed, Ashamed, and Shattered”: Moral Sentiment, Network Position, and Recovery Communities on Reddit.

Kimberly Glasgow (Johns Hopkins University Applied Physics Laboratory).

Introduction

The opioid epidemic in the United States is an ongoing public health crisis, causing devastating harms at the individual, interpersonal, community, and societal levels. Morality-based discourse around addiction and treatment, particularly medication-assisted treatment (MAT), is pervasive in society. It informs attitudes toward harm reduction efforts, treatment approaches, clinical care, and the self-image of substance users. These moral framings can stigmatize individuals, shape and limit access to supportive others, and contribute to marginalization and stigmatization (Ahern et al., 2007; Frank & Nagel, 2017).

Despite changes in how addiction has been theorized over time, addiction remains deeply moralized. Early models framed addiction as a moral-pathological phenomenon rooted in vice, weakness, and defective character (Harding, 1986). More recent models situate addiction as a chronic, relapsing condition (Leshner, 1997). Nevertheless, negative moral assessments among the public, families of substance users, and individuals themselves are common. Individuals struggling with addiction commonly feel deep shame over their perceived moral failures to control their behavior and conform to their moral self-view (Flanagan, 2013; Ohayon & Ronel, 2025). Moral judgments surrounding treatment may stigmatize MAT as “replacing one addiction with another,” shaping decisions about treatment access and acceptability (Frank, 2011; Frank & Nagel, 2017).

Social support from online communities can play a key role in recovery, particularly for individuals who lack access to supportive offline networks (Granfield & Cloud, 2001; Bluc et al., 2017). Reddit has emerged as a platform in which self-organized communities, known as subreddits, can discuss sensitive, stigmatized health conditions (De Choudhury & De, 2014). Multiple subreddits focus on opioid use disorder (OUD), recovery, and harm reduction, enabling discussion of such topics under conditions of pseudonymity. Within these communities, members share

experiences of trauma, suffering, and growth; ask questions about treatment and daily life; vent frustration; encourage and comfort one another; and shape community norms (D’Agostino et al., 2017; Gauthier et al., 2022).

This work examines four communities through the lens of Moral Foundations Theory, which conceptualizes morality as comprising a small set of fundamental dimensions that shape judgments of right and wrong (Graham et al., 2011). These subreddits differ in philosophies and approaches to recovery, including the use of medication-assisted treatment. Drawing on one year of activity, encompassing hundreds of thousands of posts and comments from tens of thousands of authors, we analyze how moral sentiment is expressed within and structured by patterns of interaction among community members. We examine both the morality expressed in specific discourse exchanges (discussion threads) and the networks of human interaction that emerge.

Objectives/Research Questions

This work examines moral sentiment within online communities focused on opioid use disorder (OUD) recovery and harm reduction. Drawing on Moral Foundations Theory, we characterize the moral tenor of discourse in these communities and examine how moral expression varies across communities that differ in their focus, membership, scale, and approaches to treatment (Graham et al., 2011).

In addition to examining discourse content, we treat each subreddit as a social network of interacting authors and examine how moral sentiment is expressed through patterns of interaction, as well as how community members respond to morally expressive content. In particular, we examine how moral language is expressed by more central or prominent actors within these networks, given their potential role in shaping norms, engagement, and community dynamics (Borgatti & Brass, 2019).

This work addresses the following research questions:

1. How is morality expressed in online OUD recovery communities?
2. Do recovery communities differ in their moral discourse?
3. How do central actors in these communities express morality in their interactions with others?

Methods

We investigated four active recovery subreddits focused on recovery and harm reduction related to opioid use disorder (OUD): r/Methadone, r/Suboxone, r/OpiatesRecovery, and r/NarcoticsAnonymous. Two of these communities focus on medication-assisted treatment (r/Methadone and r/Suboxone), one is more generally focused on recovery (r/OpiatesRecovery), and the fourth has a more 12-step-oriented approach (r/NarcoticsAnonymous). These subreddits were large, active, long-running,

and publicly accessible, and span a range of approaches to treatment with associated moral implications.

Submissions and comments to these four subreddits from January 1 to December 31, 2022 were examined. Data were retrieved using pushshift.io (Baumgartner et al., 2020) and include over half a million posts from nearly 37,000 unique authors.

To identify moral content, we applied supervised classification models to detect the presence or absence of ten moral foundations derived from Moral Foundations Theory: Care, Harm, Fairness, Cheating, Loyalty, Betrayal, Authority, Subversion, Purity, and Degradation (Graham et al., 2011). Given the scale of the data, human annotation was not feasible. Instead, models were trained using previously annotated corpora, including the Moral Foundations Twitter Corpus (Hoover et al., 2020), and applied to Reddit submissions and comments. Posts could express multiple moral foundations or none. Classifier performance was originally assessed on held-out training data, with precision, recall, and F1 scores generally in the high 0.8 to 0.9 range, and further evaluated through manual review of a sample of Reddit posts to confirm appropriate transfer to this domain, recognizing that moral foundations annotation is a task for which trained human annotators do not always agree (Hoover et al., 2020).

Reddit submissions and comments were used to construct networks representing interactions between authors within each subreddit. In these networks, nodes represent authors and directed edges represent reply relationships. Edges were enriched using moral foundations classification outputs to characterize the moral content of interactions. Network analyses considered network properties, patterns of bidirectional interaction, and actor position and centrality (Freeman, 1979). After consideration of several centrality metrics, centrality was operationalized using degree, as well as indegree and outdegree, to capture how many distinct users an author replied to and was replied to by.

Engagement with moral and nonmoral discourse was examined using a platform-specific measure (upvotes) and interaction-based measures, including the number of direct replies and total discussion thread size.

Results

In this context, Care captures expressions of support and concern for others; Cheating and Degradation reflect moral judgments related to rule-breaking, relapse, contamination, and shame; and Loyalty and Purity capture in-group commitment and abstinence- or spiritually oriented framing common in some recovery communities (Graham et al., 2011).

Moral content occurs frequently across all four recovery subreddits, appearing in a substantial proportion of posts and discussion threads. Both virtue and vice moral foundations are expressed across communities, though rates vary by subreddit. For the

medication-assisted treatment–oriented communities r/Methadone and r/Suboxone, moral content is somewhat less common, which may reflect their focus on providing factual information, while r/OpiatesRecovery and r/NarcoticsAnonymous exhibit higher overall levels of moral expression.

Consistent with expectations, Care, Cheating, Degradation, Loyalty, and Purity were the most salient moral foundations across communities. Degradation and Cheating were closely tied to discourse around stigma, relapse, and moral judgment, while Loyalty and Purity were more strongly associated with abstinence-oriented recovery framing (Frank, 2011).

The moral profiles of recovery communities were not monolithic but reflected differences in community focus and structure. r/NarcoticsAnonymous, a 12-step–oriented community, exhibited elevated expression of loyalty and purity, alongside care. In contrast, r/Methadone and r/Suboxone showed higher relative expression of the vice moral foundations of degradation and cheating, and lower use of loyalty and purity language. r/OpiatesRecovery occupied an intermediate position, with elevated care expression and lower rates of purity and ingroup loyalty language.

Future Work

Our work examined how online communities organized around opioid use disorder and recovery use moral language in discourse and interaction, both to provide care and support and to engage with potentially stigmatizing framings. Future work could extend this analysis longitudinally to examine how moral expression and network position and structure evolve over time, particularly in relation to sustained participation or disengagement. Further qualitative analysis of selected interactions could provide deeper insight into how vice-oriented moral language functions in context, including its role in validating shared experiences of stigma or enforcing community norms. Additional work might examine how moderation practices shape moral discourse, or how exposure to different moral framings relates to members' perceptions of support and stigma. More broadly, integrating moral sentiment with relational and temporal features may help clarify how online recovery communities can provide morally attuned and effective support.

References

- Ahern, J., Stuber, J., & Galea, S. (2007). Stigma, discrimination and the health of illicit drug users. *Drug and Alcohol Dependence*, 88(2), 188–196.
- Baumgartner, J., Zannettou, S., Keegan, B., Squire, M., & Blackburn, J. (2020). The Pushshift Reddit dataset. *Proceedings of the International AAAI Conference on Web and Social Media*, 14, 830–839. <https://doi.org/10.1609/icwsm.v14i1.7347>

- Bluic, A.-M., Best, D., Iqbal, M., & Upton, K. (2017). Building addiction recovery capital through online participation in a recovery community. *Social Science & Medicine*, 193, 110–117. <https://doi.org/10.1016/j.socscimed.2017.09.050>
- Borgatti, S. P., & Brass, D. J. (2019). Centrality: Concepts and measures. In *Social networks at work* (pp. 1–14). Routledge.
- De Choudhury, M., & De, S. (2014). Mental health discourse on Reddit: Self-disclosure, social support, and anonymity. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 71–80.
- D’Agostino, A. R., Optican, A. R., Sowles, S. J., Krauss, M. J., Lee, K. E., & Cavazos-Rehg, P. A. (2017). Social networking online to recover from opioid use disorder: A study of community interactions. *Drug and Alcohol Dependence*, 181, 5–10. <https://doi.org/10.1016/j.drugalcdep.2017.09.010>
- Flanagan, O. (2013). The shame of addiction. *Frontiers in Psychiatry*, 4, Article 120. <https://doi.org/10.3389/fpsy.2013.00120>
- Frank, D. (2011). The trouble with morality: The effects of 12-step discourse on addicts’ decision-making. *Journal of Psychoactive Drugs*, 43(3), 245–256. <https://doi.org/10.1080/02791072.2011.605706>
- Frank, L. E., & Nagel, S. K. (2017). Addiction and moralization: The role of the underlying model of addiction. *Journal of Bioethical Inquiry*, 14(1), 129–139. <https://doi.org/10.1007/s12152-017-9307-x>
- Freeman, L. C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks*, 1(3), 215–239.
- Gauthier, R. P., Costello, M. J., & Wallace, J. R. (2022). “I will not drink with you today”: A topic-guided thematic analysis of addiction recovery on Reddit. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3491102.3502076>
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, 101(2), 366–385. <https://doi.org/10.1037/a0021847>
- Granfield, R., & Cloud, W. (2001). Social context and “natural recovery”: The role of social capital in the resolution of drug-associated problems. *Journal of Drug Issues*, 36(1), 1543–1570. <https://doi.org/10.1081/JA-100106963>
- Hoover, J., Portillo-Wightman, G., Yeh, L., Havaladar, S., Davani, A. M., Lin, Y., Kennedy, B., Atari, M., Kamel, Z., Mendlen, M., Moreno, G., Park, C., Chang, T. E., Chin, J., Leong, C., Leung, J. Y., Mirinjian, A., & Dehghani, M. (2020). Moral Foundations Twitter Corpus: A collection of 35k tweets annotated for moral sentiment. Group

Processes & Intergroup Relations, 23(8), 1057–1071.

<https://doi.org/10.1177/1948550619876629>

Leshner, A. I. (1997). Addiction is a brain disease, and it matters. *Science*, 278(5335), 45–47.

Ohayon, S., & Ronel, N. (2025). The moral voice in addiction and the self. *Journal of Substance Abuse Treatment*, 170, Article 209624.

<https://doi.org/10.1016/j.josat.2025.209624>

2B.2 Information Wellbeing: Balancing News Knowledge and Mental Health.

Inhwa Song, Manoel Horta Ribeiro, Marianne Aubin Le Quere (Princeton).

Introduction

In a democratic society, people need to be normatively informed about current events to take civic action like voting, protesting, and organizing. Higher awareness of current news events and availability of news is associated with desirable civic behaviors, including less partisan voting, and higher civic awareness (Boulianne, 2016; Delli Carpini et al., 2018). Relatedly, ignoring emotional news about civic issues can cause people to ascribe these issues less importance moving forward (Paluck et al., 2017).

At the same time, engagement-driven algorithmic news spread on social media and the 24-hour news cycle have given rise to the perception that people have become ‘overloaded’ with information. During the Covid-19 pandemic, multiple studies found that high exposure to the news could lead to heightened stress and anxiety (Stainback et al., 2020; Strasser et al., 2022; Zappia & Makri, 2022). In extreme cases, feeling such news anxiety can cause people to detach from news altogether, in turn reducing long-term informedness and civic capacity (Schäfer et al., 2023; Toff & Nielsen, 2022).

In today’s information environment, how can people remain normatively informed while preserving their mental health? This work seeks to define a new concept, information wellbeing, which is the state when someone is normatively informed about current events while limiting news anxiety. We first tackle this question qualitatively, through 14 semi-structured interviews of U.S. participants who have high objective news knowledge and low news anxiety. To identify these participants, we constructed a survey to identify both normative news knowledge and neutral to positive individual news- anxiety. We survey people on Prolific, to identify interview participants who have news knowledge in the top 25% of participants, and also feel neutrally or positively about their relationship with the news. In these interviews, we discuss participants’ online and offline information environments, and seek to understand how they emotionally relate to news items. We identify four main themes that distinguish people with high information wellbeing: social engagement with news, active engagement with news, effective personal coping skills, and stable characteristics. We propose next steps to validate that these behaviors are indeed associated with higher information wellbeing, and can be employed to increase everyday news knowledge without sacrificing wellbeing.

Objectives/Research Questions

Prior work largely examined news consumption and emotional wellbeing through divergent lenses and priorities. Research in journalism and political communication

emphasized being informed at a civic level, often focusing on democratic participation (Boulianne, 2016; Poindexter & McCombs, 2001). In this perspective, news anxiety or avoidance are typically framed as individual barriers that endanger the collective democratic process (Toff & Nielsen, 2022). As a result, many studies focus on sustaining or increasing engagement with information, such as by making news more approachable through rewriting headlines to be more pleasant (Wessel et al., 2025).

In contrast, research on digital wellbeing and mental health focuses on the psychological and emotional consequences of information exposure. A body of work demonstrates that frequent or intensive news consumption—particularly during periods of crisis or political uncertainty—is associated with heightened stress and anxiety (Yoon et al., 2021). Within this body of work, news is often framed as an affective stressor. As a result, prior research has focused on self-regulation strategies for managing news exposure, such as disabling notifications, setting time limits on news consumption, or avoiding news altogether (Fullana et al., 2020; Schäfer et al., 2023; Zappia & Makri, 2022). Many works focused explicitly on Covid-19 or other individual news events (Silver et al., 2013).

These studies set up the idea of a tradeoff between collective good through news knowledge versus individual harms of news anxiety. Multiple studies have tried to quantify this tradeoff broadly: (Bunn & Farmer, 2021) found no relationship between self-reported news exposure and anxiety across surveys; yet (Boukes & Vliegenthart, 2017) identified that self-reported exposure to hard TV news can negatively impact wellbeing over time. A Bluesky study found that users exposed to news feeds were significantly more likely to post in ways that indicate depression, stress, or anxiety (Pal et al., 2026).

In this work, we undertake a qualitative study of people with high information wellbeing: people who appear to successfully maintain high levels of both informedness and wellbeing. These perspectives are overlooked by prior survey and experimental research; consequently, while we have information about the population-level tradeoffs of staying informed, we lack empirical insights into how information wellbeing can be constructed and maintained in everyday life. Crucially, and unlike prior work, we focus on people's news knowledge (how well users know what is happening in the news) rather than their news exposure (the extent to which users report following the news), since news knowledge is what fosters an informed population and therefore best stands to uphold democratic societies.

Accordingly, we ask the research question: What practices, attitudes, and characteristics constitute and sustain people with high information wellbeing?

Methods

Our method consisted of two steps: a screener survey to identify individuals with high information wellbeing, and 45-minute follow-up interviews. All participants were

recruited using the Prolific platform, and were paid \$13.55/hour for screeners and \$26.67/hour for interviews.

For the news knowledge scale, we use a method from (Angelucci & Prat, 2024), which focuses on people's ability to differentiate between true news events and false but plausible news events. In practice, we constructed a news knowledge scale that consisted of five stories spanning politics, health/science, technology, global affairs, and business. Each of these had one true and one false story, presented in a random order. For each story, the participant could indicate whether the story "definitely didn't happen," "probably didn't happen," "probably did happen," or "definitely did happen." We conducted this survey in two waves, one December 2025 (N = 125 screener responses) and one in January 2026 (N = 99 screener responses), updating the survey items for current events between each wave. The participants' final score was then calculated as the total score for the true stories, minus the total score for the false stories. For each wave, we calculated the top 25% of participants by score, which we considered to have high objective news knowledge.

Our news anxiety metric was based on prior work from (Betakova et al., 2025). Specifically, we used their survey items to measure news negativity, negative physical reactions to news, and news overload. Based on pilot interviews, we also added the items "I feel pressure to keep up with the news even when I don't want to" and "I find myself skipping past news content to avoid feeling anxious." We consider anyone to have limited news anxiety if their average responses across these eight items averaged neutral or better.

We consider all participants who (a) were in the top quartile of news knowledge compared with other survey respondents and (b) had low news- anxiety as having high information wellbeing. These participants were then marked as eligible for our interviews. Of all 224 participants who completed the screener survey, 56 (25%) were considered in the top quartile of news knowledge, and 94 (41.9%) gave answers indicating low news anxiety. In total, 27 (12.1%) met both of these definitions, and we interviewed 14.

Our interviewees were diverse; and included 6 women and 8 men; 6 Democrats, 4 Independents, and 4 Republicans; 8 people who self-identified as White, and 6 people who self-identified as Non-White. The average age of interviewees was 41 years, with a standard deviation of 9.7 years.

We conducted semi-structured interviews with participants, which were geared to understanding their overall information ecosystems, how they stayed informed, their overall wellbeing, and their emotional responses to the news. For this analysis, we compiled all interview notes and memos through discussion into four key themes.

Results

We identify four recurring dimensions of high information wellbeing: socially oriented news engagement, personal coping strategies, active investigation of information, and stable personal dispositions.

Many participants engaged with news in socially oriented ways, treating informedness as a shared and relational practice rather than an individual responsibility. For example, P13 rarely sought out news directly, instead encountering it through conversations with her husband and son, while P3 described how news circulated through a close-knit family call/group-chat where articles were routinely shared and discussed. These practices helped participants regulate the emotional tone of news consumption, as they emphasized respectful conversation, knowing when to de-escalate, and selectively engaging with people who supported calm and meaningful discussion.

Some participants relied on personal coping strategies to limit the emotional impact of news, especially during periods of heightened stress. Rather than disengaging entirely, participants actively calibrated how, when, and how deeply they engaged with information. For example, P14 described how she would limit her engagement depth to issues that felt too negative, unable to act on, and personally less consequential. Similarly, P5 continued to follow major local and global events but managed emotional strain by skipping distressing content, reducing time spent on highly distracting platforms, and stepping away through offline routines such as religious community involvement.

Participants often engaged with news in active ways, which helped them feel a greater sense of agency over current events. For instance, P4 described keeping small note cards as a personal repository for processing news, using them to record names, claims, or events she encountered and later verify their sources, compare accounts, and occasionally watch press conferences firsthand. Similarly, P11 described intentionally researching issues that directly affected his life—such as immigration enforcement in his city—which helped reduce uncertainty. Across participants, this active engagement often involved seeking context, following stories over time, and attending to the human stakes behind abstract events.

Finally, participants' information wellbeing were also shaped by personal characteristics that interacted with their practices of news engagement. P9 described experiencing little emotional impact from news, attributing this to a stable personality trait rather than deliberate regulation. While he observed others becoming emotionally overwhelmed, he tended to remain detached, which allowed him to stay informed through curated headline feeds without distress. In contrast, P6 reported moments of emotional activation when engaging with political news, but described a strong capacity for psychological flexibility—seeking out counterarguments, and revising his views when presented with new evidence.

Future Work

Through qualitative interviews, we arrived at four potential identifiers of people with high information wellbeing. Moving forward, we see many opportunities to deepen and extend this work:

First, we will deepen our qualitative coding and analysis to move beyond preliminary themes to hypothesized strategies for fostering information wellbeing. The current analysis is based primarily on discussions within the research team and analysis of interview notes and memos. As an immediate next step, we plan to undertake a thorough interview coding process, enabling us to move beyond immediately apparent themes to more nuanced relationships. For example, a future area we plan to explore further is how people's perceptions of their own information wellbeing matches to the objective survey-based measures we have used to sample participants.

Second, we seek to validate our qualitative findings of the characteristics of people with high information wellbeing through a quantitative survey. This survey will be aimed to understand whether the qualities we have identified indeed correlate with high information wellbeing. Additionally, a quantitative survey will help us precisely understand the additional contribution of these qualities, above non-changeable characteristics like demographics. In a way, this will help us to understand whether people with high information wellbeing are born that way; or whether behavior change can point towards long term solutions. We also aim to deploy this survey longitudinally, such that we can identify whether information wellbeing is a temporary or a stable attribute.

Third, we wish to further theorize the idea of information wellbeing. While we may wish for news knowledge to be obtainable without sacrificing mental health, this may not always be possible or desirable. Indeed, being exposed to horrible things can push us to participate in civic causes or cultivate empathy for others. In these cases, it may be permissible and even societally necessary to sacrifice some individual mental health for the sake of staying collectively informed. Though we currently posit information wellbeing as maximised when wellbeing is harmonized with news knowledge, we want to further explore whether this outcome is always the desirable goal.

References

- Angelucci, C., & Prat, A. (2024). Is Journalistic Truth Dead? Measuring How Informed Voters Are about Political News. *American Economic Review*, 114(4), 887–925.
- Betakova, D., Boomgaarden, H., Lecheler, S., & Schäfer, S. (2025). I Do Not (Want To) Know! The Relationship Between Intentional News Avoidance and Low News Consumption. *Mass Communication and Society*, 28(3), 413–440.
- Boukes, M., & Vliegenthart, R. (2017). News Consumption and Its Unpleasant Side Effect. *Journal of Media Psychology*, 29(3), 137–147.

- Boulianne, S. (2016). Online news, civic awareness, and engagement in civic and political life. *New Media & Society*, 18(9), 1840-1856.
- Bunn, R. R., & Farmer, A. S. (2021). Is ignorance bliss? The relationship between news exposure and anxiety levels in adults. *The New School Psychology Bulletin*, 18(1), 38–54.
- Delli Carpini, M. X., Morales Suarez, M., & Herman, B. (2018). Being Informed: A Study of the Information Needs and Habits of Philadelphia residents. Lenfest Institute for Journalism. <https://www.lenfestinstitute.org/being-informed/>
- Fullana, M. A., Hidalgo-Mazzei, D., Vieta, E., & Radua, J. (2020). Coping behaviors associated with decreased anxiety and depressive symptoms during the COVID-19 pandemic and lockdown. *Journal of Affective Disorders*, 275, 80–81.
- Pal, O., Goyal, A., Chandrasekharan, E., & Saha, K. (2026). The Hidden Toll of Social Media News: Causal Effects on Psychosocial Wellbeing (arXiv:2601.13487). arXiv.<https://doi.org/10.48550/arXiv.2601.13487>
- Paluck, E. L., Shafir, E., & Wu, S. J. (2017). Ignoring alarming news brings indifference: Learning about the world and the self. *Cognition*, 167, 160–171
- Poindexter, P. M., & McCombs, M. E. (2001). Revisiting the Civic Duty to Keep Informed in the New Media Environment. *Journalism & Mass Communication Quarterly*, 78(1), 113–126.
- Schäfer, S., Aaldering, L., & Lecheler, S. (2023). “Give Me a Break!” Prevalence and Predictors of Intentional News Avoidance During the COVID-19 Pandemic. *Mass Communication and Society*, 26(4), 671–694.
- Silver, R. C., Holman, E. A., Andersen, J. P., Poulin, M., McIntosh, D. N., & Gil-Rivas, V. (2013). Mental- and Physical-Health Effects of Acute Exposure to Media Images of the September 11, 2001, Attacks and the Iraq War. *Psychological Science*, 24(9), 1623–1634.
- Stainback, K., Hearne, B. N., & Trieu, M. M. (2020). COVID-19 and the 24/7 News Cycle: Does COVID-19 News Exposure Affect Mental Health? - Kevin Stainback, Brittany N. Hearne, Monica M. Trieu, 2020. *Socius* 6.
- Strasser, M. A., Sumner, P. J., & Meyer, D. (2022). COVID-19 news consumption and distress in young people: A systematic review. *Journal of Affective Disorders*, 300, 481–491.
- Toff, B., & Nielsen, R. K. (2022). How News Feels: Anticipated Anxiety as a Factor in News Avoidance and a Barrier to Political Engagement. *Political Communication*, 39(6), 697–714.

- Wessel, T. J., Trattner, C., & Starke, A. (2025). Using Large Language Models to ‘Lighten the Mood’: Satirically Reframing News Recommendations to Reduce News Avoidance.
- Yoon, S., McClean, S. T., Chawla, N., Kim, J. K., Koopman, J., Rosen, C. C., Trougakos, J. P., & McCarthy, J. M. (2021). Working through an “infodemic”: The impact of COVID-19 news consumption on employee uncertainty and work behaviors. *Journal of Applied Psychology*, 106(4), 501–517.
- Zappia, C., & Makri, S. (2022). “I Try to Find a Balance”: Investigating Strategies for Self-Regulating Covid News Consumption. *Proceedings of the Association for Information Science and Technology*, 59(1), 368–380.
-

2B.3 Navigating Health Information on YouTube: Trust, Misinformation, and Health Decision Making During COVID-19 in India.

Nisha Rani (Manipal Institute of Communication), Usha Manchanda Rodrigues (School of Information and Communication Studies), Padma Rani (Manipal Institute of Communication)

Introduction

India witnessed an unprecedented digital expansion during the pandemic, adding over 78 million new internet users from 2020 to 2021 (IAMAI & Kantar, 2022). Indian users relied on online platforms for COVID-19 information to help understand and navigate health decision making. YouTube's user engagement in India increased sharply, with subscriptions rising by 20.5% and views exceeding 300 billion all within 45 days of lockdown (Mindshare & Vidooly, 2020). The platform enabled government agencies, health organisations and content creators to disseminate COVID-19 information to its 467 million users in India. Viewers were exposed to wide range of narratives making it difficult to trust and identify credible sources. Research shows that India along with USA, Brazil and Spain was among the countries most affected by misinformation during COVID-19 (Al-Zaman, 2021).

YouTube's ability to deliver audiovisual content at scale surpasses that of traditional media in both reach and immediacy, enabling real time video uploads from government agencies to individual creators. The surge in demand for COVID-19 content also led to an infodemic, with both accurate and inaccurate information rapidly spreading across the platform. YouTube's recommendation algorithm has been shown to prioritize content that elicits strong emotional reactions and high engagement, which can amplify provocative or emotionally charged messages regardless of their informational accuracy (Hussain et al., 2022). According to Li et al. (2020), around 27.5% of COVID-19-related information on YouTube was inaccurate in the early stages of the pandemic. This created challenges around the dissemination of misinformation and information overload for users. In a study focusing on Asian countries, Rodrigues et al. (2023) found that social media platforms increased vaccine hesitancy among viewers due to misinformation and anti-vaccination campaigns during the pandemic. Research has linked the spread of misinformation and false beliefs to increased morbidity and mortality among older adults. Mueller et al. (2020) found that such misinformation led to delayed or inappropriate medical interventions, resulting in higher death rates within this vulnerable population. Given these concerns, the study is particularly relevant as it addresses the spread of misinformation, variations in content reliability, audience

trust, emotional impact, and health behaviors, offering a comprehensive overview of how COVID-19 information on YouTube shaped public response during the pandemic in India.

Objectives/Research Questions

The study aims to assess the reliability of COVID-19 content on YouTube and understand how viewers navigated trust, misinformation and health decision making during the pandemic. We examine both the creation and impact of COVID-19-related content on YouTube.

The study addresses four key questions: 1) How do Hindi and English YouTube videos related to COVID-19 differ by creator type and thematic focus? 2) Is there a relation between content source and reliability of COVID-19 information on YouTube? 3) What impact does regular YouTube viewing have on emotional well-being and health decisions during the pandemic? 4) How did viewers assess video credibility and respond to misinformation while viewing COVID-19 content on YouTube? The research draws on George Gerbner's Cultivation Theory (1976) to examine the long-term impact of media exposure on viewer's perceptions and health behaviors during the COVID-19 pandemic. While initially applied to traditional broadcast media, Cultivation Theory has since been extended to digital and social media platforms (Hermann et al., 2020; Tang et al., 2021). Unlike short term media effects, Cultivation Theory allows for an understanding of how a media platform can shape audience perception over long period of time. This is particularly relevant in an algorithm driven platform like YouTube, where exposure is structured by recommendation systems rather than user choice alone.

Methods

According to YouTube's 2020 report, 54% of the videos watched by Indians were in Hindi language, while English ranked second with 16% of total videos watched. YouTube Data Tools (Rieder, 2015) was used to retrieve the top 500 videos in Hindi and English using the search term "Coronavirus" (January 2020–December 2021). After removing duplicates and non-relevant content, the final sample comprised 701 videos (292 English, 409 Hindi). Two independent coders categorized each video by source (government/public agencies; health organizations; independent creators; private media; public media) and by type of content (news/information, political, investigative, personal narrative). To evaluate the consistency of classification, intercoder reliability was calculated using Cohen's kappa. The coefficient was 0.921 ($p < 0.001$).

Each video's information quality was evaluated using an 8-item modified DISCERN scale (Charnock et al., 1999). The DISCERN scale has been employed in studies on YouTube and other digital health information sources (Szmuda et al., 2020). Additionally, video metadata, such as views and comments, were analysed to understand user engagement patterns across different sources and content types. A chi-square test of independence was conducted to examine the relationship between content type and source type. Based on the findings of the content analysis a survey questionnaire was prepared that focussed on viewers response to the COVID-19 content and how it impacted their decision making. We surveyed 1,067 participants aged 18 and above from five Indian states, West Bengal, Odisha, Madhya Pradesh, Bihar, and Rajasthan. The states were selected based on Google Trends data indicating the highest volume of YouTube searches for COVID-19-related content.

Results

Results show that Hindi-language videos demonstrate significantly higher audience engagement across all metrics, particularly in the number of comments, averaging 3,993 comments per video compared to 435 for English videos. Private media organizations dominated the content landscape, accounting for 58.4% of the total videos in Hindi and 47.3% in English. Health organizations achieved better reach among English-speaking audiences, representing 9.2% of the videos in English compared to only 3.2% in Hindi. The relationship between the content creator and the type of content they produced also reflects on their role in building the COVID-19 information network on YouTube. Government and public agencies, along with health organizations, primarily focused on news and informative content, which included regular COVID-19 updates, and avoided politically charged or investigative content, maintaining a neutral and authoritative tone.

Despite higher engagement in Hindi videos, reliability remained low. It is important to note that reliability does not align with popularity. The disconnect between engagement and reliability can be confusing for viewers who cannot judge content accuracy or verify the information from credible sources. The overall mean mDISCERN score was higher for English videos (27.35) compared to Hindi videos (26.30), suggesting better overall quality. Independent Content Creators show lower reliability in both languages raising concerns about user-generated content accuracy.

The findings also indicate that YouTube played a significant role in shaping health decision making during the pandemic. Over 50% of respondents reported making changes to their diet or lifestyle to enhance their physical and mental well-being. Furthermore, approximately 40% adopted preventive measures such as mask-wearing and avoiding crowded spaces. Notably, more than 25% of participants stated that

access to information on YouTube influenced their decision to receive the COVID-19 vaccine. Content credibility emerged as a significant challenge on the platform. Over 90% of participants encountered false or misleading information on YouTube and only 20% reported such content to the platform. Instead, the majority chose to verify the information through alternative sources, such as news outlets, official websites, or trusted individuals, to assess its accuracy. More than 50% of respondents experienced anxiety and stress while watching COVID-19-related content on YouTube. The research reinforces Cultivation Theory showing that sustained media exposure shapes perceptions, emotional responses, and health behaviors and acknowledging YouTube's role during a public health crisis and the urgent need for credible and accessible health communication. In March 2022, YouTube launched two new features for health-related videos in India to ensure more credible health information was available to users. A health source information panel was set up to assist users identify videos

Future Work

This study contributes to health communication and social media scholarship by offering a cross linguistic, mixed methods analysis of YouTube's role as a major health information platform during the COVID-19 pandemic in India. The findings draw attention to the structural limitations of platform logics that prioritize popularity and algorithmic amplification over credibility. The study advocates targeted policies that pushes reliable sources on social media platforms while expanding digital literacy initiatives to help users verify and report false information for more beneficial health communication during crises.

Future research could extend into examining YouTube's recommendation systems directly to understand how algorithmic curation influences repeated exposure to health content. Qualitative analysis of viewer comments would provide deeper insight into how audiences interpret and respond to health information online. Additionally, expanding cross linguistic comparisons to include regional Indian languages and other digital platforms would further enhance understanding of how health information circulates within multilingual digital ecosystems and inform more inclusive and effective public health interventions.

References

- Al-Zaman, M. S. (2021). Social media and COVID-19 misinformation: How ignorant Facebook users are? *Heliyon*, 7(5), e07144.
<https://doi.org/10.1016/j.heliyon.2021.e07144>

- Charnock, D., Shepperd, S., Needham, G., & Gann, R. (1999). DISCERN: An instrument for judging the quality of written consumer health information on treatment choices. *Journal of Epidemiology & Community Health*, 53(2), 105–111.
<https://doi.org/10.1136/jech.53.2.105>
- Graham, G. (2022, March 28). Introducing new ways to help you find answers to your health questions. YouTube Official Blog. <https://blog.youtube/news-and-events/introducing-new-ways-help-you-find-answers-your-health-questions/>
- Hermann, E., Eisend, M., & Bayón, T. (2020). Facebook and the cultivation of ethnic diversity perceptions and attitudes. *Internet Research*, 30(4), 1123–1141.
<https://doi.org/10.1108/INTR-10-2019-0423>
- Hussain, M. N., Tokdemir, S., Agarwal, N., & Al-Khateeb, S. (2018). disinformation and crowd manipulation tactics on YouTube. 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 1092–1095. <https://doi.org/10.1109/asonam.2018.8508766>
- Internet and Mobile Association of India (IAMAI), & Kantar. (2022). Internet in India 2022. Retrieved 24 January 2026, from https://www.iamai.in/sites/default/files/research/Internet%20in%20India%2022_Print%20version.pdf
- Li, H. O.-Y., Bailey, A., Huynh, D., & Chan, J. (2020). YouTube as a source of information on COVID-19: A pandemic of misinformation? *BMJ Global Health*, 5(5), e002604.
<https://doi.org/10.1136/bmjgh-2020-002604>
- Mindshare and Vidooly. (2021). COVID-19 impact on video consumption trends. Retrieved 16 January 2026, from <https://vidooly.lpages.co/impact-COVID-19-video-consumption/>
- Mueller, A. L., McNamara, M. S., & Sinclair, D. A. (2020). Why does COVID-19 disproportionately affect older people? *Aging*, 12(10), 9959–9981.
<https://doi.org/10.18632/aging.103344>
- Rieder, B. (2015). YouTube Data Tools (Version 1.23) [Software]. Digital Methods Initiative. <https://tools.digitalmethods.net/netvizz/youtube/>
- Rodrigues, F., Ziade, N., Jatuworapruk, K., Caballero-Urbe, C. V., Khursheed, T., & Gupta, L. (2023). The impact of social media on vaccination: A narrative review. *Journal of Korean Medical Science*, 38(40), e326.
<https://doi.org/10.3346/jkms.2023.38.e326>

Szmuda, T., et al. (2020). YouTube as a source of patient information for COVID-19: A content-quality and audience engagement analysis. *Reviews in Medical Virology*, 30(5), e2132. <https://doi.org/10.1002/rmv.2132>

2B.4 Beyond screen time: Differentiating specific digital activities as risks for adolescent mental health problems.

Peiyao Tang, Katarzyna Kostyrka-Allchorne (Queen Mary University of London), Aja Murray (University of Edinburgh), Mariya Stoilova (LSE), Marianne Etherson (University of Glasgow), Eliz Azeri (King's College London), Ellen Townsend, Chris Hollis (University of Nottingham), Sonia Livingstone (LSE), Edmund Sonuga-Barke (King's College London)

Introduction

Concerns that excessive 'screen time' harms adolescent mental health increasingly shape research, policy, and public debate. Although some studies link greater digital technology use with depression, anxiety, and lower wellbeing (Baiden et al., 2019), the overall effects of screen time are typically small, inconsistent, and highly dependent on context (Kim et al., 2020; Orben, 2020).

Existing studies are limited in three ways. First, most studies have used undifferentiated measures of screen time, which obscure the complexity of adolescents' digital experiences and their varying mental health impacts and limit our ability to identify specific pathways of harm or meaningful intervention targets (Barr et al., 2020).

Emerging evidence indicates that particular digital activities, such as compulsive social media use, social comparison, and exposure to risky content or interactions, show stronger associations with psychopathology than screen time alone (Brailovskaia & Margraf, 2023; Kostyrka Allchorne et al., 2025).

Second, there may be specific pathways between the types of digital activities and distinct mental health outcomes, whereas few studies have set out to differentiate between clinically distinct components of mental health, such as depression and anxiety. Researching internalising problems in general or combining multiple mental health conditions could impede the characterisation of specific digital-mental health condition pathways (Kostyrka-Allchorne et al., 2023).

Third, most studies are cross-sectional or do not clearly distinguish stable differences between individuals from changes that occur within individuals over time.

Consequently, although studies often identify a correlation between digital activities and mental health, we are unable to draw inferences about any directional effects over time. Moreover, recent theory highlights the plausibility of reciprocal pathways, where

low mood could increase engagement in risky digital activities, which, in turn, may contribute to subsequent emotional and mental health problems (Coyne et al., 2024; Sonuga-Barke et al., 2024).

Objectives/Research Questions

The present longitudinal study addresses the aforementioned gaps by examining the longitudinal and reciprocal pathways between specific digital activities and adolescent mental health, adopting five empirically derived categories of digital activity: 'risky content', 'social comparison', 'risky interactions', 'leisure activity', and 'social engagement'. We report data from a 16-month longitudinal study in adolescents aged 13-14 years old, with digital activity and mental health reported at three time points. Using random-intercept cross-lagged panel models, we test bidirectional longitudinal associations between specific digital activities and depressive and anxiety symptoms, while distinguishing within-person effects from stable trait-like between-person differences.

We examined four research questions:

1. Which specific digital activities predict an increase in depressive symptoms over 16 months in middle-adolescence?
2. To differentiate between depression and anxiety, we asked: Are the observed effects of digital activity similar or different for these conditions?
3. To address the issue of reciprocity, we asked: Do depressive or anxiety symptoms alter digital activity?
4. Finally, to compare the directional and longitudinal effect of screen time versus specific digital activities, we asked: Does screen time predict changes in depressive and anxiety symptoms?

Methods

Study design and sample

This longitudinal observational study surveyed adolescents at three time points, each approximately eight months apart. Participants were recruited from UK secondary schools via the research team's networks to cover a wide socioeconomic range. Adolescents provided informed consent via Qualtrics. The baseline survey (T1) was conducted between January and April 2024, the first follow-up (T2) between September and November 2024, and the final survey (T3) between May and June 2025. The average follow-up duration was 13.8 months (range 11.9-16.1). Ethical approval was granted by the [removed for peer review] Research Ethics Committee (reference 249287).

At baseline, 552 adolescents attending Year 9 participated (mean age 13.7, SD = 0.5), with 263 (47.6%) completing T2 and 358 (64.9%) completing T3, respectively. The sample had an even distribution of sex (48% boys, 48% girls, 4% not disclosed) and a diverse representation of ethnicity (55% White, 15% South/East Asian, 8% Black, 18%

mixed or other ethnicities). Overall, 23% of the sample had disadvantaged socioeconomic background.

Measures

Digital activities were assessed using the 20-item My Online Activity section of the Digital Activity and Feelings Inventory (DAFI),¹² capturing five types: risky content, social comparison, risky interactions, leisure activity, and social engagement (Kostyrka Allchorne et al., 2025). Screen time was measured with a single self-reported item on daily phone/internet use. Depression and anxiety were assessed using the Revised Child Anxiety and Depression Scale 25 – Youth Version (RCADS-25), with separate 10-item depression and 15-item anxiety subscales (Chorpita et al., 2000).

Analysis strategy

Analyses were performed using R (version 2025.05.1; lavaan package; Rosseel et al., 2025). Random-intercept cross-lagged panel models (RI-CLPMs) examined longitudinal and reciprocal associations, separating within-person changes from stable between-person differences. Equality constraints were imposed sequentially on autoregressive paths and cross-lagged paths (Ciarrochi et al., 2016).

We conducted two sets of models for depression and anxiety symptoms, respectively, for each of the five types of digital activity (10 models overall). Two separate models additionally examined the association between screen time and depression/anxiety. We used full information maximum likelihood (FIML) to address missing data. The Comparative Fit Index (CFI; $>.90$), the Tucker-Lewis Index (TLI; $>.90$), root mean square error of approximation (RMSEA; $<.05$), and standardised root mean square residual (SRMR; $<.08$) were adopted to evaluate model fit (Hu & Bentler, 1998). Standardised effect sizes are reported.

Results

For depression, at the within-person level, higher social comparison at baseline (T1) and 8-month follow-up (T2) predicted subsequently increased symptoms at 8-month (T2) and 12-month follow-up (T3), respectively ($T1 \rightarrow T2: \beta = .212, p = .037$; $T2 \rightarrow T3: \beta = .198, p = .037$). No within-person effects were observed for other types of activities.

Between-person associations of large effect were observed linking social comparison ($\beta = .723, p < .001$), risky content ($\beta = .686, p < .001$), risky interactions ($\beta = .578, p < .001$), and social engagement ($\beta = .237, p = .048$) to depression, but not leisure activity. Screen time showed no within-person effects but a significant between-person association with depression ($\beta = .542, p < .001$).

For anxiety, within-person increases in social comparison ($\beta = .271, p = .003$) and risky interactions ($\beta = .226, p = .010$) at T1 predicted higher subsequent symptoms at T2, with consistent effects for both activities from $T2 \rightarrow T3$ ($\beta = .249, p = .003$; $\beta = .262, p = .010$). No within-person effects were observed for other types of activities.

Significant between-person associations were found for social comparison ($\beta = .728$, $p < .001$), risky content ($\beta = .546$, $p < .001$), risky interactions ($\beta = .494$, $p < .001$), and social engagement ($\beta = .223$, $p = .028$), but not for leisure activity. Screen time was associated with higher anxiety at between-person level ($\beta = .427$, $p < .001$), but not at within-person level.

These findings indicate consistent within-person longitudinal associations linking social comparison to subsequent depression and anxiety, and risky interactions to anxiety. There are also strong between-person covariances between each putative risky digital activity type and both depression and anxiety. Screen time alone explains little variance in the changes of symptoms within the same person across time.

Future Work

This study provides robust evidence that specific digital activities, rather than overall screen time, are most relevant to adolescent mental health. Social comparison consistently predicted within-person increases in both depression and anxiety, suggesting it as a common underlying risk factor. Together, the evidence points to social comparison as a key target for interventions, whether including digital literacy that helps adolescents critically evaluate idealised online portrayals and/or platform regulation that reduces comparison-inducing design features.

Risky interactions, such as peer exclusion or hurtful online exchanges, were uniquely linked to anxiety. Strengthening adolescents' resilience to cope with such interactions through raising awareness, building digital skills, and providing access to supportive on- and offline networks may help reduce this risk.

Although screen time showed robust between-person associations with both depressive and anxiety symptoms, it did not show any within-person effects. Our study provided robust evidence that screen time encompasses diverse activities and is not sufficiently reliable as a standalone measure of digital activity-mental health pathways.

However, the findings should be weighed in the context of some limitations, such as a relatively short overall follow-up duration of 16 months. Furthermore, although data were collected at shorter intervals than other longitudinal studies, the 8-month intervals may have been too long to fully capture the dynamic and ever-changing nature of youth digital activity and its impacts on mental health.

Future research should examine the mechanisms linking social comparison and risky interactions to mental health and explore potential mediators and moderators, such as cognitive and affective reactions. Also, studies with shorter follow-up intervals and more frequent assessments could better capture dynamic fluctuations in digital activity and mental health. Moreover, experimental and intervention studies could test strategies to reduce social comparisons, mitigate risky interactions, and promote positive online engagement.

References

- Baiden, P., Tadeo, S. K., & Peters, K. E. (2019). The association between excessive screen-time behaviors and insufficient sleep among adolescents: Findings from the 2017 youth risk behavior surveillance system. *Psychiatry Research*, 281, 112586. <https://doi.org/10.1016/J.PSYCHRES.2019.112586>
- Barr, R., Kirkorian, H., Radesky, J., Coyne, S., Nichols, D., Blanchfield, O., Rusnak, S., Stockdale, L., Ribner, A., Durnez, J., Epstein, M., Heimann, M., Koch, F. S., Sundqvist, A., Birberg-Thornberg, U., Konrad, C., Slussareff, M., Bus, A., Bellagamba, F., & Fitzpatrick, caroline. (2020). Beyond Screen Time: A Synergistic Approach to a More Comprehensive Assessment of Family Media Exposure During Early Childhood. *Frontiers in Psychology*, 11, 520774. <https://doi.org/10.3389/FPSYG.2020.01283/BIBTEX>
- Brailovskaia, J., & Margraf, J. (2023). Less sense of control, more anxiety, and addictive social media use: Cohort trends in German university freshmen between 2019 and 2021. *Current Research in Behavioral Sciences*, 4, 100088. <https://doi.org/10.1016/J.CRBEHA.2022.100088>
- Chorpita, B. F., Yim, L., Moffitt, C., Umemoto, L. A., & Francis, S. E. (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: a revised child anxiety and depression scale. *Behaviour Research and Therapy*, 38(8), 835–855. [https://doi.org/10.1016/S0005-7967\(99\)00130-8](https://doi.org/10.1016/S0005-7967(99)00130-8)
- Ciarrochi, J., Parker, P., Sahdra, B., Marshall, S., Jackson, C., Gloster, A. T., & Heaven, P. (2016). The development of compulsive internet use and mental health: A four-year study of adolescence. *Developmental Psychology*, 52(2), 272–283. <https://doi.org/10.1037/DEV0000070>
- Coyne, S. M., Shawcroft, J., Holmgren, H., Christensen-Duerden, C., Ashby, S., Rogers, A., Reschke, P. J., Barr, R., Domoff, S., Van Alfen, M., Meldrum, M., & Porter, C. L. (2024). The growth of problematic media use over early childhood: Associations with long-term social and emotional outcomes. *Computers in Human Behavior*, 159, 108350. <https://doi.org/10.1016/J.CHB.2024.108350>
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424–453. <https://doi.org/10.1037/1082-989X.3.4.424>
- Kim, S., Favotto, L., Halladay, J., Wang, L., Boyle, M. H., & Georgiades, K. (2020). Differential associations between passive and active forms of screen time and adolescent mood and anxiety disorders. *Social Psychiatry and Psychiatric Epidemiology*, 55(11), 1469–1478. <https://doi.org/10.1007/S00127-020-01833-9/TABLES/3>

Kostyrka Allchorne, K., Bourgaize, J., Murray, A., Stoilova, M., Abbas, I., Bridgwood, A., Azeri, E., Hollis, C., Townsend, E., Livingstone, S., & Sonuga-Barke, E. J. S. (2025). Understanding Youth Online Experiences and Mental Health: Development and Validation of the Digital Activity and Feelings Inventory (DAFI). *International Journal of Methods in Psychiatric Research*, 34(2), e70028.

<https://doi.org/10.1002/MPR.70028>

Kostyrka-Allchorne, K., Stoilova, M., Bourgaize, J., Murray, A., Azeri, E., Hollis, C., Townsend, E., Livingstone, S., & Sonuga-Barke, E. (2024). Dynamic Interplay of Online Risk and Resilience in Adolescence (DIORA): a protocol for a 12-month prospective observational study testing the associations among digital activity, affective and cognitive reactions and depression symptoms in a community sample of UK adolescents. *BMJ Open*, 14(9), e085061.

<https://doi.org/10.1136/BMJOPEN-2024-085061>

Kostyrka-Allchorne, K., Stoilova, M., Bourgaize, J., Rahali, M., Livingstone, S., & Sonuga-Barke, E. (2023). Review: Digital experiences and their impact on the lives of adolescents with pre-existing anxiety, depression, eating and nonsuicidal self-injury conditions – a systematic review. *Child and Adolescent Mental Health*, 28(1), 22–32. <https://doi.org/10.1111/CAMH.12619>

Orben, A. (2020). Teenagers, screens and social media: a narrative review of reviews and key studies. *Social Psychiatry and Psychiatric Epidemiology*, 55(4), 407–414.

<https://doi.org/10.1007/S00127-019-01825-4/METRICS>

Rosseel, Y., Jorgensen, T. D., & De Wilde, L. (2025). Latent Variable Analysis [R package lavaan version 0.6-20]. CRAN: Contributed Packages.

<https://doi.org/10.32614/CRAN.PACKAGE.LAVAN>

Sonuga-Barke, E. J. S., Stoilova, M., Kostyrka-Allchorne, K., Bourgaize, J., Murray, A., Tan, M. P. J., Hollis, C., Townsend, E., & Livingstone, S. (2024). Pathways between digital activity and depressed mood in adolescence: outlining a developmental model integrating risk, reactivity, resilience and reciprocity. *Current Opinion in Behavioral Sciences*, 58, 101411.

<https://doi.org/10.1016/J.COBEHA.2024.101411>

Session 2C - Social Media Cultures and Everyday Life (2)

Day 1 · Mon 13 July · 11:30–13:00 · Seminar Suite C, ARC

2C.1 Fentanyl Tourism on YouTube: Serial Commodification of Urban Marginality.

Elsa Soro (European Business School), Lluís Garay Tamajón (Universitat Oberta de Catalunya [UOC])

Introduction

Urban marginality, poverty, and drug use have long been present within Western media representations, particularly in documentary journalism and investigative reporting, where media coverage of urban deprivation (Dreier, 2005) has historically shaped public understandings of drug use and addiction through narratives of social disorder, crisis, and vulnerability (Orsini, 2017).

At the same time, a substantial body of scholarship shows that media representations of drug use and urban marginality frequently rely on dehumanizing frames. Labels such as “zombies” or “crackheads” recur in coverage of drug-affected urban spaces, reinforcing stigma and stripping individuals of social and political agency (Alexandrescu, 2020; Habib et al., 2023). Research further suggests that such representations contribute to punitive drug and welfare policies while deflecting attention from structural determinants of drug use, including housing precarity and inequality (Bosworth et al., 2024).

Within this context, the opioid crisis in the United States and Canada has occupied a central position in public and media debates and has been conceptualized as a socially constructed phenomenon produced through media, policy, and public discourse (Webster et al., 2019). Beyond news media, the opioid crisis has generated a substantial body of documentary content distributed by mainstream streaming platforms, such as Heroin(e), The Crime of the Century, and Dopesick, which have positioned the crisis as a national public health issue rather than a localized or episodic spectacle.

Alongside these institutional media representations, the opioid crisis has increasingly become a highly visible and recurring subject of long-form YouTube videos. This content is produced primarily by independent creators and is commonly framed

through geographically explicit titles that pair specific urban locations with opioid-related substances or dehumanizing descriptors or metaphors such as “zombies.” Titles linking places such as Kensington Avenue, Skid Row, or Vancouver’s Downtown Eastside with references to fentanyl, tranq, or drug-induced states position urban neighborhoods as destinations to be visually explored, reinforcing a spatialized and substance-centered framing of crisis. Many of these videos accumulate hundreds of thousands or millions of views, marking a significant transformation in how urban drug-related vulnerability is mediated through platform-native formats. This dynamic has also been noted in journalistic commentary, which has criticized the emergence of “tranq tourism” as a platform-driven phenomenon that turns extreme vulnerability into spectacle (The Guardian, 2023)

Objectives/Research Questions

The central hypothesis of this research is that a defining feature of these videos lies in the explicit appropriation of tourism-related language and imaginaries.

Neighborhoods affected by drug use are framed as destinations, movement through space is structured as a “tour,” and viewers are invited to digitally traverse zones of marginality. These framings echo existing discussions on slum tourism, dark tourism, and narco-tourism, while simultaneously departing from them by relocating tourism logics into a fully mediated, platform-based environment shaped by algorithmic visibility rather than physical mobility.

This pilot study analyzes a curated sample of long-form YouTube videos depicting drug-affected urban areas in the United States and Canada and examines recurring patterns across four analytical dimensions: the spectacularization of urban suffering, the circulation of tourism-like imaginaries, platform-specific aesthetic and narrative conventions, and creator practices related to engagement and monetization

RQ1

How are urban drug-related crises spectacularized through visual and narrative strategies in long-form YouTube videos commonly framed as “fentanyl tours”?

RQ2

What recurring platform-oriented formats, aesthetic conventions, and tourism-like imaginaries emerge in the algorithmically mediated representation of urban suffering on YouTube?

RQ3

How do creator practices and algorithmic affordances shape the visibility, repetition, and monetization of crisis-related content within YouTube’s long-form video ecology?

Methods

This study adopts an exploratory qualitative research design and is structured as a pilot case study aimed at examining how urban drug-related crises are represented within YouTube's long-form video ecology. Given the study's focus on formats, discursive strategies, and platform logics rather than representativeness or causal inference, a limited and purposive sample was selected to enable in-depth qualitative analysis.

Data were collected using an algorithm-guided snowball sampling approach, designed to mirror platform-native pathways of content discovery. Seed queries were constructed by combining two dimensions: geographic identifiers (cities and urban neighborhoods widely recognized as sites of concentrated opioid or street drug use) and substance-related terminology (e.g., fentanyl, tranq, opioid). Initial videos were identified through YouTube's search and recommendation features, after which the sample was iteratively expanded using the platform's "related videos" function until thematic saturation was reached and no new narrative formats or creator types emerged.

The focus on the United States and Canada reflects the centrality of these contexts within the contemporary opioid crisis, as well as the high degree of mediatization of specific urban neighborhoods. Inclusion criteria applied post hoc included temporal relevance (videos published within the last 24 months), long-form format, and the presence of explicit walkthrough or touristic framing in titles or thumbnails. The final dataset consists of a curated sample of videos produced by independent creators.

The analysis combines qualitative content analysis with critical discourse analysis (van Dijk, 2015). Qualitative content analysis (Bengtsson, 2016; Long & Johnson, 2000) was employed to systematically identify recurring formal, visual, and narrative features across videos, while CDA informed the interpretation of how meanings, power relations, and moral positioning are constructed through language, imagery, and platform-specific conventions. This combined approach allows for both pattern identification and critical interpretation without presupposing audience effects or algorithmic causality.

Videos were treated as the primary unit of analysis and examined using a theory-informed coding grid, developed iteratively and applied consistently across the dataset. The grid structures analysis around four analytical dimensions directly derived from the dataset and relevant theoretical frameworks:

- (1) Spectacularization of urban suffering (narrative framing, shock, repetition);
- (2) Touristic language and spatial framing (walkthrough practices, construction of place);
- (3) Creator positioning and commercialization of crisis (performative roles, engagement cues, monetization signals).

Algorithmic indicators such as view counts, recurrence of formats, and serial production were recorded descriptively as contextual markers of platform visibility.

Results

Preliminary analysis of the coding grid reveals recurring patterns across the four analytical dimensions, indicating a high degree of formal and discursive convergence among long-form YouTube videos depicting drug-affected urban areas.

Spectacularization of urban suffering

Across the dataset, urban suffering is frequently positioned as a central visual and narrative focal point. Videos often privilege prolonged exposure to visible vulnerability through close-ups of bodies in distress and decontextualized depictions of drug use. The grid shows that spectacularization operates not as an exception but as a recurrent structuring device, where shock and visual intensity coexist with processes of banalization (Markham, 2017; Natale, 2024) produced through repeated exposure, shaping how crisis is made visible and consumable. This process resonates with broader discussions on mediated spectatorship and moral distance (Chouliaraki, 2006).

Touristic language and spatial framing

A consistent pattern emerges in the spatial framing of neighborhoods as spaces to be traversed and observed. Even when explicit tourism vocabulary is absent, walkthrough practices, first-person movement, and exploratory narration implicitly construct urban areas as destinations to be experienced (Duarte, 2012). Titles and thumbnails frequently combine geographic identifiers with substance-related terminology. This configuration aligns with a touristic gaze (Larsen, 2014) that invites viewers to digitally “visit” zones of marginality, where everyday suffering is transformed into an experiential commodity for mediated consumption.

Creator positioning and commercialization of crisis

The grid indicates that many creators adopt hybrid positionalities, oscillating between reporter, concerned observer, and explorer. Monetization cues (Kim, 2024; Ørmen & Gregersen, 2022) such as YouTube advertising revenue, channel memberships, sponsored content, and external revenue streams such as affiliate links or creator platforms are frequently present, even in videos framed as awareness-raising or documentary. Rather than being peripheral, commercial elements are often structurally integrated into the content, blurring boundaries between witnessing, advocacy, and value extraction.

Taken together, these preliminary findings suggest that YouTube’s long-form video ecology facilitates the emergence of a standardized representational logic in which urban drug-related vulnerability is spectacularized, spatially touristicized, and rendered

a source of symbolic and economic value within creator- and platform-driven economies of visibility (Nieborg & Poell, 2018)

Within this logic, urban marginality is made persistently visible (Shabbir et al., 2024) while remaining representationally dispossessed: bodies are present, but voices and forms of agency are largely absent, reinforcing a soft yet systematic extraction of value from urban suffering.

Future Work

Future research will systematize and expand this pilot study by extending both its geographical scope and platform scope. While the present analysis focuses on long-form YouTube videos depicting drug-affected urban areas in the United States and Canada, subsequent phases will incorporate additional urban contexts and media environments in order to examine how similar imaginaries of vulnerability, crisis, and marginality circulate across different settings.

A first line of expansion concerns geographical comparison. Building on the analytical framework and coding grid developed in this pilot, future work will include European cities and peripheral urban neighborhoods that are digitally mediatized as spaces of vulnerability, exclusion, or danger. This extension will allow for comparative analysis between North American and European contexts, as well as between globally recognized “crisis hotspots” and urban areas traditionally positioned outside mainstream tourism imaginaries.

A second line of development involves platform and format diversification. While this study focuses on long-form YouTube videos, future research will incorporate additional social media platforms and content formats, including short-form video, livestreams, and cross-platform circulation. This expansion will make it possible to examine how repeated exposure and platform-native consumption contribute to the banalization of what is seen, reinforcing soft forms of representational dispossession through routine viewing.

Methodologically, future work will be responsive to dataset scale and analytical complexity. As the corpus expands, the research may integrate complementary approaches suitable for larger datasets, including computational content analysis, descriptive engagement metrics, and pattern detection across time and platforms, while maintaining a theory-driven and critical interpretative orientation. This will allow systematic comparison of formats, narratives, and visibility dynamics without reducing the analysis to purely quantitative measures.

Finally, future research will deepen the levels of analysis, including discursive and narrative analysis, visual and aesthetic analysis, and the study of touristic performativity. Where data availability allows, attention will also be given to affective

framing, ethical tensions, and the emergence of alternative or counter-narratives that attempt to engage with marginalized urban spaces in more reflexive or critical ways.

References

- Alexandrescu, L. (2020). Streets of the 'spice zombies': Dependence and poverty stigma in times of austerity. *Crime, Media, Culture: An International Journal*, 16, 113–97. <https://doi.org/10.1177/1741659019835274>
- Bengtsson, M. (2016). How to plan and perform a qualitative study using content analysis. *NursingPlus Open*, 2, 8–14. <https://doi.org/10.1016/j.npls.2016.01.001>
- Bosworth, K., Massey, Z., Boyle, M., Henry, N., McGough, K., Ashford, A., Rains, E., Battle, J., Kelly, P., Malaker, P., & Tilhou, A. (2024). Analysing media portrayals of people with substance use disorder and addiction: A scoping review. *Cultures of Science*, 7, 126–141. <https://doi.org/10.1177/20966083241313264>
- Chouliaraki, L. (2006). *The spectatorship of suffering*. Thousand Oaks, CA: SAGE.
- Dreier, P. (2005). How the media compound urban problems. *Journal of Urban Affairs*, 27, 193–201. <https://doi.org/10.1111/j.0735-2166.2005.00232.x>
- Duarte, P. (2012). Tourist experience: contemporary perspectives. *Journal of Heritage Tourism*, 7(2), 187–188. <https://doi.org/10.1080/1743873X.2011.638459>
- Habib, D., Giorgi, S., & Curtis, B. (2023). Role of the media in promoting the dehumanization of people who use drugs. *The American Journal of Drug and Alcohol Abuse*, 49, 371–380. <https://doi.org/10.1080/00952990.2023.2180383>
- Kim, D. (2024). Selling Otherness on YouTube: Digital inter-Asian Orientalism and YouTube monetization system. *International Journal of Cultural Studies*, 28(1), 39–55. <https://doi.org/10.1177/13678779241240780>
- Larsen, J. (2014). The Tourist Gaze 1.0, 2.0 and 3.0. In A. Lew, M. Hall, & A. Williams (Eds.), *The Wiley Blackwell Companion to Tourism* (pp. 304–313). Wiley-Blackwell
- Markham, T. (2017). Banal phenomenologies of violence: Media work cultures and audience engagement with distant suffering. In G. Hodgson (Ed.), *Conflict, trauma and the media: A collection of essays* (pp. xx–xx). Cambridge Scholars Publishing.
- Moeller, S. D. (1999). *Compassion fatigue: How the media sell disease, famine, war and death*. Routledge.
- Natale, S. (2024). Digital media and the banalization of deception. *Convergence: The International Journal of Research into New Media Technologies*, 31(1), 402–419. <https://doi.org/10.1177/13548565241311780>
- Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society*, 20(11), 4275–4292. <https://doi.org/10.1177/1461444818769694>

- Ørmen, J., & Gregersen, A. (2022). Institutional polymorphism: Diversification of content and monetization strategies on YouTube. *Television & New Media*, 24(4), 432–451. <https://doi.org/10.1177/15274764221110198>
- Orsini, M. (2017). Frame analysis of drug narratives in network news coverage. *Contemporary Drug Problems*, 44, 189–211. <https://doi.org/10.1177/0091450917722817>
- Quan, A., Wilson, L., Mithani, S., Zhu, D., Bota, A., Wilson, K., & Wilson, K. (2020). Reporting on the opioid crisis (2000–2018): Role of The Globe and Mail, a Canadian English-language newspaper in influencing public opinion. *Harm Reduction Journal*, 17. <https://doi.org/10.1186/s12954-020-00443-7>
- Randolph-Krisova, A. (2018). Descriptive analysis of the most viewed YouTube videos related to the opioid epidemic. <https://doi.org/10.7916/d8x36dxx>
- Rieder, B., Borra, E., Coromina, Ò., & Matamoros-Fernández, A. (2023). Making a living in the creator economy: A large-scale study of linking on YouTube. *Social Media + Society*, 9. <https://doi.org/10.1177/20563051231180628>
- Shabbir, H., Bennett, R., Kottasz, R., Vijaygopal, R., Gardasz, B., Adams, J., & Kendall, P. (2024). Poverty porn as humanitarian business: The effects of framing, affect intensity, and spokesperson characteristics. *Business & Society*, 64(7), 1354–1391. <https://doi.org/10.1177/00076503241275762>
- Taylor, S. (2008). Outside the outsiders: Media representations of drug use. *Probation Journal*, 55, 369–387. <https://doi.org/10.1177/0264550508096493>
- van Dijk, T. A. (2015). Critical discourse analysis. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds.), *The handbook of discourse analysis*. <https://doi.org/10.1002/9781118584194.ch22>
- Webster, F., Rice, K., & Sud, A. (2019). A critical content analysis of media reporting on opioids: The social construction of an epidemic. *Social Science & Medicine*, 244, 112642. <https://doi.org/10.1016/j.socscimed.2019.112642>

2C.2 Constituting Affective Reality: Platformed Nostalgia and Digital Materiality in the Chinese Hanfu Revival Movement

Erika Ningxin Wang (The Chinese University of Hong Kong, Shenzhen)

Introduction

In February 2025, a small group of Hanfu (traditional Chinese clothing) enthusiasts gathered in Nanjing, meticulously adjusting hairstyles and accessories not merely for historical accuracy, but for the camera lens. “If I tilt it forward, it looks better... not too influencer-like, but it depends on how it looks on camera,” remarked Lian, a community member, as she browsed Douyin (Chinese TikTok) for reference videos. This vignette illustrates a crucial shift in the Hanfu revival movement: it is no longer just a nationalist project of recovering a lost past, but a digital-material production where the “vibe” of history is curated for algorithmic circulation.

While existing scholarship often frames Hanfu through the lens of digital nationalism (Liu, 2012; Carrico, 2017) or youth subculture (De Kloet & Fung, 2017), this paper argues that such perspectives overlook the precise mechanisms of emotional mediation. We propose that the Hanfu revival constitutes an “affective reality”: a shared experiential space where emotional resonance and aesthetic coherence supplant historical fact as the arbiter of authenticity.

To theorize this, we introduce the concept of platformed nostalgia. Drawing on theories of digital materiality (Pink et al., 2016) and performativity (Gillespie, 2014; Butler, 1990), we define platformed nostalgia as a technologically mediated process wherein an affective sense of “pastness” is manufactured, circulated, and validated by platform aesthetics and algorithmic logic. In this context, Hanfu functions as a digital-material hybrid—simultaneously a wearable garment and a watchable content piece—where tradition is not passively inherited but actively reimagined through the interface of the body, the camera, and the platform.

Objectives/Research Questions

The primary objective of this study is to shift the scholarly focus from static symbolic analyses of Hanfu (i.e., debates on nationalism) to the dynamic, processual construction of tradition in a digital age. We aim to understand how digital platforms calibrate and circulate national sentiment, treating Hanfu not just as heritage, but as a case study in the algorithmic politics of emotion.

Specifically, this research addresses three central questions:

1. Negotiation of Authenticity: How is cultural authenticity negotiated within the

Hanfu community as meaning shifts between online discourses of historical fidelity and offline aesthetic judgments?

2. Algorithmic Governance: What are the digital-material mechanisms through which platforms like Taobao and Douyin curate Hanfu's visibility, shaping it into a consumable aesthetic?
3. Embodied Performance: How is "platformed nostalgia" enacted through embodied practices, and how do these practices contribute to the collective construction of an "affective reality"?

Methods

To address the research questions, this study employs a multi-sited ethnographic approach, combining offline participant observation with digital ethnography. Fieldwork was conducted between October 2024 and April 2025, primarily focused on "Community H," a Hanfu fandom network based in Nanjing, China.

Data Collection:

- Participant Observation: The researchers adopted the dual roles of community member and amateur photographer, participating in community walks, styling workshops, and photo shoots at sites like the Confucius Temple and the Nanjing Museum. This allowed for the observation of "objects in action"—how garments are worn, posed in, and captured for social media.
- Interviews: Formal and informal interviews were conducted with 29 participants, ranging from university students and community organizers to store owners and photographers.
- Digital Ethnography: We traced the circulation of Hanfu content across three major platforms: Taobao (e-commerce), Douyin (short video), and Xiaohongshu (lifestyle sharing). We analyzed algorithmic trends, hashtags (e.g., #HanfuOuting, #NationalStyle), and comment sections to understand how platform logics structure the affective economy of the movement.

Analysis: Data was analyzed using a digital-material lens (Pink et al., 2016), treating online content and offline material practices as deeply entangled. We examined how virtual content (videos, tutorials) "lands" as material formations (clothing choices, poses) and how these materialities are re-uploaded to reshape the digital ecosystem.

Results

1. The Politics of Authenticity: From Restoration to Reformation Contrary to the assumption of rigid historical fundamentalism, we found that "authenticity" is a fluid, performative construct. While some purists, like the participant "Bai," insist on Xingzhi (structural rules) and reject "New Chinese Style" hybrids, the broader community increasingly favors an affective consensus driven by aesthetics. Authenticity is often determined by what feels genuinely "Han" or looks beautiful on camera, rather than

historical accuracy. We observed a shift from "restorative nostalgia" (rebuilding the past) to "reflective nostalgia" (aestheticizing the past). For instance, community members like Shen and Hou argue for the democratization of elite symbols (e.g., imperial robes) based on economic accessibility rather than social rank. This negotiation redefines tradition not as a fixed history, but as a "wearable tradition" validated by contemporary consumerism and visual appeal.

2. Curating Tradition: Algorithmic Governance We found that platforms actively intervene in cultural debates by curating which versions of tradition become visible. On Taobao, keywords like "Wei-Jin elegance" or "museum restoration" serve as gateways that fuse historical imagery with modern desire, creating "affective zones of perception" (Bucher, 2018). On Douyin and Xiaohongshu, algorithms prioritize "vibes" over context. Participants Sun and Ge, for example, used saved Douyin clips to direct their offline photo shoots, explicitly seeking to replicate a "sense of immersion" found in short videos. This reliance on algorithmic recommendations creates a feedback loop where fragmented, aestheticized history is consumed as truth. The platforms effectively govern tradition by promoting content that is visually arresting and emotionally resonant, often stripping away complex historical narratives in favor of a "consumable past."

3. Crafting Affective Reality: The Practice of "Proper Dressing" The final stage is the embodied enactment of this curated past, which we term "proper dressing." This is best exemplified by the parodic performance of hierarchy. We observed members like "Zhu" wearing a Longpao (Dragon Robe)—historically reserved for the Emperor—while others playfully acted as subjects. This is not a delusion of grandeur but a form of "platformed play" or parody (Hutcheon, 1985). Through these performances, symbols of authoritarian power are decoupled from actual governance and repurposed as aesthetic resources for self-expression. When these moments are captured and uploaded, they complete the digital-material circuit. The result is an "affective reality": a tangible, felt experience of history that is socially binding for the community, even if historically inaccurate. As participants noted, the goal is not a history lesson, but an "optimistic attachment" (Berlant, 2011) to a fantasized past that offers an escape from the pressures of modern life.

Future Work

This study establishes "platformed nostalgia" as a key mechanism in the construction of contemporary cultural identity. Future work should expand this framework beyond the Chinese context to test its global applicability. Comparative studies could investigate:

1. **European Historical Reenactment:** How do groups like The Sealed Knot or Napoleonic reenactors utilize platforms like Instagram? Does the algorithmic pressure for visual "likes" alter their adherence to historical accuracy in similar ways?

2. Aesthetic Subcultures: An analysis of movements like #Cottagecore or #DarkAcademia on TikTok, which similarly construct idealized, non-historical pasts through digital-material consumption.
3. Algorithmic Politics of Emotion: Further research is needed on how "soft" cultural governance operates through platform architectures. How do algorithms implicitly encourage certain emotional structures (e.g., nationalism, nostalgia) while suppressing others?

By examining these parallel phenomena, researchers can better understand how digital platforms are reshaping collective memory and the very ontology of "tradition" in a globally connected society.

References

- Ahmed, S. (2004). Affective Economies. *Social Text*, 22(2), 117–139.
- Ahmed, S. (2014). *The cultural politics of emotion* (2nd ed.). Edinburgh University Press.
- Appadurai, A. (1986). *The Social Life of Things: Commodities in Cultural Perspective*. Cambridge University Press.
- Banet-Weiser, S. (2012). *Authentic™: The politics of ambivalence in a brand culture*. NYU Press.
- Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, 11(6), 985-1002.
- Berlant, L. (2011). *Cruel Optimism*. Duke University Press.
- Boellstorff, T. (2015). *Coming of age in Second Life: An anthropologist explores the virtually human*. Princeton University Press.
- Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. Routledge & Kegan Paul.
- Bouvier, G., & Machin, D. (2021). What gets lost in Twitter 'cancel culture' hashtags? Calling out racists reveals some limitations of social justice campaigns. *Discourse & Society*, 32(3), 307-327.
- Bucher, T. (2018). *If... Then: Algorithmic power and politics*. Oxford University Press.
- Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. Routledge.
- Carrico, K. (2017). *The Great Han: Race, nationalism, and tradition in China today*. University of California Press.
- Davis, M., Xiao, J., & Wei, L. (2025). Soft nationalism in China: The case of Hanfu style. *European Journal of Cultural Studies*.
<https://doi.org/10.1177/13675494251336606>

- De Kloet, J., & Fung, A. Y. (2017). *Youth cultures in China*. Polity Press.
- Fan, C., & Ip, P. T. T. (2023). "When showing Hanfu to foreigners, I feel very proud": The imagined community and affective economies of Hanfu among Chinese migrant youth in the United Kingdom. *European Journal of Cultural Studies*, 26(6), 781–802.
- Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. Boczkowski, & K. Foot (Eds.), *Media Technologies: Essays on Communication, Materiality, and Society* (pp. 167–194). MIT Press.
- Hutcheon, L. (1985). *A theory of parody: The teachings of twentieth-century art forms*. Methuen.
- Ip, P. T. T., & Fan, C. (2024). Reinventing Fashion as an Affective Apparatus by China's Youth. *Emerging Media*, 2(1), 109-133.
- Liu, F. (2012). Youth, Social Change and the Politics of Han Clothing in China. *The China Journal*, 68, 137-157.
- Pink, S., Ardèvol, E., & Lanzeni, D. (Eds.). (2016). *Digital materialities: Design and anthropology*. Bloomsbury Academic.
- Ritzer, G., & Jurgenson, N. (2010). Production, Consumption, Prosumption: The Nature of Capitalism in the Age of the Digital 'Prosumer'. *Journal of Consumer Culture*, 10(1), 13–36.
- van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford University Press.
- Wang, E. N. (2024). Participatory Censorship With Illusory Empowerment: Algorithmic Folklore and Interpretive Labor beyond Fandom. *Social Media + Society*, 10(4).
- Zulli, D., & Zulli, D. J. (2022). Extending the Internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform. *New Media & Society*, 24(8), 1872-1890.

2C.3 Out for the Algorithm: Queer Visibility and Self-Presentation on TikTok

Irina Bănoiu (Institute for Advanced Studies Vienna [IHS])

Introduction

Digital platforms have become central infrastructures of contemporary public spheres, actively shaping social interaction, cultural representation, and political participation. Social media platforms offer new opportunities for visibility, particularly for minoritised groups. However, rather than functioning as neutral spaces for expression, platforms govern visibility through socio-technical logics that structure participation, content circulation, and recognition, such as algorithmic selection and popularity metrics (van Dijck et al., 2018; van Dijck & Poell, 2013). For queer persons' struggles over visibility hold particular political relevance, thus aligning with or negotiating against platform norms, affordances, and economic incentives can therefore have salient implications.

TikTok constitutes an important site for examining such tensions. Its audiovisual affordances, trend-based cultures, and algorithmic feed shape both content circulation and identity expression. Existing studies describe TikTok as a meaningful space for queer self-affirmation, enabling visibility, counterpublic discourses, and solidarity (Jamet-Lange & Duguay, 2024), while also highlighting opaque governance that enables censorship (e.g. shadowbanning) and harassment (Rauchberg, 2022). This volatile visibility is deeply ambivalent: it can foster recognition, community-building, and resistance, while simultaneously heightening vulnerability. These dynamics are further intensified in socio-political contexts marked by legal and cultural marginalisation of queer people, such as the Eastern European context, where pressures to conform to heteronormative standards is higher. This paper conceptualises visibility as a dynamic and uneven process emerging from the interplay between platform governance and heteronormative constraints. Queer TikTok content is thus understood as a form of attempted resistance to both algorithmic and societal erasure.

Responding to the call of Poell and colleagues (2025) to expand platform culture research beyond Western contexts, this paper focuses on Romanian-speaking queer TikTok creators. Romania is a particularly relevant case due to the scarcity of public spaces for queer expression, prompting reliance on digital platforms to contest invisibility. TikTok functions here as a counterpublic arena, yet one constrained by platform governance and algorithmic regulation, producing tensions between empowerment and control. Thus, Sandoval's (2000) feminist modes of resistance are mobilised as sensitising concepts to analyse how creators strategically shift between assimilative and transformative practices. Adapted to the platform context, this

framework captures subtle, everyday forms of resistance (Scott, 1985) understood as minor acts that are situated, individual, and often compliant with platform rules shaped by algorithmic governance and commercial logics.

Objectives/Research Questions

The objective of this paper is to examine how TikTok creators, who position themselves at the intersection of queer and Romanian subjectivities, construct and negotiate their online presentation in the context of platformised (in)visibility and societal (hetero)normativity. It asks the following research questions:

1. How do the three most popular accounts of Romanian-speaking TikTok creators, who position themselves as queer, present their identities and experiences?
2. How are these self-presentation practices shaped by, aligned with, or resistant to TikTok's visibility regimes?
3. How are these self-presentation practices shaped by, aligned with, or resistant to heteronormative standards?

Ethical considerations include reliance on publicly accessible content, anonymisation (blurring images and paraphrasing quotes), and sensitivity to the risks associated with queer visibility in hostile socio-political environments.

Methods

The study employs a qualitative digital ethnographic approach, combining non-participatory observation of TikTok content with interpretive multimodal content analysis (MMCA).

A dedicated research account was created and used over a six-month period (June-December 2025) to train the algorithm towards content matching the study's sampling criteria. Metadata was collected in December 2025 using the Zeeschuimer scraping tool developed by the Digital Methods Initiative (2021/2025), employing a Romanian IP address and a set of searchable terms and tags related to Romanian and queer identifiers (e.g., gay/trans/queer + Romania; #lgbtromania; #wlvromania). The resulting dataset consists of 882 entries, organised in spreadsheets and analysed quantitatively to identify the three most visible accounts for in-depth qualitative analysis, based on cumulative views, likes, and comments.

Non-participatory observation, also described as "lurking", is currently being conducted following Schellewald (2021), to document patterns of self-presentation, content style, audio-visual elements, and narrative strategies. Observational notes are being systematically recorded in a spreadsheet serving as a field diary, and exemplary screen captures are occasionally collected for analytical reference.

For MMCA, videos will be coded according to content, form, and stance, including participation structures, keying, and communicative functions, using Shifman's (2013)

method, in order to obtain an understanding of the field and to answer the first research question. To evaluate the influence of TikTok's affordances on the production of content and the construction of self-presentations and to answer the second question, particular attention will be paid to trends and memetic dimensions of the content, following Literat and Kligler-Vilenchik (2019). To address the third research question, content will be coded and analysed following Sandoval's (2000) feminist modes of resistance: equal-rights, revolutionary, supremacist, separatist, and differential consciousness.

Results

As data collection and analysis are ongoing, the following findings are considered to be preliminary based on sustained engagement with the field. So far, observations suggest that queer Romanian-speaking TikTok creators display a range of self-presentation practices that reflect continuous negotiation with social normativity expectations and platformised visibility regimes.

Applying Sandoval's resistance forms, across the sample differential consciousness—theorised as movement in between the other forms—is most prevalent, manifesting through strategic shifts between covert and overt forms of queer visibility. This adaptive form of resistance is hypothesised to help creators navigate competing demands of visibility, safety, and recognition.

Some creators foreground everyday aspects of queer relationships and domestic life, aligning with an equal-rights resistance mode that emphasises similarity and normalisation. Others alternate between subtle and explicit critique, resonating with revolutionary and separatist resistance modes that seek to challenge dominant narratives, foster counterpublic discussion, or celebrate difference. These strategies are enacted through the use of platform affordances such as trends and popular sounds, often with humour, irony, and satire to render queer life visible, relatable, and culturally legible. This aligns with previous research on visibility labour (Abidin, 2016, 2020).

Guided by the concept of platformised (in)visibilisation, the preliminary analysis suggests that creators modulate the explicitness of queer expression—alternating between identity-focused content and more ambiguous or trend-aligned performances. Overall, the findings suggest that resistance on TikTok operates through adaptive, hybrid strategies, with visibility labour playing a central role in navigating platform constraints.

Future Work

Future stages of the research will focus on finalising the data collection and analysis process. Analytical categories will be refined, and the role of platform governance,

algorithmic visibility, and heteronormative social norms in shaping queer self-presentation will be further interrogated. As data collection and analysis progress, the study will deepen comparative attention to variations in visibility strategies across content types, formats, and degrees of explicitness. The paper will be developed into a full journal article by situating the empirical findings within debates on platform power, visibility labour, and digital public spheres.

Ultimately, the study aims to contribute to feminist and queer scholarship and to advance debates on queer visibility and platform power by conceptualising queer TikTok self-presentation as an adaptive form of resistance through which creators enact differential modes of contesting heteronormativity while simultaneously accommodating the platformised and societal conditions that shape who becomes visible or influential.

References

- Abidin, C. (2016). Visibility labour: Engaging with Influencers' fashion brands and #OOTD advertorial campaigns on Instagram. *Media International Australia*, 161(1), 86–100. <https://doi.org/10.1177/1329878X16665177>
- Abidin, C. (2020). Mapping Internet Celebrity on TikTok: Exploring Attention Economies and Visibility Labours. *Cultural Science Journal*, 12(1), 77–103. <https://doi.org/10.5334/csci.140>
- Digital Methods Initiative. (2025). *Digitalmethodsinitiative/zeeschuimer* [JavaScript]. Digital Methods Initiative. <https://github.com/digitalmethodsinitiative/zeeschuimer> (Original work published 2021)
- Jamet-Lange, H., & Duguay, S. (2024). “How do we do that?” An analysis of TikToks by lesbians over age 30 representing sexual identity, lived experience over time, and solidarity. *Journal of Lesbian Studies*, 1–19. <https://doi.org/10.1080/10894160.2024.2369431>
- Literat, I., & Kligler-Vilenchik, N. (2019). Youth collective political expression on social media: The role of affordances and memetic dimensions for voicing political views. *New Media & Society*, 21(9), 1988–2009. <https://doi.org/10.1177/1461444819837571>
- Poell, T., Duffy, B. E., Nieborg, D., Mutsvairo, B., Tse, T., Arriagada, A., De Kloet, J., & Sun, P. (2025). Global perspectives on platforms and cultural production. *International Journal of Cultural Studies*, 28(1), 3–20. <https://doi.org/10.1177/13678779241292736>
- Rauchberg, J. S. (2022). #Shadowbanned. In P. Paromita, *LGBTQ Digital Cultures* (1st edn, pp. 196–209). Routledge. <https://doi.org/10.4324/9781003196457-15>

- Sandoval, C. (2000). *Methodology of the oppressed*. University of Minnesota Press.
- Scott, J. C. (1985). *Weapons of the Weak: Everyday Forms of Peasant Resistance*. Yale University Press. <https://www.jstor.org/stable/j.ctt1nq836>
- Shifman, L. (2013). Memes in a Digital World: Reconciling with a Conceptual Troublemaker. *Journal of Computer-Mediated Communication*, 18(3), 362–377. <https://doi.org/10.1111/jcc4.12013>
- van Dijck, J., & Poell, T. (2013). Understanding Social Media Logic. *Media and Communication*, 1(1), 2–14.
- van Dijck, J., Poell, T., & de Waal, M. (2018). Platform Mechanisms. In *The Platform Society: Public Values in a Connective World* (pp. 31–48). Oxford University Press. <https://doi.org/10.1093/oso/9780190889760.001.0001>

2C.4 Enhancing Micro-phenomenology with Qualitative, Semi-structured, Elicitation Interviews When Combined with Interpretative Phenomenological Analysis.

Joan Owen, Jaigris Hodson (Royal Roads University)

Introduction

This paper explores the benefits of enhancing the qualitative interview method known as micro-phenomenology (Heimann et al., 2025; Hogan et al., 2017; Petitmengin et al., 2019) with social media elicitation in order to better understand the emotions driving engagement on popular social platforms such as Facebook, Instagram, and YouTube. Beginning with Interpretative Phenomenological Analysis (IPA) interviewing techniques and using social media prompts as a form of micro-phenomenology allows researchers to see the connections between social media content, emotions, and user decision-making.

Objectives/Research Questions

As a qualitative technique for interviewing and the subsequent data analysis, micro-phenomenology focuses on the entire experience of the phenomena in question (Heimann et al., 2025). According to Heiman et al. (2025), the sum of the experience includes the bodily, mental, and emotional aspects of the experience of the phenomena in question. In other words, when used as an interviewing practice, micro-phenomenology allows for the interviewee to express in detail their physical, mental, and emotional experiences to the interviewer, who can then closely and interpretatively examine the data to be able to draw meaningful conclusions and recommendations.

The qualities of micro-phenomenology align nicely with a study of affective polarization. Affective polarization arises when ideological or political polarization becomes permeated with such strong negative feelings that it results in disgust or disdain for those with opposing viewpoints (Hodson et al., 2025). The use of IPA interviewing with social media elicitation focuses on the individual's lived experience of each moment spent scrolling on a platform. In this presentation, we will discuss the relevant strengths of this method for the social media research community, and then show, via a case study of affective polarization in a mid-size Canadian city, how this method can allow for unique insights into user affect.

Methods

The combined use of IPA and semi-structured, elicitation interviews builds on the work of Hogan et al. (2016) which describes the use of evocation (the calling forward) of a memory of the experience being studied. The concept of evocation is to prompt the

participant into a present state, as opposed to speaking in the past tense (Hogan et al., 2016). This technique was later renamed as micro-phenomenology (Heimann et al., 2025; Hogan et al., 2017; Petitmengin et al., 2019). The micro-phenomenological interview maintains that a large part of our experiences of phenomena are unnoticed and should be considered. The complete lived experience of a phenomena include the five senses, as well as the evocation of the imagination and memory that occur (Hogan et al., 2017; Petitmengin et al., 2019). Examples include the memories evoked when touching a piece of fabric from your childhood clothing, seeing a photograph, or smelling your grandmother's cookie recipe baking.

As opposed to evocation, by using Facebook pages and comment threads in real-time to call upon imagination and memory as described by Hogan et al. (2016), we employed social media elicitation to call upon imagination and speculation of what someone 'might' feel and do in response to polarizing social media content. Using Facebook pages and comment threads as elicitation prompts creates opportunity to deepen the exploration of the experience of encountering content on a specific social media platform. This technique also minimizes the potential for post-rationalization of the study participant's experience of the elicitation prompts (Hogan et al., 2016). By capturing responses in real time, we explore emergent feelings and actions taken by users in a naturalistic way.

While this method of social media elicitation is a relatively novel addition to IPA, it is not unprecedented in qualitative interviewing. For example, in their 2022 study, Rivera et al. (2022) employed the use of Facebook to explore Facebook users engagement with cancer prevention and screening information. Following Rivera et al., Houlden et al. (2021) used social media elicitation with semi structured qualitative interviewing to determine how people were engaging with COVID-19 misinformation on social media sites. These approaches are useful because they can provide context to user decision making, and how they relate to the content being placed in front of people every day, as a result of social media algorithmic curation.

The use of the semi-structured, qualitative interview is consistent with IPA standards (Smith et al., 2022) and allows the researcher to develop the circular, iterative, and interpretative relationship with the data that is fundamental to IPA (Owen, 2025; Smith et al., 2022). Enhancing the semi-structured, qualitative interview with social media elicitation prompts gives opportunity to add to the data collection both clarity and the promotion of more detailed interviews (Thille et al., 2021).

Results

Based in Central Alberta, Canada, we designed a study to identify the emotions and other unconscious drivers that influence people's experiences of polarizing content on Facebook; how people are making sense of their emotions; and what explanations

people give for these emotions (Owen, 2025). Facebook as the social media platform of choice was selected because 82% of online Canadian adults report having a Facebook account (Gruzd & Mai, 2025). The intention of our study was to elicit emotional responses by showing the participants three pre-selected, public Facebook pages that represented current, salient, and enduring divisive topics represented on Alberta Facebook pages.

In keeping with IPA recruitment standards, recruitment of participants should be small and purposeful so that a larger volume of data from each individual participant can be gathered and analyzed (Smith et al., 2022; Smith & Fieldsend, 2021). With that goal in mind, eight participants were recruited using referrals, opportunities from the primary investigators' contacts, and a seed and snowball approach (Smith et al., 2022). As a result of this recruitment strategy, the participants formed a particular demographic. The participants ranged in age from 25 years to 49+ years. Of the eight participants, 3 were male presenting and 5 were female presenting. The study participant group consisted of a civil servant, new teachers, a post-secondary student, a public facing figure, a seasoned educator, and a professional communicator.

Interview conversations were guided by the interview questions and Facebook elicitation prompts. To begin the interviews, the primary investigator logged into their Facebook account and opened the predetermined public pages one at a time. For each of the three pre-selected Facebook pages, together the and the primary investigator viewed the posts and scrolled through the comment threads until the participant came upon a comment that sparked an emotional reaction. At this point, they engaged in conversation led by the interview questions. These conversations drew out the participants' lived experiences that encompassed the meaningful stories of the participants' personal experience of emotions (Mueller, 2019; Squire, 2013).

The use of the elicitation prompts gave opportunity to explore the immediacy and visceral expression of the emotional responses to the Facebook posts and comment threads that the participants found polarizing. For example, participant Beth expressed feelings of anger when reading profane and accusatory posts. Comments that belittle what others have posted angered participant Leah. Using elicitation prompts also allowed the primary investigator to observe the participants physical expression of their emotions, such as lip pursing and fist clenching (Beth) and feeling exhausted (Tim, Shawna, Leah).

Using the combination of IPA and elicitation prompts produced several novel insights that may have otherwise gone unnoticed or not revealed.

Future Work

The insights revealed in the data by using this specific combination of qualitative interview techniques include the revelation that Facebook users respond with negative

emotions to Facebook content and then, make fundamental attribution errors about those who post in comment threads; that negative emotions also result in Facebook users' identity feeling threatened, and that Facebook is not considered to be an emotionally safe place for conversation or discussion on polarizing topics

The combination of IPA as the methodology and the methods of semi-structured, qualitative, elicitation interviews contribute to the methodological canon by demonstrating that, for digital researchers, a micro-phenomenology framework can be enhanced through the use of social media elicitation specifically. Inspired by micro-phenomenology and enhanced by the qualitative, semi-structured, elicitation interview, the method detailed in this paper endorses the notion that using social media elicitation prompts to bring the interview participants into the present moment to explore emotions or other nuanced experiences of social media use. The use of semi-structured interviews in a micro-phenomenology inspired work brings a lateral movement to the work of the previous authors by simply modifying the "free form" interview technique to one that is more focused and allows for the participants and the interviewer to remain focused on and honour the specific lived experience being investigated

References

- Gruzd, A., & Mai, P. (2025). The State of Social Media in Canada 2025. figshare. <https://doi.org/10.6084/m9.figshare.28830188.v1>
- Heimann, K., Nouwens, M., Saggurthi, S., & Dalsgaard, P. (2025). Micro-phenomenology as a method for studying user experience in human-computer interaction. Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, 1–17. <https://doi.org/10.1145/3706598.3714422>
- Hodson, J., Owen, J., & Homer-Dixon, T. (2025, January 7). Mapping solutions to affective polarization: An ecological review of the literature. The Proceedings of the 58th Hawaii International Conference on System Sciences (HICSS). Hawaii International Conference on System Sciences (HICSS), Waikoloa, Hawaii. <https://hdl.handle.net/10125/109136>
- Hogan, T., Hinrichs, U., & Hornecker, E. (2016). The Elicitation Interview technique: Capturing people's experiences of data representations. IEEE Transactions on Visualization and Computer Graphics, 22(12), 2579–2593. IEEE Transactions on Visualization and Computer Graphics. <https://doi.org/10.1109/TVCG.2015.2511718>
- Hogan, T., Hinrichs, U., & Hornecker, E. (2017). The visual and beyond: Characterizing experiences with auditory, haptic and visual data representations. Proceedings of

- the 2017 Conference on Designing Interactive Systems, 797–809.
<https://doi.org/10.1145/3064663.3064702>
- Mueller, R. A. (2019). Episodic narrative interview: Capturing stories of experience with a methods fusion. *International Journal of Qualitative Methods*, 18, 1609406919866044. <https://doi.org/10.1177/1609406919866044>
- Owen, N. J. (2025). Feeling polarized? An interpretative phenomenological analysis of Central Albertans feelings in response to polarizing content on Facebook [D.Soc.Sc., Royal Roads University (Canada)].
<https://www.proquest.com/pqdtglobal/docview/3253506418/abstract/C44BB3CAAC1D4D80PQ/2>
- Petitmengin, C., Remillieux, A., & Valenzuela-Moguillansky, C. (2019). Discovering the structures of lived experience. *Phenomenology and the Cognitive Sciences*, 18(4), 691–730. <https://doi.org/10.1007/s11097-018-9597-4>
- Rivera, Y. M., Moran, M. B., Thrul, J., Joshu, C., & Smith, K. C. (2022). When engagement leads to action: Understanding the impact of cancer (mis)information among Latino/a Facebook users. *Health Communication*, 37(9), 1229–1241.
<https://doi.org/10.1080/10410236.2021.1950442>
- Smith, J., & Fieldsend, M. (2021). Interpretative phenomenological analysis. In P. M. Camic (Ed.), *Qualitative research in psychology: Expanding perspectives in methodology and design* (pp. 147–166). American Psychological Association.
<https://doi.org/10.1037/0000252-008>
- Smith, J., Flowers, P., & Larkin, M. (2022). *Interpretative phenomenological analysis: Theory, method, and research* (2nd ed.). SAGE Publications Ltd.
- Squire, C. (2013). From experience-centred to socio-culturally oriented approaches to narrative. In M. Andrews, C. Squire, & M. Tamboukou (Eds.), *Doing narrative research* (2nd ed., pp. 47–71). SAGE Publications, Ltd.
<https://doi.org/10.4135/9781526402271>
- Thille, P. H., Rotteau, L., & Webster, F. (2021). More than words: Methods to elicit talk in interviews. *Family Practice*, 38(4), 545–547.
<https://doi.org/10.1093/fampra/cmab043>

Session 2D - Discourse and Public Opinion (2)

Day 1 · Mon 13 July · 11:30–13:00 · Studio 2, ARC

2D.1 Networked Hatred: Co-evolution of Dehumanising Discourse and Channel Structure of Russian and Ukrainian Telegram During the 2022 Invasion

Elizaveta Chernenko (Oxford Internet Institute), Balazs Vedres (Department of Network and Data Science, Central European University)

Introduction

This article investigates how dehumanising discourse emerged, spread and gained users' attention within Russian and Ukrainian Telegram channels as military hostilities unfolded at the onset of the Russia's 2022 full-scale invasion of Ukraine, from a networked approach. It examines the co-evolution of discourse and network structure at the channel and post levels, tracing how linguistic and structural transformations reinforced one another, and how these dynamics differed between the two countries.

We build on our prior work and a previously developed dataset, based on a psychologically grounded operationalisation of outgroup dehumanisation.

Dehumanisation is defined as perceiving or treating others as less than human (Haslam & Stratemeyer, 2016) and may be expressed through explicit animalistic or mechanistic comparisons (Haslam et al., 2008), as well as combination of more implicit forms such as denial of agency (Kelman, 1973), deindividuation (Haslam, 2006), extreme negative evaluation (Christen et al., 2002), denial of prosocial values (Schwartz & Struch, 1989), and expressions of humiliation or disgust (Haslam, 2006). Our research is focused on blatant forms of dehumanisation of an outgroup defined in national or political terms.

The study draws on Emirbayer and Sheller (1998) conceptualisation of publics as dynamic, networked spaces in which discourse and structure are mutually constitutive. In this framework, shifts in public narratives reshape patterns of interaction, while evolving network connections influence which frames gain prominence. This relational perspective provides a powerful lens for analysing digital publics, where rapid transformations in online discourse – such as the diffusion of dehumanising narratives – are closely intertwined with changing network topologies.

We develop two complementary models using posts from the most popular political and news Telegram channels in both countries (24 December 2021 – 24 April 2022), measures of their audience reception, and their cross-channel forwards. The first model analyses the time-series dynamics of public reception to the Telegram posts (views, reactions, and shares) by affiliation, discourse, and structural factors. The second model uses channels as the unit of analysis to measure the networked evolution of information dissemination (who forwards whom) and the evolution of channel attributes (including the share of dehumanising discourse they produce) as interdependent processes unfolding over time.

Preliminary findings reveal the growing centrality of dehumanising language within both Russian and Ukrainian Telegram channels, and its gradual co-transformation with the online public sphere through a recursive spiral of verbal violence. The methodology applied in this study reveals both the structural logic behind the diffusion of hateful language and the ideological drivers of network formation that shape discursive escalation in the wartime public sphere.

Objectives/Research Questions

The study addresses the following research questions:

RQ1: How was dehumanising discourse received and amplified within the public spheres of Russia and Ukraine across different stages of the 2022 invasion?

RQ2: How did the emergence of dehumanising frames co-evolve with the topology of the Telegram network over time? Did dehumanising language influence forwarding patterns between channels (RQ2a), and were there identifiable structural mechanisms associated with its production (RQ2b)?

Methods

We employ a previously developed dataset (n = 238,190) constructed through a multi-step process that combines manual annotation with Large Language Model-assisted classification to identify blatantly dehumanising, marginally dehumanising, and non-dehumanising posts from the most popular news and political Telegram channels in Ukraine and Russia. Based on descriptive analysis of dynamics in dehumanising language use, we identified four substantially different periods in the discourse evolution, including ‘Pre-Invasion’, ‘Invasion Shock’, ‘Early Adaptation’, and ‘Bucha Shock’, with the latter aligning with the emergence of widely circulated reports of the Bucha massacre in early April 2022.

This paper is currently in progress, with several models still under or awaiting estimation. Model specifications may be refined as the data analysis proceeds.

To assess how audiences responded to dehumanising language, we estimate a series of models predicting the count user engagement metrics (number of views, reactions, and forwards) using Generalized Linear Mixed Model (GLMM) framework (Breslow & Clayton, 1993) such as Multilevel mixed effects Poisson regression. This framework allows to account for the hierarchical group structure of the data (posts nested in channels) and overtime dynamics, as well as effects of presence of dehumanising language in a post.

There are several options for the model of networked evolution of dehumanising language. It can be implemented using Stochastic Actor-Oriented Model (SAOM) approach that enables analysis of parallel changes in network and discourse, accounting for structural effects, such as reciprocity and transitivity, as well as channel-level attributes, including country affiliation, language, type, public reception, and the share of dehumanising content (Snijders et al., 2010). General characteristics of the dataset, that is moderate network size, more than three meaningful time periods, considerably sparse network, sufficient variation in both structure and 'behaviour' (share of dehumanising posts) is favourable for successful estimation. To validate findings from SAOM model and in the scenario where the model will not converge, we will test the same effects using dynamic time-series frameworks such as Granger causality (Granger, 1969) and related dynamic panel models, which capture temporal interdependence between structural and behavioural changes. This approach and SAOM share a common logic: they treat the structure and content of communication as mutually shaping, temporally coupled systems, rather than independent or static components.

Results

This paper explores how dehumanising discourse and the network structure of Russian and Ukrainian Telegram channels co-evolved during the early stages of the 2022 invasion, and how such discourse was received and amplified by audiences of the two countries.

Preliminary findings suggest that dehumanising language gained traction as the conflict progressed, gradually growing its presence in wartime communication – especially within Russian networks.

Patterns of engagement indicate that audience responsiveness to such narratives also intensified over time, with such posts attracting higher numbers of views, reactions, and forwards. This trend emerged earlier and more strongly in Russian channels, while Ukrainian audiences exhibited a similar but delayed increase, suggesting asymmetries in how dehumanising narratives are perceived by the audiences in the aggressor and defender countries.

Network analysis indicates a systematic association between tendency to produce dehumanising discourse and network structure of channels. Telegram's forwarding network exhibits discourse-based homophily, with channels preferentially forwarding content from others producing similar levels of dehumanising language. Homophily values within strong-tie networks reach 0.247 for blatantly dehumanising and 0.408 for all non-neutral posts.

Channels producing higher (combined) share of dehumanising language also tend take more central positions in the network, illustrated by small, but significant positive correlation between dehumanisation and indegree centrality (0.198 and 0.217 for weighted and unweighted indegree, respectively).

These patterns are situated within broader structural and discursive dynamics. Specifically, the full-scale invasion intensified network polarisation, producing distinct Russian and Ukrainian clusters characterised by strong country-based homophily. Over time, both the volume of dehumanising content and the number of channels producing it increased. Russian channels disseminate substantially more dehumanising posts and occupy more central positions than Ukrainian ones. Preliminary evidence also suggests channel-type effects, with Russian state-affiliated and military blogger channels, as well as Ukrainian grassroots and regional channels, being particularly active in disseminating dehumanising content.

Future Work

While full estimation is pending, future model-based analysis will formally measure the interdependent dynamics of Russian and Ukrainian Telegram channels described above, providing a more robust account of how dehumanising discourse, network structure, and public reception co-develop under conditions of conflict.

References

- Breslow, N. E., & Clayton, D. G. (1993). Approximate Inference in Generalized Linear Mixed Models. *Journal of the American Statistical Association*, 88(421), 9–25.
<https://doi.org/10.1080/01621459.1993.10594284>
- Christen, C. T., Kannaovakun, P., & Gunther, A. C. (2002). Hostile Media Perceptions: Partisan Assessments of Press and Public during the 1997 United Parcel Service Strike. *Political Communication*, 19(4), 423–436.
<https://doi.org/10.1080/10584600290109988>
- Emirbayer, M., & Sheller, M. (1998). Publics in history. *Theory and Society*, 27(6), 727–779. <https://doi.org/10.1023/A:1006957015604>
- Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3), 424.
<https://doi.org/10.2307/1912791>

- Haslam, N. (2006). Dehumanization: An Integrative Review. *Personality and Social Psychology Review*, 10(3), 252–264.
https://doi.org/10.1207/s15327957pspr1003_4
- Haslam, N., Loughnan, S., Kashima, Y., & Bain, P. (2008). Attributing and denying humanness to others. *European Review of Social Psychology*, 19(1), 55–85.
<https://doi.org/10.1080/10463280801981645>
- Haslam, N., & Stratemeyer, M. (2016). Recent research on dehumanization. *Current Opinion in Psychology*, 11, 25–29. <https://doi.org/10.1016/j.copsyc.2016.03.009>
- Kelman, H. G. (1973). Violence without Moral Restraint: Reflections on the Dehumanization of Victims and Victimizers. *Journal of Social Issues*, 29(4), 25–61.
<https://doi.org/10.1111/j.1540-4560.1973.tb00102.x>
- Schwartz, S. H., & Struch, N. (1989). Values, Stereotypes, and Intergroup Antagonism. In D. Bar-Tal, C. F. Graumann, A. W. Kruglanski, & W. Stroebe (Eds), *Stereotyping and Prejudice* (pp. 151–167). Springer New York. https://doi.org/10.1007/978-1-4612-3582-8_7
- Snijders, T. A. B., van de Bunt, G. G., & Steglich, C. E. G. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, 32(1), 44–60. <https://doi.org/10.1016/j.socnet.2009.02.004>

2D.2 Speaking Back Between Misogyny and Censorship: Three Counterpublic Strategies by Chinese Women

Jiahao Sun (Hitotsubashi University)

Introduction

Recent years have witnessed social media platforms' progress. Specifically, they have emerged as critical sites for gender-political conflict. While digital spaces have given unprecedented visibility to issues concerning women's bodies, rights, and social roles, they have also exposed women to a surge of misogynistic discourse, anti-feminist attacks, and nationalist encirclement. Existing scholarship has extensively discussed digital feminist mobilization and online misogyny. What should be noticed is that, the literature largely focuses on liberal democratic contexts and global platforms such as Twitter, Reddit, or YouTube (Baer, 2016; Yoder et al., 2023). There remains a paucity of empirical research and theoretical reflection on how women resist online misogyny and develop counterpublics in non-Western contexts characterized by the coexistence of systemic censorship and algorithmic governance.

This paper delves into a public controversy on China's Weibo concerning "whether high-speed trains should sell sanitary pads", so as to the gap mentioned above. This dispute originated when a woman shared her personal experience of unexpectedly getting her period on a train and being unable to find sanitary products. Her post quickly ignited a fierce and polarized debate. In the ensuing discussion, prominent male anti-feminist influencers publicly opposed selling pads on trains. They sought to invalidate the demand by employing "menstrual stigma" or framing the request as "Western ideology." In response, Chinese feminist counterpublics constructed three distinct discursive strategies to sustain and propagate their demands. This case provides a representative entry point for analyzing how online misogyny is constructed in the Chinese context and how women respond within a highly restricted digital environment.

China offers a particularly significant case for examining this phenomenon. Culturally, enduring and deep-rooted Confucian gender norms continue to exert a profound influence on social understandings of female bodies and emotions. Besides, practices such as menstrual stigma persist in casting female physiological experiences as "private," shameful, and unsuitable for public discussion (Chu, 1980). Politically, feminist appeals are often reframed as foreign, Western, or ideologically threatening. Women who voice these demands often face nationalist pushback, which seeks to undermine their legitimacy (Huang, 2023). Digitally, Chinese social media operates within a nexus of algorithmic governance and state censorship, where the public sphere exhibits intermittent visibility, sudden disappearances, and highly asymmetrical

power structures (Liu, 2025). Under such conditions, women cannot rely on open protest, institutional advocacy, or sustained public debate. Instead, they must negotiate limited spaces for expression through strategic language use, affective framing, and the strategic drawing of discursive boundaries.

Objectives/Research Questions

This study aims to explore how Chinese women exercise discursive agency and respond to online misogyny within a digital environment where censorship intersects with gendered power structures. The specific research questions are as follows:

- 1) What forms of online misogyny emerged in the Weibo controversy over the sale of sanitary pads on high-speed trains?
- 2) What specific discursive strategies did Chinese women employ to respond to or resist these misogynistic attacks?
- 3) What defines the nature of counterpublics in China?

By addressing these questions, this paper seeks to extend Fraser's counterpublic theory by situating it within an authoritative and highly regulated media context (1990). It argues that China, counterpublics cannot form through open and democratic debate. In fact, they have frame their discourse within state-approved boundaries to avoid censorship. Furthermore, the cohesion of these counterpublics stems not from a clear political manifesto, but from shared affective experiences. Last but not least, they operate through decentralized networks—such as comment sections, threads, and hashtags—rather than centralized organizations.

Methods

This study employs a qualitative research approach based on data collected from Weibo. Data was gathered between July and August 2023, which centered on posts from two prominent anti-feminist influencers involved in the controversy, as well as the original media post that ignited the public discussion. For each of these posts, the top 50 most-liked comments were selected. After excluding advertisements, obvious duplicates, and comments unrelated to the event, a final sample of 100 representative comments was retained for analysis.

The methodology combines Thematic Analysis (TA) with Critical Discourse Analysis (CDA). Following Braun and Clarke's framework, Thematic Analysis was used to identify recurrent discursive patterns and strategies used by female users in response to misogynistic discourse. Critical Discourse Analysis was then applied to situate these strategies within broader social structures. This approach is intended for examining how gender hierarchies, nationalist discourse, and platform governance are reproduced or challenged through linguistic practices.

Results

The study identified two intertwining forms of misogynistic discourse within the controversy. The first is menstrual stigma, which stems from deep-seated cultural views in China that treat menstruation as a private and unclean matter that should not be discussed publicly. This discourse was evident in the official statement by China Railway defining sanitary pads as "private items," and in male influencers who juxtaposed "menstruation" with "public consumption spaces" (like dining cars) to cultivate a sense of the "unclean." They further criticized women for "not planning ahead", in order to shift responsibility onto them.

The second is politicized misogyny. In this line of attack, women's gendered demands are reframed as the importation of "Western values" and accused of threatening national culture and identity. Labels such as "baizuo" (white left), "nvquan" (a derogatory term for radical feminists), and "jingmei" (spiritually American) were used to erode the legitimacy of women's voices, which painted them as ideologically suspect or disloyal.

Through thematic analysis, this study identified three primary counter-discursive strategies employed by women:

1)"Half the Sky" Universalizing Discourse: Women re-appropriated the state-sanctioned rhetoric that "women hold up half the sky". They used this concept to reframe menstrual needs as a general public issue rather than a "minority" concern or identity politics. This strategy not only lowered censorship risks but also established emotional resonance through familiar symbolic language.

2)Empathy-Driven Discourse: By using analogies (comparing menstruation to uncontrollable physiological needs like using the toilet or eating) and appealing to familial roles (mothers, daughters, wives), women sought to elicit emotional understanding from male users. In this context, menstruation as were redefined as a "normal and public" experience.

3)Experience-Based Exclusionary Discourse: Women directly challenged men's epistemic authority on menstrual issues by pointing to their lack of personal bodily experience. This approach did not seek consensus; rather, it sought to construct a relatively safe, female-centered discussion space by establishing discursive boundaries.

Future Work

Despite providing a comprehensive analysis of female counter-misogyny practices on Chinese social media, this study comes with limitations. First and foremost, it relies on a single event and a limited sample size, which may not fully capture the diverse morphology of Chinese digital feminism. Future research could extend to different platforms, events, or longer timeframes to capture a richer array of discursive practices. Second, comparative studies—particularly between different authoritarian

or semi-authoritarian contexts— could improve our understanding of how political environments shape pathways of digital resistance. Last but not least, integrating qualitative discourse analysis with computational methods could further reveal the scale of propagation and the structural characteristics of these counter-discursive strategies.

References

- Baer, H. (2016). Redoing feminism: digital activism, body politics, and neoliberalism. *Feminist Media Studies*, 16(1), 17–34.
<https://doi.org/10.1080/14680777.2015.1093070>.
- Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. Sage Publications.
- Chu, C. M.-Y. (1980). Menstrual Beliefs and Practices of Chinese Women. *Journal of the Folklore Institute*, 17(1), 38–55. <https://doi.org/10.2307/3814221>
- Fraser, Nancy. 1990. “Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy.” *Social Text*, no. 25/26: 56.
<https://doi.org/10.2307/466240>.
- Huang, Q. (2023). Anti-Feminism: Four strategies for the demonisation and depoliticisation of feminism on Chinese social media. *Feminist Media Studies*, 23(7), 3583–3598. <https://doi.org/10.1080/14680777.2022.2129412>
- Liu, M. (2025). #MeToo in China: Visceral accounts of rape culture, a non-domestic feminist counterpublic, and networked solidarity. *Feminist Media Studies*, 25(2), 357–372. <https://doi.org/10.1080/14680777.2023.2285718>
- Yoder, M. M., Perry, C., Brown, D. W., Carley, K. M., & Pruden, M. (2023). Identity construction in a misogynist incels forum. In *The 7th Workshop on Online Abuse and Harms (WOAH)* (pp. 1–13). Association for Computational Linguistics.

2D.3 The Duality of Media Actors and Culture: The White Afrikaner Refugee Controversy on YouTube

Jihye Kim (Umass Amherst)

Introduction

In May 2025, fifty-nine White South Africans arrived in the United States as government-granted refugees. Public controversy further escalated soon afterward when Trump accused the South African government of permitting a “white farmer genocide” during a televised Oval Office meeting with President Ramaphosa. At the same time, South Africa’s role within BRICS+ and its diplomatic tensions with the United States heightened transnational attention from multiple countries. These conditions create a unique opportunity to examine how competing frames circulate, stabilize, and contend within a networked media environment. Importantly, this case foregrounds the role of heterogeneous media actors on YouTube in diffusing narratives of “white genocide.”

To analyze these dynamics, this study adopts a duality of media actors and culture perspective, extending the sociological insight that individuals and culture are mutually constitutive (Breiger, 1974; Pachucki & Breiger, 2010) to contemporary digital ecosystems. Meaning does not flow unidirectionally from mainstream news media to passive audiences. It emerges through overlapping discourses among diverse actors operating at micro, meso, and macro levels of communication. Cultural network research provides tools for tracing these processes. Within this approach, shared frames are treated as relational structures embedded in text, enabling analysis of how meanings align, diverge, and stabilize across heterogeneous actors (Bail, 2016; Shugars & González-Bailón, 2024; Wu & Xu, 2024).

This perspective directly addresses limitations in existing communication research on frame diffusion. Much prior work relies on hashtag-based data collection and user-to-user networks, treating frames as if they spread through interpersonal contagion. Such approaches are ill-suited to YouTube, where discourse is organized around spoken content rather than hashtags and where circulation depends less on dyadic ties than on channel-level visibility and cross-media reinforcement. As a result, previous research has revealed patterns of connection without capturing how interpretive frameworks gain durability through repeated alignment across media actors. By analyzing video transcripts and modeling channel-frame bipartite networks, this study conceptualizes frame diffusion as a cultural process driven by reinforcement and alignment rather than by simple information transfer (Guilbeault, Becker & Centola, 2018; Centola, 2015). Frames are therefore defined here as interpretive, stance-infused meaning structures through which channels construct and contest social realities (Berger & Luckmann,

1966; Goffman, 1974). Topics indicate what discourse is about, whereas interpretive frames reveal how issues are evaluated and understood. Building on this distinction, the study examines how actors with different levels of visibility participate in frame circulation around a racially charged and geopolitically situated controversy.

Objectives/Research Questions

The objective of this study is to examine how interpretive frames related to the White Afrikaner refugee controversy are produced, circulated, and stabilized within a transnational YouTube media ecosystem. Drawing on a duality of media actors and culture perspective, the study conceptualizes frame diffusion as a cultural process shaped by overlapping discourses among heterogeneous media actors with unequal visibility and influence. Specifically, the study aims to identify which frames are more likely to gain durability and prominence, which types of channels play key roles in initiating, stabilizing, and bridging these frames, and how these dynamics vary across different regions. Through a combined analysis of interpretive frames and inter-channel networks, the study seeks to clarify how fringe narratives become normalized within the transnational YouTube media ecosystem.

RQ1. What types of frames concerning White South Africans and the Afrikaner refugee controversy are more likely to be sustained and amplified across YouTube channels?

RQ2. What types of channels facilitate the initiation, stabilization, and bridging of these frames within the transnational media ecosystem?

RQ3. How do patterns of frame diffusion differ among channels from South Africa, the United States, the Global North, and the Global South?

Methods

To examine frame diffusion on YouTube, I collected video- and channel-level information, including subscriber counts and stated locations, using the YouTube API from 2022 through June 2025. Search queries covered “Afrikaner,” “white south African,” and related terms. The initial dataset contained 1,021 videos. Video transcripts were generated using GPT-based audio extraction tools. To ensure comparability across highly imbalanced content length, I standardized transcription by extracting the first 1,395 seconds of audio from each video. All transcripts were processed as multilingual text.

Each channel was assigned three attributes: impact level, media type, and location. Impact levels were defined based on subscriber counts (1,000-99,999; $\geq 100,000$). Media types were manually coded into three categories: mainstream, personal brand channels, and alternative media. To validate media type classification, Gemini 2.5 Pro (hereafter, Gemini) independently coded all channels; intercoder reliability between Gemini and the human coder reached 80.2% (Cohen’s κ). For channels with missing

location information (89 out of 527 channels), Gemini conducted web-based inference to assign plausible geographic origins.

Representative frames were extracted through a two-step LLM (Large Language Model)-assisted qualitative procedure. First, Gemini conducted interpretive readings of transcripts to identify preliminary topics. Second, each topic was classified according to stance. To capture distinct interpretive meaning structures, frames were operationalized as topic-stance pairs (e.g., White Genocide Narrative_Affirming Genocide Claims). After manual review, a unified human-AI codebook of 30 interpretive frames was developed. For frame assignment, Gemini was instructed to assign up to three frames per video using both transcript excerpts and video descriptions.

To address RQ1, frame occurrence counts were analyzed across periods to identify sustained and amplified frames, followed by the construction of inter-frame networks to assess the co-presence and competitive relations among opposing frames. To address RQ2 and RQ3, inter-channel networks based on shared frames were constructed to identify initiators, stabilizers, and bridges within each period. Initiators were defined as early frame introducers, operationalized as those in the top 3% of first mentions per frame. Stabilizers were identified as channels ranking in the top 30% of weighted degree centrality within same-media-type subnetworks, while bridges were defined as channels exhibiting both (1) positive E-I index with respect to media type and (2) top 30% betweenness centrality.

Results

The analysis reveals three main findings regarding frame competition, media roles, and transnational diffusion dynamics. First, frame level analysis shows that right-wing and left-wing interpretive frames compete evenly. Contrary to expectations that right-wing narratives would dominate discourse around the Afrikaner controversy, evidence from frame occurrence counts and inter-frame networks indicates sustained visibility of both supportive and critical frames over time.

Second, results point to a clear division of labor across media types in the circulation of frames. In later periods, frame initiation and stabilization become increasingly centralized within U.S.-based mainstream media channels. These actors play a major role in consolidating interpretive frames. At the same time, personal brand channels and alternative media disproportionately occupy bridging positions. This reflects a process of fringe narrative mainstreaming.

Third, the inter-channel network analysis reveals a pronounced temporal and geopolitical shift. In period 4, before the arrival of the refugees, South African and Global South channels occupy central positions. By period 5, after the arrival, U.S.-based channels increase in both number and centrality, particularly as stabilizers of dominant frames. By Period 6, following the Oval Office meeting, interpretive authority

becomes concentrated among U.S. and Global North channels, which cluster at the network core. In contrast, South African and other Global South actors are displaced to fragmented peripheral positions rather than forming a unified counterpublic. Taken together, these findings suggest that Global South channels struggle to sustain coherent counter-narratives to dominant frames. Instead, frame diffusion within YouTube shows competition between white defensive public narratives (Jackson & Kreiss, 2023) and elite-mediated discourses circulating through mainstream media, while voices from marginalized groups remain weakly amplified on YouTube.

Future Work

This study opens several avenues for future research. First, to further validate the robustness of the frame extraction and diffusion patterns identified here, I will compare the Gemini-based interpretive frame classification with alternative computational approaches, including word-embedding and open-source LLMs. This validation step will help clarify whether results reflect stable underlying dynamics. Second, this research design can be extended to comparative analyses of other transnational controversies. By applying the same LLM-assisted frame extraction and channel-frame network framework to additional cases, future studies can examine whether the characteristics of frame diffusion observed here generalize beyond the White Afrikaner issue. More broadly, this approach offers a transferable methodological template for communication research in non-text-based platforms. By combining LLM-applied qualitative frame analysis with cultural network analysis, future work can better capture how interpretive frames are constructed, consolidated, and contested in contemporary media ecosystems.

References

- Bail, C. A. (2016). Combining natural language processing and network analysis to examine how advocacy organizations stimulate conversation on social media. *Proceedings of the National Academy of Sciences*, 113(42), 11823-11828.
- Berger, P., & Luckmann, T. (1966). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. New York: Doubleday.
- Breiger, R. L. (1974). The duality of persons and groups. *Social forces*, 53(2), 181-190.
- Centola, D. (2015). The social origins of networks and diffusion. *American journal of sociology*, 120(5), 1295-1338.
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Northeastern UP.
- Guilbeault, D., Becker, J., & Centola, D. (2018). Complex contagions: A decade in review. *Complex spreading phenomena in social systems: Influence and contagion in real-world social networks*, 3-25.

- Jackson, S. J., & Kreiss, D. (2023). Recentring power: Conceptualizing counterpublics and defensive publics. *Communication Theory*, 33(2-3), 102-111.
- Pachucki, M. A., & Breiger, R. L. (2010). Cultural holes: Beyond relationality in social networks and culture. *Annual review of sociology*, 36(1), 205-224.
- Shugars, S., & González-Bailón, S. (2023). Semantic and cultural networks. *The Sage Handbook of Social Network Analysis*, 202.
- Wu, V. C. S., & Xu, W. W. (2024). Who leads and who echoes? Tracing message similarity network of #ClimateChange advocacy on Twitter. *Nonprofit and Voluntary Sector Quarterly*, 53(2), 372-401.

2D.4 How Topics Flow Across Reddit Communities Through Shared User Participation: Comparative Study of Pre and Post COVID-19 Vaccine Discourse

Joohee Kim (University of Massachusetts)

Introduction

Reddit is a digital media platform where users share and discuss content in topical communities called subreddits. Each subreddit is organized through community moderation, anonymity, and salience shaped by user upvotes (Caplan, 2018; Habib et al., 2024; Prakasam & Huxtable-Thomas, 2021). These affordances construct subreddit-specific discursive practices, ultimately shaping the topics and culture of each community. However, because Reddit's design encourages interaction to remain largely within individual subreddits, most research on Reddit discourse focuses on community-specific characteristics and pays limited attention to inter-subreddit interactions. This presents a gap in the literature, given that users freely participate in one subreddit and another.

A small body of subreddit-level research has demonstrated that relationships between subreddits exhibit structured patterns, including conflict and modular communities (Datta et al., 2017; Olson & Neal, 2015). However, what these subreddit-to-subreddit connections mean for the topical communities is less investigated. Prior research on homophily between subreddits found meaningful homophily between communities and presents that it has predictive power in which community individuals choose to participate (Massachs et al., 2020; Valensise et al., 2019; Yuan et al., 2025). The patterns of homophily showed that those with author similarity share term similarity, suggesting that subreddit topical homophily is constructed by membership (Datta et al., 2019). A combination of these studies suggests that subreddit-level structural relationships, based on shared members, are noteworthy for understanding the construction of discourse in these topical communities.

This study suggests that shared member network constitute a discursively meaningful environment that shapes what content users encounter across communities. The structural pattern of shared participation across subreddits, which reshapes the relational environment in which discourse is constructed, can be theorized by drawing on networked gatekeeping. Networked gatekeeping emphasizes the gatekeeping as a dynamic and relational process between the gatekeeper and the gated within the network (Barzilai-Nahon, 2008). In this framework, ordinary participants within a subreddit may function as gatekeepers and gated as they collectively shape visibility, salience, and topical focus. Therefore, when participants engage across multiple

subreddits, their participation contributes to the homophily between communities by linking distinct topical spaces through shared membership. Because subreddits are organized as topical communities, such structural alignment is discursively meaningful, as it is the environment within which gatekeeping processes emerge across communities.

Objectives/Research Questions

This study investigates subreddit-to-subreddit networks on Reddit to examine how shared membership structures shape cross-community relationships in discourse. Rather than directly analyzing gatekeeping acts, this analysis identifies structural conditions of exposure and brokerage that can emerge across topical communities. This approach offers two contributions: first, it provides a structural understanding of communities on Reddit beyond individual subreddits; second, it foregrounds how unintentional participation patterns reorganize the relational terrain in which gatekeeping operates.

For the purpose of empirical research, this study focuses on vaccine discourse, as it is a topic with growing controversies surrounding anti-vaxx and vaccine-hesitant movements (WHO, 2019). Because vaccine discourse is highly polarized and politicized, it provides a suitable case for examining how the gatekeeping dynamics operate across subreddits via networked positions. There are three reasons for focusing on vaccine discourse at this moment. First, the topic has undergone significant events in recent years, including the COVID-19 pandemic, global vaccine rollout, and the AstraZeneca recall. Second, public reliance on online discourse to collect health-related information has increased. Finally, the presence of relatively clear stance distinctions enables investigation of how subreddit-level gatekeeping may be interpreted.

Therefore, this study investigates how the subreddit-level network surrounding r/DebateVaccines changes before, during, and after COVID-19. Based on previous studies, I expect the network to reveal two important trends. First, over time, r/DebateVaccines will exhibit increased homophily with subreddits associated with anti-vaxx and vaccine-hesitant discourse. Second, as vaccine discourse becomes increasingly politicized and polarized, I expect the subreddit will show stronger homophily with conservative subreddits after COVID-19 period (Howard, 2024).

Methods

To address the proposed research question, I focus on the subreddit r/DebateVaccines. Created in 2017, the subreddit has approximately 23,000 members and describes themselves as “debate and discuss issues surrounding vaccination.” This positioning suggests that there is no presupposed alignment with either pro- or anti-vaccine discourse, making it well-suited for this study.

I collected posts and comments from r/DebateVaccines across three periods: December 2019, December 2022, and October 2024, with each period spanning one month. Data was collected using Arctic-Shift API and Python. From this data, I identified users participating in vaccine discourse on r/DebateVaccines and retrieved their complete post and comment history during the same periods. Using this information, I constructed an egocentric network on r/DebateVaccines, identifying other subreddits that share members with it.

For each identified subreddit, I then collected all posts and comments of these subreddits within each period to construct a complete shared-membership network among these communities. In the final network, nodes represent subreddits and edges represent a normalized number of shared members. The number of shared members was normalized using cosine similarity to account for differences in subreddit size. The resulting networks were analyzed by creating clusters and calculating betweenness centrality to understand who exhibits strong homophily with r/DebateVaccines and which topical communities mediate the discourse.

Results

Overall, r/DebateVaccines exhibits increasing homophily with anti-vaxx and conservative subreddits over time, supporting both hypotheses. Importantly, structural shifts are observable without direct analysis of textual content, demonstrating that shared membership networks alone can reveal meaningful discursive alignment and polarization.

From 2019 to 2024, r/DebateVaccines shows increasing homophily with anti-vaxx subreddits and presents crucial changes in network position. Across all three periods, r/DebateVaccines belongs to a cluster that consists of other subreddits that discuss vaccine. However, the topics of connected subreddits and their positions change from pro-vaccine to anti-vaxx over time. In 2019, r/DebateVaccines is primarily connected to pro-vaccine subreddits. By 2022, it is connected to anti-vaxx and vaccine-hesitant subreddits. This pattern persists in 2024, with anti-vaxx subreddits positioned as a bridge that links the vaccine discourse subreddit with the other topical subreddits. These findings suggest that, in the post-COVID period, anti-vaxx subreddits occupy structural positions consistent with gatekeeping or brokerage roles.

The connection from r/DebateVaccines with other political topics also presents a clear change from liberal to conservative in its closeness and position across the period. In 2019, political subreddits are few and positioned distant from the r/DebateVaccines, with liberal-oriented subreddits more prevalent than conservative ones. By 2022, conservative subreddits become more numerous and are positioned closer to r/DebateVaccines. This trend intensifies in 2024, when conservative subreddits

dominate the political portion of the network and are positioned closest to r/DebateVaccines, while liberal subreddits become comparatively sparse.

Beyond hypothesis-driven findings, the networks reveal additional structural features of the vaccine discourse. First, the subreddits with high betweenness centrality shift from predominantly entertainment-focused subreddits to conspiracy subreddits. Second, as anti-vaxx, conservative, and conspiracy subreddits appear in 2022, the 2024 network shows emerging ties to alternative-medicine and parenting subreddits, signaling alignment with anti-vaxx narrative.

Future Work

These results raise three points of discussion. First, the alignment between shifts in network structure and shifts in discourse suggests that shared membership facilitates the transfer of ideas and values across subreddits, even when communities are not explicitly related by topic. This highlights the value of subreddit-level network analysis for understanding discursive dynamics. Second, the findings suggest that structural positions of the topical communities can enable unintentional gatekeeping, pointing to the need to extend the existing theory of networked gatekeeping beyond intentional actions. While prior work emphasizes deliberate gatekeeping practices, these results indicate that structural alignment alone can produce gatekeeping effects that shape discourse. Finally, because these processes operate unintentionally, users who do not actively seek polarized content may nonetheless encounter it when participating in seemingly neutral topical communities. This raises the possibility that subtle structural pathways of information flow can be exploited to amplify polarization.

References

- Barzilai-Nahon, K. (2008). Toward a theory of network gatekeeping: A framework for exploring information control. *Journal of the American Society for Information Science and Technology*, 59(9), 1493–1512.
- Caplan, R. (2018). Content or context moderation? Data & Society Research Institute.
- Datta, S., Phelan, C., & Adar, E. (2017). Identifying misaligned inter-group links and communities. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW), 1–23.
- Habib, H., Stoldt, R., Maragh-Lloyd, R., Ekdale, B., & Nithyanand, R. (2024). Uncovering the interaction equation: Quantifying the effect of user interactions on social media homepage recommendations. *arXiv*.
- Howard, M. C. (2024). Investigating the relation of political orientation and vaccination outcomes: Identifying the roles of political ideology, party affiliation, and vaccine hesitancy. *Psychological Reports*, 127(4), 1796–1817.

- Massachs, J., Monti, C., Morales, G. D. F., & Bonchi, F. (2020). Roots of Trumpism: Homophily and social feedback in Donald Trump support on Reddit. In Proceedings of the 12th ACM Conference on Web Science (pp. 49–58).
- Olson, R. S., & Neal, Z. P. (2015). Navigating the massive world of Reddit: Using backbone networks to map user interests in social media. PeerJ Computer Science, 1, e4.
- Prakasam, N., & Huxtable-Thomas, L. (2021). Reddit: Affordances as an enabler for shifting loyalties. Information Systems Frontiers, 23(3), 723–751.
- Valensise, C. M., Cinelli, M., Galeazzi, A., & Quattrociocchi, W. (2019). Drifts and shifts: Characterizing the evolution of users' interests on Reddit. arXiv.
- WHO. (2019). Ten threats to global health in 2019. <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019>
- Yuan, L., Schneider, P. J., & Rizoïu, M. A. (2025). Behavioral homophily in social media via inverse reinforcement learning: A Reddit case study. In Proceedings of the ACM Web Conference 2025 (pp. 576–589).

Session 3A - Algorithms and AI (3)

Day 1 · Mon 13 July · 14:00–15:00 RD Suite, ARC

3A.1 Privacy Concerns and Self-Disclosure Motivation to AI: A Multidimensional Perspective

Kunqiang Wu, Wanting Huang (School of Communication, Hong Kong Baptist University)

Introduction

Conversational AI products based on large language models are expanding application across various domains of our life. They are increasingly considered not merely as functional tools to carry out instrumental tasks but as social agents bearing the capacity to form relationships with users, provide advice, and offer psychological comfort (Kelly et al., 2025; Meng & Dai, 2021). Self-disclosure, the act of sharing personal relevant information, thoughts, and feelings to another (Ashuri & Halperin, 2025; Derlega et al., 1993), is thus becoming central in human-AI interactions to obtain satisfied and personalized responses.

The Computers Are Social Actors (CASA) paradigm indicates that people tend to treat machines as social entities if they possess human-like social features (Nass et al., 1994; Nass & Moon, 2000). In this sense, the highly anthropomorphic attributes of conversational AI agents enable them to foster users' experiences of being together with real and present entities, being connected, and understood (Flavián et al., 2024; Kreijns et al., 2022; Novak & Biocca, 2003). This idea aligns with Social Presence Theory (Short et al., 1976) that highlights the perceived closeness and intimacy of para-social interactions within virtual settings. In other words, AI's social presence alleviates users' self-disclosure burden through constructing a psychologically supportive environment and establishing mutual trust (Konya-Baumbach et al., 2023).

Although existing literature generally indicates the positive association between self-disclosure through social media and users' psychological well-being improvements (Chu et al., 2023), the proliferation of similar practices with AI still raises public concerns towards information privacy. Privacy Calculus Theory suggests that disclosure decisions result from a cost-benefit tradeoff (Cain & Imre, 2021; Lim & Shim, 2022), where individuals compare the perceived benefits and potential risks. It explains the privacy paradox: sometimes people still disclose lots of personal information

despite expressing high-level privacy concerns (Debatin et al., 2009; Hallam & Zanella, 2017; Taddicken, 2014).

Communication Privacy Management theory regards privacy as a dynamic process entailing the ongoing boundary negotiations (Child et al., 2009; Taddicken, 2014), rather than people's general attitudes (Joinson et al., 2010), suggesting take contextualized meanings into consideration when investigating the role of privacy concerns in self-disclosure. However, previous studies largely operationalized privacy concerns as a unidimensional construct, overlooking the complexity of empirical contexts and significant individual or cultural differences in how it is perceived (Lutz & Tamò-Larrieux, 2021). Compared with self-disclosure in social media, AI may not only infer users' personal attributes through continuous interaction but also retain these data for further purposes such as model training (Wu et al., 2024).

Objectives/Research Questions

Addressing the multi-faceted nature of privacy concerns, this study proposes a nuanced theoretical perspective to explore how AI users perceive different dimensions of privacy concerns and how these understandings are associated with their self-disclosure motivations. We adopt the working definition of information privacy concerns (IPC) from Burgoon et al., (1989): individuals' control over when, how and to what extent their information is obtained and used by others, thereby encompassing all aspects of personal data processing. Six dimensions of IPC have been identified in previous studies: control, awareness of privacy practices, collection, errors, unauthorized secondary use, and improper access (Hong & Thong, 2013; Malhotra et al., 2004; Smith et al., 1996).

As a unidimensional construct centered on the potential misuse of personal information, existing research generally suggest a negative relationship between IPC and self-disclosure motivation on SNSs and advertising chatbots (Kim et al, 2024; Krasnova et al., 2009; Wu et al., 2012; Zlatola et al., 2015). Therefore, this study proposes:

H1: AI users' concerns over information control will be negatively associated with self-disclosure motivation.

H2: AI users' awareness of privacy practices will be negatively associated with self-disclosure motivation.

H3: AI users' concerns over information collection will be negatively associated with self-disclosure motivation.

H4: AI users' concerns over information errors will be negatively associated with self-disclosure motivation.

H5: AI users' concerns over unauthorized secondary use of their information will be negatively associated with self-disclosure motivation.

H6: AI users' concerns over improper access to their information will be negatively associated with self-disclosure motivation.

Methods

The data was collected via an online survey, where all individuals above 18-year-old with experience in using AI were invited to participate through online survey platforms, www.wjx.cn or www.qualtrics.com. The survey was initially distributed using snowball sampling within researchers' personal networks, followed by posting survey notices on Rednote (i.e., a popular Chinese social media platform) to enhance the sample's diversity. All participants who completed the survey were offered 5 CNY or 6 HKD as incentives. The data was gathered in three consecutive days from October 8th to 10th, 2025, resulting in 514 original responses. The average filling time of the questionnaires was approximately 7 minutes. Participants were from various occupations or industries, with the top three being students (28.40%), professionals (22.10%), and technicians (20.00%).

All the measures used in this survey were adapted from previous studies, i.e., the 25-item scale from Malhotra et al. (2004) to measure IPC and the 6-item scale from Lim & Shim (2022) to measure self-disclosure motivation with conversational AI agents. To align with our research context, appropriate wording modifications were made to the original scale, for example, "AI agents" or "AI companies" replaced "online companies" in the original questions. A 5-point Likert scale (1 = strongly disagree, 5 = strongly agree) was used to record responses, with higher scores indicating higher level of privacy concerns or self-disclosure motivation. The reliability check showed that all concepts demonstrated acceptable internal consistency, with all Cronbach's $\alpha \geq .68$ (Nunnally, 1978).

A series of cleaning procedures were undertaken towards the raw data ($N = 514$). First, participants under 18 years old ($n = 7$), with completion time length unreasonably short (less than 135 seconds, one third of the average completion time, $n = 12$), and with missing values on key variables ($n = 1$) were excluded. Second, to ensure data quality and address straight-line responses, we screened for and excluded cases with identical responses on more than one sub-scale ($n = 41$). The above steps resulted in a final analytic sample of $N = 453$.

Results

A hierarchical multiple regression analysis was performed to test hypotheses after considering control variables including gender, education level, and AI use frequency. The results indicated that education and AI use frequency in Model 1 had significant relationships with self-disclosure motivation, $R^2 = .16$, $F(7, 445) = 11.70$, $p < .001$. After adding independent variables in Model 2, there was a statistically significant increase in the variance explained, $\Delta R^2 = .13$, $\Delta F(6, 439) = 13.19$, $p < .001$. The full model was also

significant, $R^2 = .28$, $F(13, 439) = 13.42$, $p < .001$. In Model 2, users with a postgraduate degree reported lower self-disclosure motivation than those with a bachelor degree ($\beta = -.29$, $p < .001$). Daily users reported higher self-disclosure motivation than non-daily users ($\beta = .21$, $p < .001$). Regarding privacy concerns, only two of the six dimensions were found significantly associated with self-disclosure motivation. Concerns over unauthorized secondary use was negatively related to self-disclosure motivation ($\beta = -.28$, $p < .001$) while concerns over information errors was positively related to self-disclosure motivation ($\beta = .28$, $p < .001$). Thus, only H5 is supported.

Our results suggest that users were sensitive to the inappropriate deployment of the personal information they shared with AI agents or the companies behind, acting as a cognitive barrier suppressing self-disclosure motivation despite psychological connectedness with AI. Another interesting finding is the positive relationship between concern over informational errors and the self-disclose motivation. Our explanation is that the accuracy of information held by AI entities also matters for users, meaning that they will adjust boundary permeability rules (Kang and Oh, 2023) to allow higher levels of openness to get accurate and quality responses. Here the perceived benefits of reducing uncertainty outweighs the risk of increased disclosure.

In general, self-disclosure behaviors in human-AI interactions may be characterized by a pragmatic negotiation. Though usually hindered by perceived privacy risk of unauthorized information use, individuals will disclose more details to reduce misunderstandings and obtain personalized outcomes. Additionally, individual characteristics such as education level and usage experiences also shape people's understanding and perception of self-disclosure to AI.

Future Work

While we frame information privacy concerns as a key factor, it is possible that the previous AI usage or self-disclosure experience could influence the salience of privacy concerns. The difference in self-disclosure motivation between daily and non-daily users in our sample initially supported this idea. Sensitivity to privacy concerns may also evolve in divergent directions over time (Tsay-Vogel et al., 2018), influenced by factors such as digital literacy (Park, 2013) and engagement level with novel technology (Rogers, 2003). Therefore, we plan to employ a panel longitudinal approach to examine how these relationships would change over time. We also will conduct experimental designs to establish causal directions between user characteristics, information type (e.g., emotional versus factual), primacy motivation to use AI (e.g., instrumental versus entertainment), and self-disclosure practices.

References

- Ashuri, T., & Halperin, R. (2025). Online self-disclosure: An interdisciplinary literature review of 10 years of research. *New Media & Society*, 27(10), 5886–5912.
<https://doi.org/10.1177/14614448241247313>
- Burgoon, J. K., Parrott, R., Le Poire, B. A., Kelley, D. L., Walther, J. B., & Perry, D. (1989). Maintaining and restoring privacy through communication in different types of relationships. *Journal of Social and Personal Relationships*, 6(2), 131–158.
<https://doi.org/10.1177/026540758900600201>
- Cain, J. A., & Imre, I. (2022). Everybody wants some: Collection and control of personal information, privacy concerns, and social media use. *New Media & Society*, 24(12), 2705–2724. <https://doi.org/10.1177/14614448211000327>
- Child, J. T., Pearson, J. C., & Petronio, S. (2009). Blogging, communication, and privacy management: Development of the blogging privacy management measure. *Journal of the American Society for Information Science and Technology*, 60(10), 2079–2094. <https://doi.org/10.1002/asi.21122>
- Chu, T. H., Sun, M., & Crystal Jiang, L. (2023). Self-disclosure in social media and psychological well-being: A meta-analysis. *Journal of Social and Personal Relationships*, 40(2), 576–599. <https://doi.org/10.1177/02654075221119429>
- Debatin, B., Lovejoy, J. P., Horn, A.-K., & Hughes, B. N. (2009). Facebook and online privacy: Attitudes, behaviors, and unintended consequences. *Journal of Computer-Mediated Communication*, 15(1), 83–108.
<https://doi.org/10.1111/j.1083-6101.2009.01494.x>
- Derlega, V. J., & Grzelak, J. (1979). Appropriateness of self-disclosure. In G. Chelune (Ed), *Self-disclosure: Origins, patterns, and implications of openness in interpersonal relationships*. Jossey-Bass.
- Flavián, C., Belk, R. W., Belanche, D., & Casaló, L. V. (2024). Automated social presence in AI: Avoiding consumer psychological tensions to improve service value. *Journal of Business Research*, 175, 114545.
<https://doi.org/10.1016/j.jbusres.2024.114545>
- Hallam, C., & Zanella, G. (2017). Online self-disclosure: The privacy paradox explained as a temporally discounted balance between concerns and rewards. *Computers in Human Behavior*, 68, 217–227. <https://doi.org/10.1016/j.chb.2016.11.033>
- Hong, W., & Thong, J. Y. L. (2013). Internet privacy concerns: An integrated conceptualization and four empirical studies. *MIS Quarterly*, 37(1), 275–298.
<http://www.jstor.org/stable/43825946>

- Joinson, A. N., Reips, U. D., Buchanan, T., & Schofield, C. B. P. (2010). Privacy, trust, and self-disclosure online. *Human-Computer Interaction*, 25(1), 1–24. <https://doi.org/10.1080/07370020903586662>
- Kang, H., & Oh, J. (2023). Communication privacy management for smart speaker use: Integrating the role of privacy self-efficacy and the multidimensional view. *New Media & Society*, 25(5), 1153–1175. <https://doi.org/10.1177/14614448211026611>
- Kelly, S., White, K. M., Kaye, S.-A., & Oviedo-Trespalacios, O. (2025). Tapping into key drivers: Self-disclosure in sensitive health conversations with ChatGPT. *International Journal of Human-Computer Interaction*, 1–11. Advance online publication. <https://doi.org/10.1080/10447318.2025.2499656>
- Kim, W., Ryoo, Y., & Choi, Y. K. (2024). That uncanny valley of mind: When anthropomorphic AI agents disrupt personalized advertising. *International Journal of Advertising*, 1–30. Advance online publication. <https://doi.org/10.1080/02650487.2024.2411669>
- Konya-Baumbach, E., Biller, M., & von Janda, S. (2023). Someone out there? A study on the social presence of anthropomorphized chatbots. *Computers in Human Behavior*, 139, 107513. <https://doi.org/10.1016/j.chb.2022.107513>
- Krasnova, H., Kolesnikova, E., & Guenther, O. (2009). “It won’t happen to me!”: Self-disclosure in online social networks. *AMCIS 2009 Proceedings*, 343. <https://aisel.aisnet.org/amcis2009/343>
- Kreijns, K., Xu, K., & Weidlich, J. (2022). Social presence: Conceptualization and measurement. *Educational Psychology Review*, 34(1), 139–170. <https://doi.org/10.1007/s10648-021-09623-8>
- Lim, S., & Shim, H. (2022). No secrets between the two of us: Privacy concerns over using AI agents. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 16(4). <https://doi.org/10.5817/CP2022-4-3>
- Lutz, C., & Tamò-Larrieux, A. (2021). Do privacy concerns about social robots affect use intentions? Evidence from an experimental vignette study. *Frontiers in Robotics and AI*, 8, 627958. <https://doi.org/10.3389/frobt.2021.627958>
- Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet users’ information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research*, 15(4), 336–355. <https://doi.org/10.1287/isre.1040.0032>
- Meng, J., & Dai, Y. (Nancy). (2021). Emotional support from AI chatbots: Should a supportive partner self-disclose or not? *Journal of Computer-Mediated Communication*, 26(4), 207–222. <https://doi.org/10.1093/jcmc/zmab005>

- Nass, C., Steuer, J., & Tauber, E. R. (1994, April). Computers are social actors. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94, pp. 72–78). Association for Computing Machinery.
<https://doi.org/10.1145/191666.191703>
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103. <https://doi.org/10.1111/0022-4537.00153>
- Nowak, K. L., & Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 12(5), 481–494.
<https://doi.org/10.1162/105474603322761289>
- Nunnally, J. C. (1978). An overview of psychological measurement. In B. B. Wolman (Ed.), *Clinical diagnosis of mental disorders: A handbook* (pp. 97–146). Springer US. https://doi.org/10.1007/978-1-4684-2490-4_4
- Park, Y. J. (2013). Digital literacy and privacy behavior online. *Communication Research*, 40(2), 215–236. <https://doi.org/10.1177/0093650211418338>
- Rogers, E. M. (2004). A prospective and retrospective look at the diffusion model. *Journal of health communication*, 9(S1), 13–19.
<https://doi.org/10.1080/10810730490271449>
- Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. Wiley.
- Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly*, 20(2), 167–196. <https://doi.org/10.2307/249477>
- Taddicken, M. (2014). The 'privacy paradox' in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of Computer-Mediated Communication*, 19(2), 248–273. <https://doi.org/10.1111/jcc4.12052>
- Tsay-Vogel, M., Shanahan, J., & Signorielli, N. (2018). Social media cultivating perceptions of privacy: A 5-year analysis of privacy attitudes and self-disclosure behaviors among Facebook users. *New Media & Society*, 20(1), 141–161.
<https://doi.org/10.1177/1461444816660731>
- Wu, X., Duan, R., & Ni, J. (2024). Unveiling security, privacy, and ethical concerns of ChatGPT. *Journal of Information and Intelligence*, 2(2), 102–115.
<https://doi.org/10.1016/j.jiixd.2023.10.007>

Wu, K. W., Huang, S. Y., Yen, D. C., & Popova, I. (2012). The effect of online privacy policy on consumer privacy concern and trust. *Computers in Human Behavior*, 28(3), 889–897. <https://doi.org/10.1016/j.chb.2011.12.008>

Zlatolas, L. N., Welzer, T., Heričko, M., & Hölbl, M. (2015). Privacy antecedents for SNS self-disclosure: The case of Facebook. *Computers in Human Behavior*, 45, 158–167. <https://doi.org/10.1016/j.chb.2014.12.012>

3A.2 What Else Can We Do? Pessimistic Proactivity and Coerced Domestication in Chinese Gay Men's Negotiations with Algorithms

Longxuan Zhao, Yuyang Cao (Tongji University)

Introduction

This study is situated within the trajectory of critical algorithm studies: algorithms are not neutral technologies; they often carry biases and exclusionary mechanisms embedded in broader sociocultural systems, and thus continually shape the digital lives of marginalized groups through distribution, ranking, and the governance of visibility (Hoffmann, 2019; Haimson et al., 2021). For gay men, such impacts frequently take the form of stricter algorithmic scrutiny, diminished deserved visibility, and even risks to account survival (Thiago et al., 2021; Southerton et al., 2021; Simpson & Semaan, 2021). In the Chinese context, although homosexuality has long been decriminalized and depathologized, official attitudes remain relatively ambiguous and social acceptance is limited (Bao, 2021; Kong, 2010). As a result, platform algorithms are more likely to align with dominant norms and policy expectations, thereby continuously marginalizing gay men through everyday content governance (Li et al., 2025; Wang & Zhou, 2024). At the same time, existing scholarship also suggests that users are not merely passive recipients of algorithmic power; rather, they develop forms of resistance and strategic engagement grounded in their own algorithmic understandings. For instance, some studies highlight how queer creators in China practice creative compliance by deploying symbols, coded vocabulary, and emoji to circumvent censorship while sustaining expression (Simpson et al., 2022; Ai et al., 2023; Shi, 2024).

Objectives/Research Questions

Building on the above scholarship, we identify an underdeveloped research area: existing studies have largely focused on circumvention and expressive tactics on the “posting/publishing” side (e.g., Ai et al., 2023; Shi, 2024), yet there is still a lack of detailed and systematic explanation of how gay men, beyond posting, engage in sustained strategic interactions with algorithms through broader everyday media practices, such as browsing, searching, consuming recommendations, and participating in comment-based interactions. Therefore, we draw on domestication theory (Silverstone et al., 1992), treating algorithms as governable media objects that can be incorporated into everyday life, endowed with meaning, used, and adjusted. This move advances the analysis from publishing strategies to a processual account of everyday domestication.

Under conditions where algorithmic exclusion persists and platform power is structurally asymmetric, we aim to explain how Chinese gay men continue to enact agentic resistance: how they repeatedly probe, learn, and recalibrate their ways of interacting with algorithms in daily use, attempting to domesticate algorithms into usable resources that support the (re)claiming and practice of their sexual identities and intimate lives, while also making sense of their relationship with algorithms and the constraints and costs that this relationship entails. Accordingly, this study centers on the following research question: Under conditions in which algorithms mediate their daily media practices, how do Chinese gay men exercise agency and attempt to domesticate algorithms to support the expression and practice of their sexual identities?

Methods

We employ qualitative interviews and thematic analysis. The data draw on semi-structured interviews with 35 participants, who were recruited through two primary routes. First, recruitment posts were circulated on platforms such as Douyin and Zhihu to invite eligible individuals to volunteer. Second, the sample was expanded via snowballing through our social networks, to capture more diverse experiences of media use and encounters with platform governance. Interviews began with participants' "most recent / recent everyday media practices", asking them to describe concrete situations (e.g., how they scroll through recommendations, whether and how they search, how they read comment sections, and whether they save or share content). We then gradually probed participants' understandings and imaginaries of algorithms, their experiences of visibility shifts and governance interventions, and the actions and self-adjustments they adopted in response.

For analysis, we follow Saldaña's (2021) two-cycle coding procedure to strengthen the systematicity and transparency of interpretation. The first cycle primarily used *in vivo* coding to stay close to participants' own words and everyday concepts (e.g., "safe words"), supplemented by emotion coding and process/action coding to capture participants' affective experiences and action logics when negotiating with algorithms (e.g., anxiety, caution, excitement, numbness, as well as iterative processes of "probing to observing to adjusting"). The second cycle employed focused coding to cluster and refine codes directly related to "domesticating algorithms", generating higher-order themes and mapping the conditions, mechanisms, and consequences across themes.

Regarding ethics and anonymization, participants received information sheet and provided informed consent prior to interviewing. Given that even pseudonyms resembling online handles may enable reverse identification (Gerrard, 2021), the article uses Arabic numerals rather than names to refer to participants, thereby further reducing identifiability. In addition, we reflect methodologically that our non-

heterosexual identity and long-term involvement in queer communities helped build rapport and contextual understanding, but could also shape what becomes salient in questioning and interpretation. Accordingly, the writing remains reflexive and seeks to ground concept development in the empirical materials and the coding process.

Results

The findings can be summarized as a sequential chain of “Objectification and Incorporation, Conversion, and Mutual Domestication”.

First, participants treat algorithms as a manipulable, everyday gatekeeper. By leaving cues for the algorithm at key moments and deliberately redirecting its attention, they attempt to recalibrate recommendation outputs. This effort to “reshape the flow of information” is driven not only by aesthetic fatigue but also by identity-related safety concerns: when one’s device screen might be seen by others, having the feed on homepage “flooded” with recommended content can push private boundaries outward. Some participants therefore reduce searching and liking to lower the risk of unwanted exposure. At the same time, they situate algorithm-mediated browsing within existing physical spaces of privacy—such as bedrooms—so that they can consume recommended content more “at ease,” effectively pulling risk back into a controllable range. Second, to ensure that content can “enter the flow,” participants engage in pre-posting self-censorship by removing explicit keywords and relying on “safe words” and polysemous substitutes (e.g., “rainbow,” “roommate,” “athlete”), as well as repeatedly layering implicit cues, to evade classification and scrutiny. Third, in leveraging the flow, the algorithm also becomes a “secondary verifier”: participants read comment sections and interaction signals to identify the gay subtext of ambiguous content, and on that basis pursue cross-temporal fantasies and affective tracking. Finally, as these individualized tactics circulate through recommendation and search, they contribute to the communal stabilization of coded language and symbols, enabling cultural re-creation and the rebuilding of networks, while also revealing the algorithm’s counter-domestication of users and a shared sense that “algorithms ultimately cannot be fully domesticated”.

Grounded in these empirical data, we further proposes two concepts, “Coerced Domestication” and “Pessimistic Proactivity”, to better capture the meanings and affective dispositions that characterize marginalized groups’ engagements with technology.

Future Work

Future work can proceed along three lines. First, expanding the sample and contexts: the current sample is primarily composed of young, urban participants, which helps capture the fine-grained tactics of high-frequency platform users but may obscure differences across age groups, occupations, and urban-rural divides in terms of risk

perceptions and algorithmic practices. Future research could include a broader range of participants and compare governance logics and spaces for domestication across different platforms (e.g., short-video platforms, Q&A sites, social networking services, and dating apps).

Second, deepening the temporal dimension: this study intentionally “brackets” the dynamic cycle of domestication and therefore does not trace how tactics are re-domesticated or counter-domesticated as platform rules, censorship lexicons, and life stages change. Future work could adopt longitudinal interviews, digital ethnography, or diary/log methods to observe processes such as the updating of “safe words,” fluctuations in visibility, and shifts in community network nodes over time.

Third, extending and testing the explanatory framework: future studies could further specify the conditions under which “Pessimistic Proactivity” turns into exit, silence, or more confrontational tactics, and connect “Coerced Domestication” to platform political economy as well as regimes of labor and affective governance. In addition, incorporating platform-side materials (e.g., appeals, ban records, and review rules) or quantitative indicators of visibility could provide stronger evidence for the experiential claim that “algorithms cannot be fully domesticated,” while also facilitating comparison with other marginalized groups.

References

- Ai, Q., Song, Y., & Zhan, N. (2023). Creative compliance and selective visibility: How Chinese queer uploaders performing identities on the Douyin platform. *Media, Culture & Society*, 45(8), 1686–1695.
<https://doi.org/10.1177/01634437231174345>
- Bao, H. (2021). *Queer media in China*. Routledge.
- Gerrard, Y. (2021). What’s in a (pseudo) name? Ethical conundrums for the principles of anonymisation in social media research. *Qualitative Research*, 21(5), 686–702.
<https://doi.org/10.1177/1468794120922070>
- Haimson, O. L., Delmonaco, D., Nie, P., & Wegner, A. (2021). Disproportionate removals and differing content moderation experiences for conservative, transgender, and black social media users: Marginalization and moderation gray areas. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–35. <https://doi.org/10.1145/3479610>
- Hoffmann, A. L. (2019). Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society*, 22(7), 900–915. <https://doi.org/10.1080/1369118X.2019.1573912>
- Kong, T. S. (2010). *Chinese male homosexualities: Memba, tongzhi and golden boy*. Routledge.

- Li, Y., Wu, J., & Gao, H. (2025). Understanding LGBTQ+ Family Dynamics on Chinese Social Media: A Content Analysis and Sentiment Study on Douyin. *LGBTQ+ Family: An Interdisciplinary Journal*, 1–19.
<https://doi.org/10.1080/27703371.2025.2480586>
- Saldaña, J. (2021). *The coding manual for qualitative researchers*. London: SAGE Publications.
- Shi, J. (2024). Queering algorithms: LGBTQ+ content creators' non-conforming and non-confronting workarounds to digital normativity in China. *Convergence*, 31(3), 728–751. <https://doi.org/10.1177/13548565241299281>
- Silverstone, R., Hirsch, E., & Morley, D. (1992). Information and communication technologies and the moral economy of the household. In R. Silverstone & E. Hirsch (Eds.), *Consuming technologies: Media and information in domestic spaces* (pp. 15–31). London: Routledge.
- Simpson, E., & Semaan, B. (2021). For you, or for “you”? Everyday LGBTQ+ encounters with TikTok. *Proceedings of the ACM on human-computer interaction*, 4(CSCW3), 1–34. <https://doi.org/10.1145/3432951>
- Simpson, E., Hamann, A., & Semaan, B. (2022). How to tame “your” algorithm: LGBTQ+ users' domestication of TikTok. *Proceedings of the ACM on Human-computer Interaction*, 6, 1–27. <https://doi.org/10.1145/3492841>
- Southerton, C., Marshall, D., Aggleton, P., Rasmussen, M. L., & Cover, R. (2021). Restricted modes: Social media, content classification and LGBTQ sexual citizenship. *New Media & Society*, 23(5), 920–938.
<https://doi.org/10.1177/1461444820904362>
- Thiago, D. O., Marcelo, A. D., & Gomes, A. (2021). Fighting hate speech, silencing drag queens? artificial intelligence in content moderation and risks to LGBTQ voices online. *Sexuality & culture*, 25(2), 700–732. <https://doi.org/10.1007/s12119-020-09790-w>
- Wang, S., & Zhou, O. T. (2024). Being recognized in an algorithmic system: Cruel optimism in gay visibility on Douyin and Zhihu. *Sexualities*, 27(7), 1185–1204.
<https://doi.org/10.1177/13634607221106912>

3A.3 “Cutting Ties” with AI? Perceptions and Management of AI Shaming among Young Art Creators in China

Yuyang Cao, Yixiang Zhang, Longxuan Zhao (Tongji University)

Introduction

Alongside the rapid diffusion of AI technologies across industries, norms around “appropriate use” and “how to evaluate AI-involved work” remain in flux and highly fragmented. For example, publishers vary substantially in what they allow regarding generative AI (GAI), where disclosure should appear, and how detailed such disclosure needs to be; this lack of standardization effectively shifts additional compliance burdens onto authors and practitioners (Ganjavi et al., 2024). It is precisely in this transitional period of normative instability that AI shaming has emerged as a visible cultural phenomenon. By AI shaming, we refer to the negative judgments imposed on AI tool users and their outputs, where AI assistance is equated with laziness, deception, or low-quality labor, thereby devaluing the originality, authenticity, and legitimacy of both the AI users and the AI output (Giray, 2024; McNealis, 2025; Sarkar, 2025).

The artistic domain is often regarded as a high-incidence site of AI shaming. On the one hand, artistic communities tend to uphold stronger value narratives around originality, authorship, and human creativity, making AI’s involvement easier to frame as a threat to the human core of art (Colton & Wiggins, 2012); accordingly, audiences frequently develop negative impressions of the aesthetic value and expressive capacity of AI-generated art (Rapp et al., 2025). On the other hand, the reality is not that art rejects AI outright, but rather that controversy and adoption coexist. GAI has markedly lowered barriers for certain creative practices, enabling broader participation and human-AI co-creation, and thereby advancing the democratization of artistic production (Park, 2025). Related studies also show observable diffusion and growth in both tool adoption and output volume within AI art communities (Zhou & Lee, 2024). This raises a more tension-filled question: once creative practice has already begun to change, how do creators interpret the value judgments embedded in AI shaming? Under what contexts do they choose to disclose or conceal AI use, and how do they negotiate their own positionality amid competing demands of efficiency, aesthetics, ethics, and reputational risk?

Objectives/Research Questions

As discussed above, we take the artistic domain as our case. We focus on this field because value narratives around originality, authorship, and human creativity tend to

be particularly strong, and thus more likely to produce stringent evaluative standards toward AI use. Against this backdrop, we ask: in artistic practice, how do AI-using creators perceive, interpret, and develop their evaluative frameworks for AI shaming? Building on our empirical materials, this study pursues two main objectives.

At the experience level, we aim to enrich existing research with grounded accounts of AI shaming and to provide more empirical materials for its conceptualization. Using artistic creation as a high-tension context, where boundaries are more concentrated and more readily contested, we aim to map how AI shaming is perceived, constructed, and circulated through everyday interactions among art creators. Specifically, we examine where art creators experience AI shaming as coming from, in what situations they most acutely recognize the presence of such stigmatization, and what strategies of coping and negotiation they adopt under these pressures.

At the theoretical level, we adopt a more open perspective toward the issues of critique in AI controversies and their social meanings. Prior discussions often treat certain phenomena (e.g., AI hallucinations) as AI traces or negative effects, yet some scholars also remind us that these phenomena may at the same time signal the creative potential of GAI as a posthuman medium for human-AI co-creation (e.g., Ge, 2025). Similar to AI hallucinations, AI shaming carries a critical impulse toward AI; however, we further argue that AI shaming should not be understood merely as a straightforward rejection of technological flaws. Rather, it may be a form of boundary defense and resistant practice through which humans seek to safeguard subjectivity, authorship, and the value of labor amid the rapid permeation of AI. The problem is that such defense is often accompanied by one-size-fits-all stigmatization and an almost obsessive purging of so-called AI traces. As a result, the uncertainties and defect-related costs of AI development are harshly offloaded onto individual creators, while public debate is displaced, from questions of how to define responsibility, labor, and value to a witch-hunt-like adjudication based on symbolic cues, thereby obscuring the more consequential institutional and ethical issues at stake.

Methods

We collected data through semi-structured interviews with young art creators. Each interview lasted 85 to 119 minutes. With participants' consent, interviews were audio-recorded in full, transcribed, and anonymized. Specifically, we combined snowball sampling with purposive sampling and ultimately recruited 15 participants via Xiaohongshu (RED) for in-depth interviews. All participants had several months of experience using AI for artistic creation, with the longest usage history reaching five years. Interviews typically began with participants' initial motivations for adopting AI in their creative practice, and then moved to their processes of human-AI co-creation, their evaluations of AI's involvement in art, and the norms and expectations

surrounding AI use within their industries. Each participant received an interview honorarium of RMB 50 to 100.

For data analysis, we followed Braun and Clarke's (2006, 2019) approach to reflexive thematic analysis. First, each author engaged in iterative readings of the dataset and conducted open coding by drawing on their own knowledge to identify features potentially related to AI shaming. Second, we held online meetings to jointly discuss the initial codes, after which we merged, condensed, and re-labeled them. Finally, guided by a logic of endogeneity and exogeneity, we developed an analytical framework for the situational production of AI shaming, comprising three mechanisms ("quality-driven," "reflexivity-driven," and "coercion-driven"), as well as a typology of creators' practices in response to AI shaming, captured by three themes: "cleaning", "concealment", and "non-disclosure".

Results

Quality-driven AI shaming is endogenous to artistic presentation and technical performance. When AI outputs exhibit recognizable stiffness or awkwardness, creators may feel dissatisfaction or embarrassment because the work falls short of their standards. For instance, I7 expressed displeasure with formulaic and unnatural sentence patterns commonly found in AI-generated text (e.g., templated constructions such as "not X, but Y"), viewing them as undermining the texture and credibility of the work. Similarly, when students pointed out AI-generated qualities in a video clip embedded in I14's teaching slides, I14 described feeling uneasy and embarrassed.

Reflexivity-driven AI shaming is endogenous to art creators' norm-governed selves and is propelled by moral judgments about whether one "should" do this. For example, I12 noted using AI-generated images as references but deliberately avoiding replicating their visual techniques or compositional pathways, both to resist being "pulled" by AI's expressive conventions and to reduce potential copyright or infringement risks. I14 likewise explicitly framed the use of AI generation by illustrators as unethical, suggesting that it amounts to piecing together and appropriating prior artists' work.

Coercion-driven AI shaming, by contrast, is exogenous to the field and its audiences. It is a pressure-laden affect shaped by anticipated external evaluation and punishment. Its core logic is less "I think this is wrong" than "if others find out, there will be consequences". I8 emphasized that they would never let readers know they used AI-generated images as references for character design, fearing widespread backlash or even abuse that could jeopardize their writing career. I2, also an art teacher, similarly worried that if students' parents learned that AI was used in art training classes, it could harm both business prospects and professional reputation.

Importantly, such AI shaming mechanisms do not remain at the level of attitudes; they translate into a set of strategic practices in artistic labor, which we thematize as

“cleaning,” “concealment,” and “non-disclosure”. For instance, I5 described removing the “AI vibe” from dialogue in AI-generated videos to reduce the likelihood that the work would be perceived as AI-made. I2 concealed the fact that images in teaching materials were AI-generated, offering an alternative explanation that they were simply “found online”. In addition, I4 stated that he would not upload or publicly release songs created with AI, treating such outputs instead as part of a private process of learning and experimentation.

Future Work

In our study, we also observed that some art creators did not report experiencing AI-related shaming. Possible reasons include a more tolerant stance toward AI’s still-limited development, or an audience environment that is relatively accepting of AI use. For example, I1 stated that she would directly tell students that the poetry video embedded in their teaching slides was AI-generated, treating it as an emerging medium and a new aesthetic form. For I3 and I6, their supervisors and colleagues did not mind their use of AI at work. I10 likewise indicated that he was not particularly concerned with whether AI was involved in the creative process; as long as the final output was presented well, AI’s participation was not seen as problematic. Future research should continue to examine the contexts and mechanisms under which users do not develop AI-related shame.

In addition, AI shaming is not an either-or attribute. For any given individual, different topics and contexts may generate different perceptions and intensities of AI shaming. For example, although I8 refused to commission creators who had used AI tools in their past work (suggesting the presence of AI shaming), in her own writing practice she sometimes deliberately inserted words commonly associated with AI-generated text to make it difficult for readers to tell whether a passage was written by a human or by AI (suggesting the absence of such shame in that context). Future research could build on this observation to explore more dynamic and nuanced trajectories of AI shaming across contexts and over time.

This study also has limitations. All of our participants were AI-using art creators, which may partly account for the complex and ambivalent configurations of AI shaming we document. We did not include art creators who never use AI; these individuals may hold more unequivocal or strongly negative positions toward AI use and thus exhibit more consistent forms of AI shaming. While our findings illuminate the complexity and tension in how AI shaming is produced across different contexts, future research on AI shaming would benefit from incorporating creators who have never used AI in artistic production.

References

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health*, 11(4), 589-597.
- Colton, S., & Wiggins, G. A. (2012). Computational creativity: The final frontier?. In *ECAI 2012* (pp. 21-26). IOS Press.
- Ge, L. (2025). Spectral imaginings and sympoietic creativity: AI hallucinations and the ethics of posthuman creativity. *Big Data & Society*, 12(4), 20539517251400696.
- Giray, L. (2024). AI shaming: the silent stigma among academic writers and researchers. *Annals of Biomedical Engineering*, 52(9), 2319-2324.
- McNealis, R. (2025). Shame in the machine: affective accountability and the ethics of AI. *AI & SOCIETY*, 1-11.
- Park, S. (2025). The work of art in the age of generative AI: aura, liberation, and democratization. *AI & society*, 40(3), 1807-1816.
- Rapp, A., Di Lodovico, C., Torrielli, F., & Di Caro, L. (2025). How do people experience the images created by generative artificial intelligence? An exploration of people's perceptions, appraisals, and emotions related to a Gen-AI text-to-image model and its creations. *International Journal of Human-Computer Studies*, 193, 103375.
- Sarkar, A. (2025, April). AI Could Have Written This: Birth of a Classist Slur in Knowledge Work. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (pp. 1-12).
- Zhou, E., & Lee, D. (2024). Generative artificial intelligence, human creativity, and art. *PNAS nexus*, 3(3), pgae052.

Session 3B - Digital Intimacies

Day 1 · Mon 13 July · 14:00–15:00 Seminar Suite A, ARC

3B.1 “Amazing body. Does the face match:” A Qualitative Study of how Queer Embodied Masculinities on Dating Apps Practice, Communicate and Build Trust with Anonymity

John Hodson (University of Salford)

Introduction

Queer embodied masculinities (QEM) – a catch-all term used to describe non-cisgender and/or non-heterosexual people who embody masculinity in some way – have a long history with digitally mediated dating. This history is closely entwined with anonymity (Campbell, 2014; Slater, 1998), creating complex negotiations of if, when, and how images of the self are shared. Since the introduction of smartphones, this has predominantly been found on dating apps. In this paper, anonymity is defined as obscuring the user’s face in their publicly viewable profile. This research relies on a qualitative multi-method approach to understand in depth how anonymity is experienced on both sides of the spectrum: both by those who are anonymous, and those who are identifiable.

With a long history of anonymous intimacies, this vital tradition is at a juncture in the digital world where privacy is traded for authenticity and profit, and where digital presences are expected to exist accurately and with clarity (van Dijck, 2013). For some, this is a valued trade, however for others, this becomes a source of anxiety, and increases the labour of managing identity and privacy. The existing dichotomy of anonymous users and identifiable users communicating with themselves and each other raises valuable discussions of how privacy, safety and trust are practiced – especially in times of heightened emotional states.

Objectives/Research Questions

This research holds 1 aim, and 3 objectives to guide the data gathering process. The research aim is: ‘To discover how QEM on dating apps in the United Kingdom position anonymity.’ Critically, this is inclusive of both anonymous and identifiable QEM,

privileging neither identity. The three objectives that follow are constructed to achieve this aim:

1. To gauge QEM expectations and tensions of anonymous users on dating apps.
2. To deeply analyse dating app profiles to understand how the profile in its many forms facilitates the communication of the user's identity.
3. To assess how users produce and impose self-defined boundaries in order to build trust, and maintain safety and privacy.

The aim and objectives have been generated as a way of forcing anonymity into the discussion of QEM on dating apps, as anonymity in interview-based research has often been referenced in relation to other highly valuable aims and objectives in other research projects, whilst not becoming the centre of discussions surrounding dating apps (Albury et al., 2019; Blackwell et al., 2015; Dhoest & Szulc, 2016).

Methods

Three qualitative research methods, forming an interpretivist multi-method approach were used as part of a wider discussion regarding data privacy, dating app architecture and anonymity. Firstly, the walkthrough method (Light et al., 2018) was used on ten sampled dating apps: Grindr; Growlr; Scruff; Tinder; 3way; Taimi; Hornet; Onlylads; Happn; Jack'd. This included reading and analysing the Terms and Conditions, Privacy Policy, Cookie Policy, Community Guidelines, their web pages, and the use of the app itself to understand how each handled data, and encourage users to practice identity disclosure. Written text, as well as screenshots and recordings of the reading and viewing of each app, were analysed and coded.

A netnography (Kozinets, 2015) was conducted, using an immersion journal approach with screen and a manual tracking of some relevant statistical data. Specifically, the number of anonymous profiles per each app and location, done by manually viewing every profile image of the geographically closest profiles, and viewing all publicly visible profile images the user had uploaded. Three locations in the United Kingdom were used: Norfolk, Surrey and Greater Manchester. Netnography data used general observations to find broad themes amongst both anonymous and identifiable profiles. The types of general observations were what users wrote in their profiles, the extent to which statistical data such as users 'height,' 'weight,' 'sexual position' and 'tribes' were filled out, and how and what profile images were used. This data was coded and themed manually, as much of the data was visual such as screenshots.

Finally, 14 semi-structured interviews were hosted online. Dating apps were the predominant method of participant enrolment, using a deliberately created recruitment account on each sampled dating app, where users were encouraged to enquire about the ongoing study. Further recruitment through mutual friends and connections were also used. Though snowball sampling and external advertising was used, these

methods garnered no additional participants. Interviews ranged between 50 minutes and 174 minutes, with 7 anonymous QEM, and 7 identifiable. Though most questions were asked to all participants, such as ‘Can you describe the public profile image that you use?’ and ‘what drew you to choosing them’, questions were also generated from participants personal profiles. This often revolved around why some statistical data was included or omitted, or why they had used certain phrases that was relevant to the research aim and objectives, such as why they had used the phrase ‘no pic, no chat.’ Data was coded in NVIVO, with three repetitions. Codes were then themed.

This paper focuses primarily on the interview data. Walkthrough and netnography data is occasionally referenced when it strongly relates to the interview data discussed. Ethical approval was granted by the University of Salford under code 0055.

Results

3 themes and 4 subthemes were found within the 14 interview transcripts: external governance of privacy and anonymity; reasoning, feelings and experiences of personal anonymity (subthemes; hypothesising anonymity; user legitimacy); the management of privacy throughout negotiations and communications (subthemes; the public facing identity; the considerations of taking connections offline).

Each theme is briefly introduced to provide context for the research, with feelings and experiences of personal anonymity being the focus of this paper. The theme that explores how both identifiable and anonymous users approach anonymity from differing ends of the spectrum. It discusses what participants felt that anonymous users want from their dating app usage, whilst they also hypothesise and even state the exact reason that they are anonymous. This also reflects on the expectations that users hold about anonymous users. It found a large variety of hypothetical and actual reasons that users were anonymous, that demonstrates a collective sympathy for users in the subtheme ‘hypothesising anonymity’, whilst equally expressing frustration that anonymity is a common practice. It ultimately led to an accepted discomfort whereby anonymity was tolerated, so long as other criteria were met. Of particular note, there was evidence of ‘elitism’ amongst anonymous users, who expressed that whilst they acknowledged that they contributed to issues of trust and aggravation, the labour invested into their images and profiles exceeded other anonymous users.

The second subtheme ‘user legitimacy is discussed in regards to fake profiles, bots, and in-app verification systems – such as badges or checkmarks as forms of validation. This subtheme emerged as anonymity was often associated with fake profiles and bots, conflating genuine anonymous and discrete users as scammers, AI bots or high-risk encounters. Users reported using “gut instincts” to tell if a profile they saw was owned by a real user, or a ‘catfish’. Some users went so far as to find patterns in fake profiles, with detailed descriptions such as the order of profile images, the types of images

used, the clarity and quality of pictures and the responsiveness of users. These perspectives were seen across anonymous and identifiable users on dating apps, and there was a consistent awareness across all participants that illegitimate users and accounts exist on dating apps. Though in-app verification systems were appreciated, users felt there could still be ways to fake them, and preferred to use their own judgements of other users to develop trust. Interestingly, participants consistently referred to bots as fake profiles (whereby fake profiles relied on overt deception, bots did not); yet this was found to be inconsequential in terms of protecting themselves, as participants regardless of distinction did not trust these profiles.

Future Work

Future work from this research should continue to use different methods, as well as further explore different national backgrounds. Critically, using methods such as qualitative surveys may drastically increase the recruitment of anonymous QEM. This study experienced a large number of users contacting the recruitment profile expecting a survey link, in which their interest diminished once realising the study was primarily interview-based. Using surveys to build trust amongst potential participants, and inviting them for interviews at the end of the survey, using small financial incentives would likely increase the recruitment of anonymous QEM on dating apps to further excavate the personal experiences of them. Though surveys been used previously (Miller, 2020), there could be further qualitative discussions to be developed further in regards to feelings and experiences. With appropriately focused questions that probe longer responses, those who are still uncomfortable to come forward for interviews on such sensitive topics, survey responses may uncover further details on the extent to which anonymity is experienced by anonymous users.

The same methodology may also be used to further research transgender and 'chaser' identities, as only 2 of the 14 participants identified as transgender at the time of the interviews. This is inclusive of transgender women specifically, of whom were not a part of the target sample of the research project. Based on discussions with participants more widely, more intensive sampling of anonymous QEM in high South-East population areas may also prove to highlight highly charged tensions between users' religion and sexuality, which was not able to be fully explored in this research project.

References

- Albury, K., Byron, P., McCosker, A., Pym, T., Walshe, J., Race, K., Salon, D., Wark, T., Botfield, J., Reeders, D., & Dietzel, C. (2019). Safety, risk and wellbeing on dating apps: final report (LP160101687). S. U. o. Technology.
<https://apo.org.au/node/268156>

- Blackwell, C., Birnholtz, J., & Abbott, C. (2015). Seeing and being seen: Co-situation and impression formation using Grindr, a location-aware gay dating app. *New Media & Society*, 17(7), 1117-1136. <https://doi.org/10.1177/1461444814521595>
- Campbell, J. E. (2014). *Getting It On Online : Cyberspace, Gay Male Sexuality, and Embodied Identity*. [Online Non-fiction Electronic document]. Taylor and Francis. <https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=cat07845a&AN=uea.876512899&authtype=sso&custid=s8993828&site=eds-live&scope=site>
- Dhoest, A., & Szulc, L. (2016). Navigating Online Selves: Social, Cultural, and Material Contexts of Social Media Use by Diasporic Gay Men. *Social Media + Society*, 2(4). <https://doi.org/10.1177/2056305116672485>
- Kozinets, R. V. (2015). *Netnography : redefined* (Second edition. ed.). SAGE.
- Light, B., Burgess, J., & Duguay, S. (2018). The walkthrough method: An approach to the study of apps [Article]. *New Media and Society*, 20(3), 881-900. <https://doi.org/10.1177/1461444816675438>
- Miller, B. (2020). A Picture is Worth 1000 Messages: Investigating Face and Body Photos on Mobile Dating Apps for Men Who Have Sex with Men. *Journal of Homosexuality*, 67(13), 1798-1822. <https://doi.org/10.1080/00918369.2019.1610630>
- Slater, D. (1998). Trading Sexpics on IRC: Embodiment and Authenticity on the Internet. *Body & Society*, 4(4), 91-117. <https://doi.org/10.1177/1357034x98004004005>
- van Dijck, J. (2013). 'You have one identity': Performing the self on Facebook and LinkedIn [Article]. *Media, Culture and Society*, 35(2), 199-215. <https://doi.org/10.1177/0163443712468605>

3B.2 Narrative Singularity and Collective Resonance in TikTok Sexual Violence Disclosures and Networked Publics

Talia Fiester (University of Pennsylvania)

Introduction

A legacy of the viral 2017 #MeToo movement, social media remains a conducive environment for first-person narrativization of sexual violence by ordinary users. These networked disclosures constitute a distinct digital genre due to their vulnerability, visibility, and culturally contested status. Narrators experience exposure to risk and questions of credibility, and their disclosures are constrained by racialized and gendered hegemonic scripts, patriarchal logics of believability, and platform affordances and governance. A well-conceptualized approach to studying these disclosures, thus, requires them to be analyzed as both personal narratives and mediated texts.

Mendes, Keller, and Ringrose (2018) theorized “digitized narratives of sexual violence,” as narratives “shaped by distinctive... conventions, affordances, and restrictions of the platforms in which they appear” (p. 1290). While a networked environment allows for more immediate and accessible relational support and community care (O'Neill, 2024), they also expose creators to pervasive misogyny, public scrutiny, and algorithmic invisibility. In order to endure the “public stage” on which digitized narratives unfold (Hewa, 2021), survivors are tasked with the “affective labor” of calibrating authenticity, vulnerability, and credibility (Hochschild, 2012).

On TikTok, the visibility of digital testimonies of sexual violence is structurally mediated by relationality, with platform affordances encouraging “participatory mimesis” through comment sections, share functions, and video “stitches” (Zulli & Zulli, 2022). These interactions form “networked publics” and produce “public intimacy,” linking relational dynamics to algorithmic visibility outcomes (boyd, 2010; Baym, 2015; Kaplan, 2021). Moreover, algorithmic design infused with “data discrimination” (Noble, 2018) and empowered by racist believability logics has substantial implications for visibility and engagement (Benjamin, 2023; Bucher, 2018).

This study conceptualizes TikTok disclosures of sexual violence through the application of two theories. Reckwitz’s (2020) “logic of singularity” holds that narratives must successfully affect others to secure engagement and visibility. Theories of resonance describe the mechanism in which digital narratives of sexual violence become relational cultural texts through the affecting processes of empathy, identification, and

recognition (Rosa, 2019; McDonnell et al., 2017; Mühlhoff, 2019).

The implications of singularity and resonance on digitized narratives produces a paradox that animates this paper: stories must appear singular to be visible online; however, resonance absorbs narratives into collective forms and diffuses the individuality that enabled its exposure. I refer to this as the “paradox of collective resonance,” which highlights how TikTok’s platform design both amplifies and redistributes narrative authority.

Objectives/Research Questions

Scholars have yet to theorize how boundaries around digital authorship of sexual violence narratives become rhetorically contested. While prior work explores visibility logics and affective engagement, it does not currently conceptualize the required singularity or facilitation of collective resonance. This project addresses that gap by introducing the paradox of collective resonance, where narratives achieve visibility through singularity but are simultaneously absorbed into collective forms. By examining the relational-affective dynamics, this paper theorizes singularity, collective resonance, and distributed authorship as key mechanisms in the circulation of “digitized narratives” (Mendes, Keller, & Ringrose, 2018).

This paper asks: How do TikTok users negotiate control over their narratives while audiences co-create meaning in comment sections? How does the platform’s design influence both visibility and collective authorship of sexual violence disclosures? Analysis proceeds in three sections: creators’ mediation of singular narrative authority; comment-level “networked recognition” that destabilizes narrative authorship; and the synthesis of these findings to theorize the paradox of collective resonance.

Methods

The methodological approach foregrounds intentional selection and interpretive transparency; data serve as illustrative rather than representative evidence. The corpus consists of five TikTok videos with a focus on women’s sexual violence disclosures. TikTok was selected for its algorithmic prioritization of relatable content (Zulli & Zulli, 2020; Abidin, 2021) and its robust comment culture (Nainggolan & Agustini, 2024; Alafwan et al., 2023; Slater, 2022). These affordances make TikTok a key site for studying relational-affective engagement, networked authorship, and visibility dynamics of disclosures. The videos are selected from the “digitized narrative” genre (Mendes, Keller, & Ringrose, 2018) that present truth-telling, vulnerability, and authorship via self-disclosure. Purposive sampling identified videos narrative ownership. Search terms included “my sa story,” with “sa” as “algospeak” for sexual assault (Lorenz, 2022). Videos included autobiographical disclosure with substantial comment section engagement to analyze relational resonance.

Videos were analyzed using qualitative content analysis informed by affect theory and critical discourse analysis. Initial coding involved transcribing and reviewing 150-200 top comments to identify emerging engagement patterns. Coding proceeded in two phases: descriptive coding of textual and interactional features, followed by theoretical coding to interpret recurring patterns and their role in resonance, singularity, and authorship.

Analysis used Fairclough's (1992, 1995) three-dimensional CDA model, mapped onto creator, comment, and structural levels. Four descriptive coding categories emerged: forms of resonance, affective tone, ownership markers, and modes of address. Theoretical codes included networked recognition, pedagogy through recognition, discursive ownership, relational empathic positioning, rewarding vulnerability, and identification resistance.

Results

Analysis explores how sexual violence disclosures circulate online, elicit affective recognition, and become sites of collective meaning-making. These dynamics form the paradox of collective resonance, a structural condition of visibility: narratives attract engagement through their singular capacity to affect others, while destabilizing boundaries of narrative ownership. On TikTok, platform affordances amplify these dynamics by privileging relational engagement and algorithmic visibility.

Creators encounter this paradox from the outset. They must present their experience as singular to secure algorithmic visibility and anticipate reception by "imagined audiences" (Marwick & boyd, 2011). Hegemonic logics of believability constrain this performance, and singularity is demonstrated through signifiers of authenticity and veracity. Creators establish authorship through story-claiming language, invitations for recognition, and modes of address—invitational, declarative, and collective voice—to structure authority in anticipation of engagement.

Commenters engage in what I term "networked recognition," producing co-constructed engagement within the comments of an individual disclosure that expands it into a communal text. Four forms emerge: direct identification, asserting shared experience and subtly claiming narrative authority; partial recognition, expressing empathy/sympathy while maintaining the creator's narrative centrality; independent comment disclosures, replicating the structure and emotional register of the original post; and universalizing recognition, invoking collective voice to recode individual narratives as shared cultural texts. Commenters also negotiate ownership boundaries through identification resistance and pedagogical interactions, reflecting on the disclosure without claiming co-authorship.

Creators and commenters thus co-produce meaning and ownership. Videos and comment sections function as collectively authored cultural texts, in which resonance

produces engagement, validation, and the redistribution of narrative authority. These practices demonstrate how TikTok's platform design mediates collective engagement and structural authorship. Micro- and mezzo-level practices—including ownership markers, invitations for recognition, modes of address, and forms of networked recognition—accumulate to produce structural outcomes: comment sections become iterative sites of meaning-making, narratives circulate as networked cultural texts, and authorship is redistributed at the cost of experiential singularity.

TikTok disclosures are fundamentally relational and affective. Singularity, resonance, and networked recognition operate in tandem to inform how narratives are perceived, circulated, and co-authored. As resonance expands beyond the bounds of the original post, digital sexual violence disclosures become culturally networked compositions implicated within the paradox of collective resonance.

Future Work

Historical #MeToo parallels show that aggregation of case-by-case testimonies disrupts the status quo of isolation and silence, but it also risks veiling distinct sociocultural realities. Within this corpus, multiple narratives and positionings coexist in each disclosure: commenters assert narrative equivalence, authorize affective legibility, or contribute additional testimonies. Collective resonance empowers participants, but the compression of diverse experiences threatens to erase significant narrative differences on the basis of identity and marginalization. Social media platforms can simultaneously amplify and obscure individual narratives in accordance with hegemonic structures which raises important questions for platform design, moderation, and digital culture research. Further work should explore the potential of problematic consequences for structural authorship, in which “my story” is displaced by “our experience.”

References

- Abidin, C. (2021). From “Networked Publics” to “Refracted Publics”: A Companion Framework for Researching “Below the Radar” Studies. *Social Media + Society*, 7(1), 2056305120984458. <https://doi.org/10.1177/2056305120984458>
- Alafwan, B., Siallagan, M., & Putro, U. S. (2023). Comments Analysis on Social Media: A Review. *EAI Endorsed Transactions on Scalable Information Systems*, 10(6), 1. <https://doi.org/10.4108/eetsis.3843>
- Baym, N. K. (2015). *Personal connections in the digital age*. John Wiley & Sons.
- Benjamin, R. (2023). *Race After Technology*. Routledge.
- Boyd, D. (2010). Social network sites as networked publics: Affordances, dynamics, and implications. In *A networked self* (pp. 47-66). Routledge.

- Bucher, T. (2018). *If... then: Algorithmic power and politics*. Oxford University Press.
- Fairclough, N. (1992). Discourse and Text: Linguistic and Intertextual Analysis within Discourse Analysis. *Discourse & Society*, 3(2), 193–217.
<https://doi.org/10.1177/0957926592003002004>
- Fairclough, N. (1995). *Critical discourse analysis The critical study of language*. London and New York Longman. - References—Scientific Research Publishing. (n.d.). Retrieved December 13, 2025, from
<https://www.scirp.org/reference/referencespapers?referenceid=821968>
- Hewa, N. (2022). The mouth of the internet, the eyes of the public: Sexual violence survivorship in an economy of visibility. *Feminist Media Studies*, 22(8), 1990–2001. <https://doi.org/10.1080/14680777.2021.1922483>
- Hochschild, A. R. (2003). *The Managed Heart: Commercialization of Human Feeling, Twentieth Anniversary Edition, With a New Afterword (2nd ed.)*. University of California Press. <https://www.jstor.org/stable/10.1525/j.ctt1pp0cf>
- Kaplan, D. (2021). Public intimacy in social media: The mass audience as a third party. *Media, Culture & Society*, 43(4), 595–612.
<https://doi.org/10.1177/0163443721991087>
- Lorenz, T. (2022, April 8). Internet ‘algospeak’ is changing our language in real time, from ‘nip nops’ to ‘le dollar bean.’ *The Washington Post*.
<https://www.washingtonpost.com/technology/2022/04/08/algospeak-tiktok-le-dollar-bean/>
- Marwick, A. E., & boyd, danah. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, 13(1), 114–133. <https://doi.org/10.1177/1461444810365313>
- McDonnell, T. E., Bail, C. A., & Tavor, I. (2017). A Theory of Resonance. *Sociological Theory*, 35(1), 1–14. <https://doi.org/10.1177/0735275117692837>
- Mendes, K., Keller, J., & Ringrose, J. (2019a). Digitized narratives of sexual violence: Making sexual violence felt and known through digital disclosures. *New Media & Society*, 21(6), 1290–1310. <https://doi.org/10.1177/1461444818820069>
- Mühlhoff, R. (2019). Affective resonance. In *Affective Societies*. Routledge.
- Nainggolan, V., & Agustini, M. Y. D. H. (2024). From Comments To Cart: The Role Of Comments On Purchase Decision Process On Tiktok. *International Seminar*, 6, 963–975.

- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press.
<https://doi.org/10.18574/nyu/9781479833641.001.0001>
- O'Neill, T. (2024). The Politics of 'Speaking In': Forming Online Survivor Communities. In *Disclosing Sexual Violence in a Digital Society: Storytelling, Activism, and Justice* (pp. 59-80). Cham: Springer Nature Switzerland.
- Reckwitz, A. (2020). The society of singularities. *Futures of the study of culture: Interdisciplinary perspectives, global challenges*, 8, 141.
- Rosa, H. (2019). *Resonance: A sociology of our relationship to the world*. John Wiley & Sons.
- Slater, K. (2022). Theorizing Cultures of Oversharing on Tiktok. In *LGBTQ Digital Cultures*. Routledge.
- Zulli, D., & Zulli, D. J. (2022). Extending the Internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform. *New Media & Society*, 24(8), 1872–1890. <https://doi.org/10.1177/1461444820983603>

3B.3 Intimacy capital, recognition, and ordinary creators on TikTok

Irida Ntalla (University of the Arts London)

Introduction

The paper aims to contribute to the scholarship on creator economies by accounting the production and exchange of meaning making, symbolic values, and social recognition. Focusing on TikTok, a key site within the social platform ecosystem with competing systems of value, the study moves beyond the dominant lenses of creator visibility tied to monetisation and promotion. It proposes that for ordinary creators, visibility is linked to struggles for voice and recognition, being heard and seen, which remains a fundamental human and creative motivation (Sujon & Ntalla, 2025). The study therefore explores how ordinary creators, as part of intimate publics navigate the platform's logics of community building and algorithmic virality through intimate narratives and self-telling.

On TikTok, creators share intimate moments with the platform's algorithms to form networked and digital intimate publics. These publics are built on common interests, and a shared aspiration for connection creating worlds through narratives of shared experiences (Berlant, 2008: 281). This process blurs publics that blur private life and public performance, with platforms facilitating an infrastructure of intimacy that supports digital kinships and affairs of the heart. Dobson et al. (2018) extend Berlant's work by developing the idea of digital intimacies and argue that social media facilitate new intimate imaginaries, forming spaces for connection, solidarity, and "less bounded forms of intimate care and pleasure" (p. 7). On TikTok, this includes creators and users forming intimate relationships with the platform itself via its algorithm (Bhandari & Bimo, 2022). Intimacy involves a search for recognition that is provided by other humans (Berlant, 2008), which in digital culture also comes from platforms, affordances and infrastructures that shape human subjectivity (Paasonen et al., 2023).

The paper considers that the ordinary creator on TikTok must be understood as an agent navigating a complex field where the pursuit of recognition key for identity-making, and belonging is woven fabric of digital communication (Wells & Friedland, 2023), both a meaningful social practice and a source of economic value within the platform. Visibility has taken a social value making attentional logics in a media environment a central principle and close partner of recognition (Davies, 2021). Intimate publics and intimate narrative sharing can foster counter-publics of care and solidarity, embedded within TikTok's sides and silos, and platform capitalist logics. The

accumulation of intimacy capital could be seen as key in this process converting personal narratives into a form of value that is symbolic and economic.

Objectives/Research Questions

The study is guided by the core research questions: How do ordinary TikTok creators acquire intimacy capital? What practices and activities related facilitate this as indicated in their public content? And, does achieving visibility in the platform require more accumulation of intimacy capital for certain creators? The objective is to analyse approaches of intimacy capital accumulation, particularly asking about disparities in acquiring for marginalised groups of creators, contributing to scholarship in the field of digital intimacies and understandings of the symbolic and affective economies at play where the human need for recognition is articulated, circulated and monetised.

Methods

Methodologically, the study employs an exploratory digital ethnography on TikTok to analyse self-narratives and creative practices of diverse ordinary creators. The analysis considers TikTok videos as audiovisual media artefacts that carry multidimensional and complex messages, key to the creators' practices of sharing within intimate publics through interaction and reaction. The content analysis targets the language, format and interaction which creators express their communicative goal and how they construct the video's audiovisual message (Vázquez-Herrero et al., 2023: 52). This involves a close reading of video content, comments, duets, and hashtag networks to trace how intimate narratives are constructed, validated, and circulated, and ultimately contribute to accumulation of intimacy capital. The approach is situated within the platform ecology of virality (Lee & Umbach, 2025), examining how the selected videos align with algorithmic affordances, trending sounds, and meme formats to achieve visibility and recognition.

The study purposely moves beyond a limited case-study approach to examine a broader spectrum of intimate self-telling through the mapping of thirty selected different videos that present original content of personal intimate narratives. This includes content aligning with existing work on single mother creators, who share narratives of vulnerability, 'suffering,' and 'experiential expertise' to forge community and entrepreneurial possibility, such as single mother sharing journey in temporary accommodation and a trans single mother. Beyond that the sample includes original videos from diverse ordinary TikTok creators posting intimate content about love, breakups and mental health journeys, couple and individuals documenting everyday life during hard and impossible circumstances such war. There are limitations and ethical considerations to consider in the methodological approach such as the focus on the public performances of intimacies, and the ethical research approaches to engaging with public data.

Results

The analysis of the empirical data for this study is currently in development. However, the findings will be conceptually framed and supported by scholarship, which indicates that intimacy has become a key content creation strategy. Existing research within social media indicates that it is fostered through performances of authenticity, vulnerability and resilience. It is built through shared trust from perceived authentic self-presentation, personal disclosures, shared identities and sometimes oversharing (see Andreassen, 2017; Kennedy, 2018). This is evidenced in the practices of creators like single mothers, whose digital narratives of ‘imperfect–resilient’ motherhood reframe struggle into aspirational stories of entrepreneurial independence, converting lived experience into educational and social capital that shape their mediated visibilities (Ntalla, 2025). Intimacy becomes central to the platform’s political economy, through performances of authenticity that turn social capital into economic capital (Heřmanová, 2024) with these ‘communicative intimacies’ not contradicting the commercial nature of the relationships (Abidin, 2016). These practices can be therefore understood as the accumulation of intimacy capital, a form of cultural behaviour that converts into social and economic benefit. Lambert (2016) understands this form of capital as requiring the cultural and socio-technical competencies to perform intimacy within platform infrastructures, competencies embedded within a classed, raced, and gendered emotional habitus (Dobson, 2018). Bourdieu’s notion of habitus relates as a “feel for the game” that informs strategies, and activities and practices developed and oriented towards accumulating capital in the field. Being part of a field, a ‘social space’ (Bourdieu, 1984) and striving for a capital is not solely about economic gain, but also pursuing what we care about, giving us meaning and defining our beings.

Future Work

Future work could extend the digital ethnography to interviews with creators to understand their lived experiences, strategies, and the practices of accumulating and managing intimacy capital. Further, the research could develop a more intersectional lens in examining the diverse creators, as well as consider other platforms and how strategies may vary with the broader digital ecosystem.

References

- Abidin, C. (2016). Communicative ♥ Intimacies: Influencers and Perceived Interconnectedness. *Ada: A Journal of Gender, New Media & Technology* 8.
- Andreassen R., Petersen M. N., Harrison K., Raun T. (Eds.). (2018). *Mediated intimacies: Connectivities, relationalities and proximities*. Routledge.
- Berlant, L. (2008). *The female complaint: The unfinished business of sentimentality in American culture*. Duke University Press.

- Bhandari, A. & Bimo, S. (2022). Why's Everyone on TikTok Now? The Algorithmized Self and the Future of Self-Making on Social Media. *Social Media + Society* 8(1) <https://doi.org/10.1177/20563051221086241>.
- Bourdieu, P. (1984). *Distinction*. Routledge.
- Davies, W. (2021). The politics of recognition in the age of social media. *New Left Review*, (128), 83–99.
- Dobson, S., Brady R. A., & Carah, N. (2018). *Digital Intimate Publics and Social Media*. Cham: Palgrave Macmillan.
- Heřmanová, M. (2024). Hands Labouring For Safety: Mediated Intimacy in Influencer Communities on Instagram. *Anthropology of Consciousness* 35(2): 186–200. <https://doi.org/10.1111/anoc.12238>.
- Kennedy, J. (2018). Oversharing is the norm. In A. S. Dobson, B. Robards, & N. Carah (Eds.), *Digital intimate publics and social media* (pp. 265–280). Palgrave Macmillan.
- Lambert, A. (2015). Intimacy and social capital on Facebook: Beyond the psychological perspective. *New Media & Society*, 18(11), 2559-2575.
- Lee, J. & Umback, J. (2025). From Visibility to Virality: Unpacking the Virality-Centered Ecology and the Evolving Creator Economy on TikTok. Continuum.
- Ntalla, I. (2025). Single mothers as ‘imperfect-resilient’ subjects and their entrepreneurial narratives on TikTok. Continuum, 1–20. <https://doi.org/10.1080/10304312.2025.2545526>
- Paasonen, S., Jaaksi, V., Koivunen, A., Nikunen, K., Talvitie-Lamberg, K. & Vänskä A. (2023). Intimate infrastructures we depend upon: Living with data. *Media Theory*, 7(2), 285–308.
- Sujon, Z., & Ntalla, I. (2025). Special issue introduction: reflecting on TikTok creators and digital economies. Continuum, 1–19. <https://doi.org/10.1080/10304312.2025.2549003>
- Vázquez-Herrero, J., Negreira-Rey, M., & Klug, D. (2023). Key Concepts and Research Methods to Analyze TikTok Journalism. *Cultural Science*, 15 (1) 42-62. <https://doi.org/10.2478/cs-j-2023-0005>
- Wells, C., & Friedland, L. A. (2023). Recognition Crisis: Coming to Terms with Identity, Attention and Political Communication in the Twenty-First Century. *Political*

Communication, 40(6), 681–699.

<https://doi.org/10.1080/10584609.2023.2222267>

Session 3C - Social Media Cultures and Everyday Life (3)

Day 1 · Mon 13 July · 14:00–15:00 Seminar Suite C, ARC

3C.1 Status in the Palm of the Hand: Class-Based Patterns of Digital Leisure

Kata Számel (ELTE Centre for Social Sciences)

Introduction

The utilisation of leisure time has long been a subject of interest to social scientists, with numerous studies demonstrating its correlation with social class. Although the influence of digitalisation is evident not only in professional spheres but also in private life, research in this field has been limited in terms of its focus on social stratification within the digital space and the utilisation of leisure time.

Digital technologies have permeated all aspects of everyday life. However, the digital space reflects the inequalities of social stratification that exist in the non-digital world. While analyzing digital inequalities, researchers initially centred their efforts on the gap between those with and without physical access to the internet – the first level of the digital divide – but with the global prevalence of internet access, the focus shifted to disparities of usage patterns and digital skills, defined as the second level of the divide (DiMaggio, 2004; Spracklen, 2015). A comprehensive understanding of digital inequalities is included in the theory of digital capital, which can be defined as the accumulation of digital competencies and ICT devices, that can be transformed into other forms of capital (Ragnedda, 2018).

Historically, the domain of leisure time has been a space for the articulation of personal taste and individual choices, serving as a signal of social status (Scott, 2017). Given the tendency of the digital space to mirror social stratification, the manner in which social classes utilise their leisure time in this domain has become an area of interest. The present study examines how different social classes navigate the digital space during their leisure time.

Objectives/Research Questions

The two focal points of this research are as follows: firstly, the extent of leisure time social classes have, and the times at which it occurs during the day; and secondly, the patterns in which they spend their leisure time on their phones. The objective of this study is to analyse the patterns of application usage, the types of applications utilised by different classes, and the number of applications between which they alternate during their leisure time. The results of the study contribute to the discussion on digital capital and the second level of the digital divide.

Methods

The database under consideration contains data on 500 individuals and is representative of adult Hungarian internet users. The initial data set was weighted using data from the 2022 Hungarian census to enhance representativity. The database was compiled from traditional survey responses and data collected using the Octopus Research Tool, and then the two data sources were linked by anonymous respondent IDs. The survey contains data on a number of factors, including the socio-demographic characteristics and social status of the participants. The Octopus Research Tool is an application that, once installed on a smartphone, logs the duration for which various applications are occupying the screen, along with the geographical coordinates of the users' location, recorded at five-minute intervals via GPS. Each participant provided a minimum of 14 days of data.

The analysis is founded upon a theoretical starting point, where social classes are defined based on the European Socio-Economic Classification (ESeC). The ESeC is based on occupational status, as outlined by the data collected in the survey. The five-class aggregated version of the ESeC is utilised, incorporating the following classes: Higher-level service class, Lower-level service class, Intermediate class, Skilled workers and Unskilled workers.

The definition of leisure time is based on the residual time approach, which defines leisure time as the time left after paid and unpaid work, and sleeping time. In the absence of data concerning unpaid work completed by respondents, leisure time is measured by deducting time spent working and sleeping from the day. In this study, sleep time is measured as the time between the last screen time in the evening and the first device use in the morning. The duration of working time is derived from the time spent at the respondent's place of employment, as determined by GPS coordinates. Consequently, these two indicators are unique to each individual. The analysis of application usage patterns is conducted with a distinction between weekdays, and weekends grouped together with public holidays.

The analysis of applications is conducted within the framework of two classifications: topic and type. Type is determined by hedonistic or utilitarian usage (Vujić, 2022). Utilitarian use is defined as applications designed for administration, productivity,

learning, organisation, and information gathering, e.g. email, PDF readers, internet browsers, language learning applications, navigation apps and health apps. Hedonistic use applies to applications designed for recreational, entertainment, and relaxation purposes, including gaming, streaming, and social media platforms.

Results

Preliminary results indicate that there is no significant difference in the amount of leisure time across social classes, nor in the amount of screen time within that leisure time. The timing of leisure time during the day is not uniform; skilled and unskilled workers check their phones earlier in the morning for the first time, indicating that they probably wake up earlier than the upper classes.

With regard to the application types, it appears that all classes are characterised by predominantly hedonistic application use during leisure time. However, a higher propensity for utilitarian applications outside of working hours is demonstrated by the upper classes when compared to the lower classes. Analyses of the prevalence of different categories of applications reveals that streaming apps are most prevalent among the Lower-level service class. Unskilled and Skilled workers allocate the most leisure time to gaming and using social media. The selection of social media applications is influenced by social class. Facebook is used most by the lower classes, while Instagram use is more common among the upper classes.

Future Work

The research can be taken in many further directions, one of which is particularly promising for future investigation. The theoretical starting point of this work enables me to examine whether differences in application usage patterns can be observed along traditional social classes. A future data-driven exploratory approach of this same question will reveal what differences in app usage can be observed that differ from traditional classes, and what socio-demographic characteristics can be used to define data-driven groups.

References

- DiMaggio, P., Hargittai, E., Celeste, C., & Shafer, S. (2004). From unequal access to differentiated use: A literature review and agenda for research on digital inequality. In K. Neckerman (Ed.), *Social inequality*, Russell Sage Foundation, 1(1), 355–400.
- Vujić, A., & Szabo, A. (2022). Hedonic use, stress, and life satisfaction as predictors of smartphone addiction. *Addictive behaviors reports*, 15, 100411. <https://doi.org/10.1016/j.abrep.2022.100411>
- Ragnedda, A. (2018). Conceptualizing digital capital. *Telematics and Informatics*, 35(8), 2366–2375. <https://doi.org/10.1016/j.tele.2018.10.006>.

Scott, D. (2017). Why Veblen Matters: The Role of Status Seeking in Contemporary Leisure. In: Spracklen, K., Lashua, B., Sharpe, E., Swain, S. (eds) The Palgrave Handbook of Leisure Theory. Palgrave Macmillan, London.

https://doi.org/10.1057/978-1-137-56479-5_22

Spracklen, K. (2015). Digital Leisure, the Internet and Popular Culture. Leisure Studies in a Global Era. Palgrave Macmillan, London.

https://doi.org/10.1057/9781137405876_1

3C.2 Impatient feminism: Platform-mediated intimacy advice on Chinese social media

Jiayi Chen (University of Warwick)

Introduction

In contemporary China, marriage and intimacy have become salient sites where women seek to negotiate their life chances, contributing to the proliferation of relationship advice on social media. As women's economic security and life prospects become increasingly contingent on their marital choices, and as regressive state policies and worsening employment conditions raise relational risks and the costs of exit, many turn to women-focused online advice communities to navigate intimate decisions. Yet, these spaces are also populated by pejorative labels such as “jiaoqi” (a compliant and pleasing “tradwife”), “love-addled brain,” and “marriage donkey.” The dispersion of these stigmas across platforms normalises ridicule, mockery, and derision as routine advice-giving modes to press for feminist “awakening,” targeting women perceived to remain “fooled” by patriarchal scripts of romance. This paradoxical symbiosis between help and policing in women's everyday relationship advice practices on social media warrants critical scrutiny.

Existing scholarship has examined mediated relationship advice as a key site shaped by neoliberal ethics of market rationality and self-entrepreneurship (Barker et al., 2018), producing postfeminist “feeling rules” (Hochschild, 2012; Gill & Kanai, 2018) that delimit permissible feelings, for instance, valorising confidence, positivity, and resilience while repudiating self-doubt and vulnerability. These emotional rules obscure structural inequalities by responsibilising women for their wellbeing through psychological self-work. In post-socialist China, similar affective expectations circulate within platform-mediated advice culture, yet not in the absence of, but alongside, a heightened awareness of structural gender injustice. This complicates a reductive reading of Chinese feminism as convergent with Euro-American neoliberal capitalism (Wu & Dong, 2019), and assumptions that structural awareness straightforwardly translates into collective action.

While neoliberal ideals of successful career women and the “big heroine” (dà nǚ zhǔ) have gained prominence, feminist discourse in China has also been punctuated by concerns about structural gendered oppressions, such as son preference, sexual harassment, unequal division of household labour, and domestic violence. However, within a broader state-market collaboration that mobilises social media platforms as tools of regulation and unevenly targets feminist critiques (Liao, 2024), feminist discussions often have to avoid confrontation with systemic injustice. This state-

platform-user configuration (Ye et al., 2024) plays a key role in shaping the framing and terms of mediated intimacy advice. Against these backdrops, this study examines relationship-advice communities on Douban, a major Chinese social media platform. It foregrounds how feminist advice practices may turn into disciplinary mechanisms under repressive digital contexts.

Objectives/Research Questions

This study aims to contribute a timely analytical intervention and a situated critique of platform-mediated intimacy advice flourishing in the Chinese digital media landscape. It examines how feminist advice communities, while oriented towards support and empowerment, also generate normative ideals and affective expectations that discipline women's feelings and responses to relationship injuries. Rather than positioning these dynamics merely as a "failure" of popular feminist critiques, the study approaches them as situated practices shaped by platform and wider socio-political conditions.

Through a case study of relationship-advice communities on Douban, a widely used platform for gender- and relationship-related discussions, the study explores how tacit emotional norms around intimacy are produced, circulated, and negotiated in everyday digital interactions, and how these advice practices organise relations of authority and exclusion in feminist spaces, shaping boundaries of feminist legitimacy. By focusing on Douban's discussion groups, it investigates how community-based regulation operates in relation to platform governance and state censorship, and highlights how platform design and technological affordances shape the affective dynamics of advice-giving within a tightly moderated discursive space.

Methods

The empirical data were collected from Douban through a digital ethnography approach (Lindgren & Krutrök, 2024). With urban women constituting its user majority, Douban features various interest-based, user-organised forums (called "Groups") that enable members to seek advice, share experiences, and debate gender issues, making it a hub for Chinese popular feminism (Tan & Liu, 2025). From Douban's official category of "relationships and emotions," I selected four of the most active and popular relationship-advice groups. The data collection began with non-participatory, systematic observations (Androutsopoulos, 2008), aiming to map recurring topics, interaction patterns, governance practices, and platform visibility mechanisms. As part of the ethnographic fieldwork, I developed a corpus for close analysis, centred on approximately 120 advice-seeking posts curated as "featured" content by group moderators and the top-liked comments foregrounded by the platform for each post (typically up to five, yielding around 600 comments). Attention to engagement metrics allows me to capture dominant discursive formations and norm-setting affects. This

corpus was iteratively revisited and expanded in response to emerging analytical concerns, enabling a close tracing of how key figures, associations, and affective norms were elaborated across different interactional contexts.

The analysis draws on Wetherell's (2012) affective-discursive approach to examine how emotional norms are produced and propagated through everyday advice practices. The coding process involved a theory-driven reading of discussion threads, focusing on two interrelated analytical dimensions. First, I developed a set of interpretative repertoires to capture recurring ways of making sense of romantic relationships, emotional ambivalences, and gendered vulnerabilities. These repertoires were identified through buzzwords, memes, repeated tropes, narrative framings, and sticky associations. Second, I scrutinised how the groups' affective-discursive practices, as patterned forms of activity, organise broader relational patterns of proximity and distance, affiliation and attachment (Wetherell et al., 2015). Drawing on Ahmed's (2004) theorisation of affective economies, I explored how affects move through the groups and mediate relationships between advice-seekers and advice-givers by producing affective identification, fostering belonging and alliances, and demarcating boundaries of inclusion. I paid particular attention to members' practices of emotional self-regulation and affective performance in their efforts to render their relationship dilemmas and emotional experiences legible and legitimate in the advice communities.

Results

To understand how derision and mockery become normalised in platform-mediated advice practices, I position impatience as a dominant mediating affect through which popular feminist advice culture in China is organised. In Douban's discussion groups, members repeatedly call on advice-seekers to move on from toxic or unworthy relationships and police those who remain emotionally attached, or ambivalent about breakup or divorce. As a pedagogical mode of advice-giving, impatience structures the relationality between posters and commenters. Comment sections are dominated by dismissive judgement and sharp ridicule that mock the posters' failure to detect, exit, or recover from relationship harm swiftly. While commenters sometimes offer quick, diagnostic clarity—for example, by identifying manipulative behaviours or citing high-profile cases of domestic violence—such clarity is restricted to individualised and moralised labelling, ultimately responsibilising women to predict and prevent relational risks.

The circulation of impatience is further amplified by Douban's platform features and affordances. Online anonymity reduces social inhibition (Suler, 2018), enabling harsher judgement. Top-liked comments—those typically marked by bluntness—are foregrounded as “trending,” while formulaic “move-on and feel-good” narratives are liked, reposted, and curated as featured content. The economy of visibility (Banet-

Weiser, 2018) sustains a feedback loop that rewards impatience as the affective norm in the Douban communities.

This platformised impatience also turns inwards when negative feelings such as pain or self-doubt disrupt the prevailing ideal of an independent and emotionally invulnerable “big heroine.” Women thus become impatient with their own injuries, resorting to vigilant emotional self-management, peer surveillance, and the public performance of rapid recovery.

Impatience indexes how platform governance simultaneously constrains political frustration while amplifying individual-targeted disappointment and urgency for personal change. Where platform governance and group-level moderation attune users to the speakable, and institutions are not experienced as addressable targets of action but given and monolithic, feminist frustration is rerouted towards those perceived as the only imaginable agents of change—individual women themselves, often manifesting as impatience towards the “unawakened” other. Positioning psychological self-transformation as the necessary—and the only—path to reclaim marital and reproductive autonomy, women enact impatience in advice-giving as a mechanism of “accelerating awakening” and a perceived feminist duty of mutual help. Feminist awareness paradoxically “necessitates” emotional discipline directed at individual women, producing entanglements between protection and surveillance, help and policing that are themselves situated yet curtailed responses to structural inequalities.

Future Work

This study brings feminist affect theories into dialogue with platform and social media studies. Rather than positioning social media platforms as a mere background against which the cultural politics of emotions unfold, future research could examine how platforms shape the visibility, circulation, and trajectories of different affective expressions in online communities. When upward political frustration is constrained under conditions of digital authoritarianism, such frustration may travel sideways and inwards, taking the form of interpersonal impatience and moral urgency. This calls for a closer examination of the internal dynamics of user-driven, community-based governance and its interplay with platform-level censorship within multi-layered governance structures, beyond a linear model of top-down governance. Future work could further trace how non-cathartic, mundane, and censorship-enduring affects—such as impatience, disappointment, and resignation—come to organise women’s everyday digital interactions and shape community cultures.

References

- Ahmed, S. (2004). Affective economies. *Social Text*, 22(2), 117–139.
https://doi.org/10.1215/01642472-22-2_79-117

- Androutsopoulos, J. (2008). Potentials and limitations of discourse-centred online ethnography. *Language@Internet*, 5(8).
<https://web.archive.org/web/20150331080838/http://www.languageatinternet.org/articles/2008/1610/metadata>
- Banet-Weiser, S. (2018). *Empowered: Popular feminism and popular misogyny*. Duke University Press.
- Barker, M.-J., Gill, R., & Harvey, L. (2018). Mediated intimacy: Sex advice in media culture. *Sexualities*, 21(8), 1337–1345.
- Gill, R., & Kanai, A. (2018). Mediating neoliberal capitalism: affect, subjectivity and inequality. *Journal of Communication*, 68(2), 318–326.
<https://doi.org/10.1093/joc/jqy002>
- Hochschild, A. R. (2012). *The managed heart: Commercialization of human feeling*. University of California Press.
<http://ebookcentral.proquest.com/lib/warw/detail.action?docID=870020>
- Liao, S. (2024). The platformization of misogyny: Popular media, gender politics, and misogyny in China's state-market nexus. *Media, Culture & Society*, 46(1), 191–203. <https://doi.org/10.1177/01634437221146905>
- Lindgren, S., & Krutrök, M. (2024). *Researching digital media and society* (1st ed.). SAGE Publications, Ltd. (UK). <https://bookshelf-ref.vitalsource.com/books/9781529679281>
- Suler, J. (2018). Cyberpsychology. In *Textbook of applied psychoanalysis* (1st ed.). Routledge. <https://www.taylorfrancis.com/reader/download/59eb4da2-76c3-48d9-806b-ef6593ed0d8a/chapter/pdf?context=ubx>
- Tan, Y., & Liu, S.-D. (2025). “Our advice is to break up”: Douban's intimate public and the rise of girlfriend culture. *New Media & Society*, 27(11), 6156–6177.
<https://doi.org/10.1177/14614448241264450>
- Wetherell, M. (2012). Negotiating affect: discourse, representation and affective Meaning-Making. In *Affect and emotion: A new social science understanding* (Chap.3). SAGE Publications Ltd. <https://doi.org/10.4135/9781446250945>
- Wetherell, M., McCreanor, T., McConville, A., Moewaka Barnes, H., & le Grice, J. (2015). Settling space and covering the nation: Some conceptual considerations in analysing affect and discourse. *Emotion, Space and Society*, 16, 56–64.
<https://doi.org/10.1016/j.emospa.2015.07.005>
- Wu, A. X., & Dong, Y. (2019). What is made-in-China feminism(s)? Gender discontent and class friction in post-socialist China. *Critical Asian Studies*, 51(4), 471–492.
<https://doi.org/10.1080/14672715.2019.1656538>

Ye, Z., Huang, Q., & Krijnen, T. (2024). Douyin's playful platform governance: Platform's self-regulation and content creators' participatory surveillance. *International Journal of Cultural Studies*. <https://doi.org/10.1177/13678779241247065>

3C.3 “My cat asks you to buy a drone”: How civilians in Ukraine strategically navigate social media logic to run small grassroots military fundraising campaigns

Kateryna Bystrytska (Rutgers University)

Introduction

The full-scale invasion of Ukraine catalyzed a civil-rights movement rooted in Euromaidan volunteering, now serving as a vital “condition for the consolidation of the nation” (Pankiv 2023, p. 3). This dynamic underpins “participatory warfare,” where civilians contribute to national defense through daily resource mobilization and information sharing (Lokot, 2023). While existing research addresses macro-sociopolitical aspects of wartime volunteering, the micro-level practices of non-professional grassroots fundraisers remain under-explored.

This study argues that small civilian fundraising campaigns help better understand the nature of war participation and map the role of social media platforms in wartime or crisis. It examines how small-scale civilian campaigns on Instagram navigate the complexities of digital activism.

There are a few things that complicate small grassroots fundraising campaigns in the war-torn country.

First, such campaigns are usually run by people with no prior fundraising experience. Technological advancement and social media affordances make these campaigns possible even for individuals with personal social media accounts, in comparison with NGO fundraising, which is often run by a team of professionals (Lauro et al., 2019).

Second, people learn in practice how to boost military fundraising posts on personal social media (mainly Instagram and Facebook) and can already create a guide on how to build successful campaigns, manage and understand platform affordances and algorithms, engage with others, and make each campaign unique. At the same time, they need to learn how to make the campaign visible, as none of those platforms prioritize political content (Bast, 2021). It illustrates how folk theories about social media algorithms affect online activities.

Those small fundraising campaigns are often personal, as people fundraise for their family or friends on the frontline, which makes every campaign a puzzle for understanding how civilians live through war trauma and become a part of participatory warfare. This study offers a systematic digital ethnographic observation of 50 grassroots campaigns on Instagram, filling the gap between NGO-led fundraising and personal efforts to support the army.

Objectives/Research Questions

Small grassroots fundraising for the military is a unique issue, not only in Ukraine but in any other country that has lived through war or crisis. This research aims to examine how this participatory social media fundraising culture emerged, what strategies people use to reach fundraising goals through social media, and how platform vernacular and affordances are adopted for military fundraising in wartime. While previous research focused broadly on volunteers in wartime (Boichak & McKernan, 2024), this study aims to explore the role of social media in small-scale, grassroots fundraising among people who may not identify themselves as volunteers or social media influencers.

This study focuses on military volunteers in Ukraine, but its insights are relevant to all political fundraising efforts conducted individually on social media. Previous research indicates that Instagram restricts political content, encouraging users to create posts that generate higher engagement (Trevisan et al., 2019). Nevertheless, society leverages social media platforms not only for entertainment but also as effective tools for political participation and digital activism. In the context of the war in Ukraine, social media became an effective tool for fundraising to support the military's needs.

The study was guided by the following questions:

RQ1: How do Instagram users use platform affordances and vernacular to construct military fundraising campaigns in Ukraine?

RQ2: What are the common attributes of military fundraising campaigns for the Ukrainian army on Instagram?

Methods

This study is grounded in digital ethnographic observation (Varis, 2015) of selected Instagram fundraising campaigns.

Acknowledging that hashtags are valuable for understanding online content flow (Highfield & Leaver, 2015), the sampling focused on hashtags previously used in Ukraine for social media fundraising campaigns. To select campaigns for observation on Instagram, the research used Ukrainian hashtags #36ip #ДопомогаАрмії #36ip3CV (translated as #Fundraising, #SupportArmy, #FundriseArmy). Then, a critical approach was applied to select posts that illustrate small grass-roots campaigns. The criteria for the selection were: 1) the content is about fundraising for the Ukrainian military; 2) the fundraiser does not position itself as a social media influencer. As a result, the research corpus consisted of 50 Instagram fundraising campaigns, published between May and September 2025.

Following the principles of digital ethnography research on social media platforms (Caliandro, 2018), this study aimed to observe and understand the “online social

formations emerging from different practices of the use of digital devices enacted by users as well as the meanings they attribute to activities deploying such social formations” (p 554). For this study, digital ethnography prioritizes openness to engagement with a field, as this ethnography is a first step in a bigger project, and interaction and communication with creators via platforms can enhance the understanding of the principles of fundraising campaigns.

The goal of this ethnographic study is to identify patterns of interaction, recurring themes, and strategies used by campaign creators and supporters to mobilize resources and attention. Through an analysis of posts, comments, and engagement metrics, the study documents how people create fundraising campaigns and leverage social media affordances and algorithms to boost them.

Following Rogers' (2021) suggestions for observing visuals on social media, this study focuses on image trends and vernaculars, dominant images, image circulation, and engagement to create a metapicture of how a fundraising campaign is represented on Instagram.

Results

For this paper, the findings are preliminary and can be systematized as follows. Small grassroots campaigns on Instagram that aim to fundraise for military needs have a list of similar patterns.

1. Network-driven visibility. Personal connections are important for campaign visibility: users organize campaigns for friends and colleagues and rely on their close circle for both donations and amplification. The story of the battalion or personal connection with a soldier is an important puzzle in fundraising storytelling. Thus, some fundraisers ask to add a few personal words about the campaign when sharing (“Please, when sharing the image with the campaign, mention that you know me personally, and this is the fundraising for my brother”).
2. Interactive and experimental learning is a part of campaigning. As the observation shows, a fundraising campaign often changes tactics during an experiment. Fundraisers learned a lot from other campaigns and improved their practice by implementing new methods to bring attention to the military's needs. As users learn about Instagram algorithms while running campaigns, they often adjust their approach to audience reach when fundraising is moving slowly.
3. Collaborative micro teams can be more effective in fundraising. Most initiatives are run by small, ad hoc teams (typically 2 - 5 people, often with relatives or friends in the army) who coordinate resources, share content libraries, and co-manage donation streams, effectively scaling individual efforts into broader support networks. When the

fundraising goal is big for one person, friends or family can support it by opening smaller “donation jars” (a bank’s tool for volunteering and fundraising).

4. Strategic use of Instagram affordances. Fundraisers blend personal photographs and narrative captions with high engagement formats such as Reels, short animations, and story highlights to signal legitimacy and urgency. Direct, unambiguous calls to action (“Donate now”) are paired with visual storytelling to boost algorithmic favorability, for example, by using pets in information campaigns (“My cat/dog asks you to donate”). The majority of analyzed campaigns used a human image (a photo of someone collecting money or someone who needs it) or pets to boost audience interaction. They also manage the amount of personal content before the campaign launch (for example, posting on Instagram Stories about their daily life for a few days before launching the fundraising).

5. Direct messaging as a momentum rescue tool. When public engagement wanes, creators shift to personalized DMs, targeting previously engaged followers with tailored reminders and progress updates; this approach yields a measurable uptick in late-stage contributions compared with relying solely on feed visibility. Participants often mention on social media that they plan to DM friends or ask them to react to posts or stories to make their content more visible.

Future Work

After using digital observation to analyze existing online campaigns, this study aims to conduct an interview with the creators of the campaigns to explore their understanding of Instagram fundraising campaigns and the strategies they use to make them successful. Qualitative semi-structured interviews may deepen the discussion about the role of such campaigners in war perception. An experimental online study from the other side can help explore which types of campaigns are most effective for civilian participation. Another goal is to explore fundraising on other social media platforms to get a broader picture of participatory practices in wartime.

References

- Bast, J. & University of Koblenz-Landau, Germany. (2021). Politicians, Parties, and Government Representatives on Instagram: A Review on Research Approaches, Usage Patterns, and Effects. *Review of Communication Research*, 9, 193.
<https://doi.org/10.12840/ISSN.2255-4165.032>
- Boichak, O., & McKernan, B. (2024). Narratives of Volunteering and Social Change in Wartime Ukraine. *Cultural Sociology*, 18(1), 48–71.
<https://doi.org/10.1177/17499755221127877>

- Caliandro, A. (2018). Digital Methods for Ethnography: Analytical Concepts for Ethnographers Exploring Social Media Environments. *Journal of Contemporary Ethnography*, 47(5), 551–578. <https://doi.org/10.1177/0891241617702960>
- Highfield, T., & Leaver, T. (2015). A methodology for mapping Instagram hashtags. *First Monday*, 20(1), Article 1.
- Lauro, S. D., Tursunbayeva, A., & Antonelli, G. (2019). How Nonprofit Organizations Use Social Media for Fundraising: A Systematic Literature Review. *International Journal of Business and Management*, 14(7), p1. <https://doi.org/10.5539/ijbm.v14n7p1>
- Lokot, T. (2023). The role of citizens' affective media practices in participatory warfare during Russia's invasion of Ukraine. *Journal of International Relations and Development*, 26(4), 776–790. <https://doi.org/10.1057/s41268-023-00317-y>
- Pankiv, O., Onyshchuk, O., & Stebelska, O. (2023). MILITARY VOLUNTEERING AS A UNIQUE EXPERIENCE AND ETHICAL PRACTICE. *Trames. Journal of the Humanities and Social Sciences*, 27, 175–188. <https://doi.org/10.3176/tr.2023.2.05>
- Rogers, R. (2021). Visual media analysis for Instagram and other online platforms. *Big Data & Society*, 8(1), 20539517211022370. <https://doi.org/10.1177/20539517211022370>
- Trevisan, M., Vassio, L., Drago, I., Mellia, M., Murai, F., Figueiredo, F., Couto da Silva, A. P., & Almeida, J. M. (2019). Towards Understanding Political Interactions on Instagram. *Proceedings of the 30th ACM Conference on Hypertext and Social Media*, 247–251. <https://doi.org/10.1145/3342220.3343657>
- Varis, P. (2015). Digital ethnography. In *The Routledge Handbook of Language and Digital Communication*. Routledge.

Session 3D - Discourse and Public Opinion (3)

Day 1 · Mon 13 July · 14:00–15:00 Studio 2, ARC

3D.1 "Chat, Is This Real?": Truth-Seeking Heuristics & Distributed Trust on Reddit's r/isthisAI

Kaitlyn McNamee (Marist University)

Introduction

In the current age of generative AI, internet users encounter new challenges to trust as AI-generated images and texts populate user spaces. As this technology rapidly advances, determining whether media is human made or AI-generated has become significantly more challenging. To gain epistemic clarity, humans have introduced AI text checkers, AI image watermarking detectors, and AI audio taggers as methods for detecting AI-generated material (Crew & Savage, 2025). While these programs attempt to offer proof of human or AI generation, their capabilities are limited and may lead to false determinations of AI generation in human-created work (Liang et al., 2023). At present, there is no technological method to determine entirely if media has been AI generated. As such, the responsibility to make this determination now falls to humans.

Decentralized information sharing in digital spaces is a key aspect of online culture, and one that allows for collective knowledge to be both utilized and tracked. Apps like iNaturalist and the ornithology eBird app have revolutionized the ability to identify wildlife through online citizen science. Github's repository for open source code has changed the landscape of coding. Now, in the age of AI, internet users are increasingly relying on this type of collective knowledge sharing to identify AI-generated work. One such space is Reddit's r/isthisAI community, an online forum where users determine whether or not a digital artifact is AI generated following a collective reasoning, or "wisdom of the crowd" approach (Surowiecki, 2004). This paper examines the r/isthisAI community to assess the methods by which internet users collectively determine if digital artifacts are AI generated or human made. The research explores how this type of assessment creates rules, referred to in the paper as truth-seeking heuristics, through the affordances of the Reddit platform. This research also considers how positive or negative language functions within these truth-seeking heuristics, specifically looking at positive and negative sentiment toward the person who originally posted the digital artifact.

Objectives/Research Questions

The objective of this research is to map the mechanisms of collective trust and assess how communities organize themselves in response to uncertainty. The aim of this research is to answer the question of: how does a decentralized group reach a consensus about truth? In doing so, how does the community develop heuristics, and how many unique pieces of information are required in order to satisfy collective truth seeking? Finally, how does language and sentiment contribute to collective trust building?

This study applies social epistemology and theories of distributed cognition to assess how trust and truth seeking become a socially distributed process under conditions of uncertainty (Hollan et al., 2000). To understand ways in which sentiment and language contribute to rule development, this study also uses sentiment analysis to assess positive, negative, or neutral linguistic markers in the comments. Specifically, the research looks at how users address the original poster of the digital artifact and consider if user sentiment influences the number of new contributions to the heuristic, as well as how these sentiments may sway the collective decision.

Methods

This research takes an interdisciplinary approach to understanding group behavior in digital spaces. First, I present the case study of r/isthisAI by first giving an overview of the affordances of the community and a description of how it is moderated, including the hierarchy of moderators and users, as well as an overview of the community upvote and downvote system. I then review four Reddit posts: two in which the subject of the post was determined to be AI generated, and two in which the subject was determined not to be AI generated. All posts consist of over a thousand "upvotes" and at least three hundred comments. I assess the user comments to aggregate original epistemic claims that explain why the subject is or is not AI generated. I analyze the comments that contribute to the outcome, looking specifically at how logical reasoning is portrayed via discourse, and how many original pieces of information contributed to the determined outcome of real or AI generated. I also consider dissenting opinions and counter narratives to understand how these contribute to the overall argument. Within these comments, I examine whether the overall discourse contains positive, negative, or neutral sentiment markers aimed at the original poster, and I assess the ways in which language contributes to outcomes. Finally, I compare the results from the posts and consider the overall differences in reasoning and discourse when posts are determined to be AI generated versus human made.

Results

Preliminary results show that posts in which the artifact is collectively determined to be AI generated present more pieces of unique information to contribute to the outcome

than posts with artifacts that are determined to be human made. These posts also contain more positive language and affirmations within the group, as well as more positive sentiments aimed at the original poster. Posts determined to be human made contain more negative discourse toward the original poster than posts determined to be AI.

Overall, this community presents a case study in collective truth seeking that demonstrates effective group reasoning. The rules and language of the community contribute to the development of heuristics for collective truth seeking, and this process successfully identifies AI-generated content with relatively strong certainty. As AI-generated content becomes increasingly difficult to identify, this method of collective reasoning and human-centered assessment of digital artifacts will also become increasingly difficult to maintain. Developing heuristics early based on collective knowledge and agreed upon standards may aid in future attempts to curtail uncertainty caused by AI-generated content. Understanding the foundations of collective truth seeking in response to AI will provide important insight into community trust as generative AI becomes a more established issue in digital spaces.

Future Work

The pilot study presented in this research demonstrates multiple needs when handling uncertainty around AI generated media. Within digital communities, there is a need for more epistemic clarity now that AI generation has introduced doubt into user spaces. In creating collective truth seeking, there is a need for rules and heuristic development based on accurate knowledge sharing. Within this structure, once rules have been established, users must contribute to outcomes or clarity cannot be established. There is room to explore how discourse contributes to these outcomes, and there is a need to better understand how epistemic clarity is perceived by reviewing the language of the decision-making process.

Future work can take multiple approaches. One approach would be aggregating and assessing more posts to create further data. With these data, future work could take a linguistics approach to give insight into group sentiment and degree of certainty using linguistic markers. Additionally, the work may be expanded to consider other digital platforms that utilize distributed trust. One example is Wikipedia, where changes are traced, argued, and verified by a truth seeking collective. Future research might also involve examining problems with collectively created heuristics in online communities by considering how persuasive rhetoric might sway groups toward inaccuracies or reinforce false beliefs.

References

- Crew, J., & Savage, M. (2025, November 21). How can you tell if music is AI-generated? BBC. Retrieved January 17, 2026, from <https://www.bbc.com/news/articles/c5ylzjj5wzwo>
- Hollan, J., Hutchins, E., & Kirsch, D. (2000, June 01). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(2), 174 - 196. <https://dl.acm.org/doi/10.1145/353485.353487>
- Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4(7). <https://arxiv.org/abs/2304.02819>
- Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. Doubleday.

3D.2 From 'Diaploki' to Pluralism - Does Social Media Still Drive Change in Greece's Political and Media Spheres?

Michael Nevradakis (College Year in Athens)

Introduction

The previous decade was one of economic, political and social upheaval in Greece. The Greek financial crisis saw the rise of a widespread protest movement that occupied city squares, which gave rise to several new political parties and contributed to the electoral rise of previously marginal parties and the sharp decline of the establishment parties that had governed Greece since the end of the military dictatorship in the 1970s. This phenomenon was characterized by a growing credibility crisis impacting most societal institutions, including the mainstream news media (Kapa Research, 2016), and by the electoral victory of the left-wing SYRIZA party in 2015, having run on a platform that rejected the austerity measures imposed on Greece by its 'troika' of creditors as part of their so-called 'bailout' of the country's economy.

Social media played a catalytic role in this process. The movement of the Greek 'Indignants' in 2011 started as a Facebook invitation that went 'viral', while numerous new political parties and initiatives engaged in civic activism instrumentalized social media to engage with the public, bypassing what was widely seen as the gatekeeping of legacy news media outlets (Nevradakis, 2018) and the interplay of the owners of these media organizations — wealthy oligarchs — with Greece's political elite, a phenomenon encapsulated by the uniquely Greek term 'diaploki' (Sims, 2003).

Beginning in the mid-2010s, however, the public sentiment in Greece began to shift — a change attributed to 'protest fatigue' as the financial crisis wore on, but also to what was widely seen as the betrayal of the Greek anti-austerity electorate by SYRIZA and its governing partner, the populist-right Independent Greeks political party. Having campaigned on an anti-austerity platform and having organized the 2015 referendum in which voters overwhelmingly rejected further austerity — the SYRIZA-Independent Greeks coalition subsequently agreed to new austerity measures. In 2018, the same government signed the Prespes Agreement, resolving a longstanding name dispute between Greece and the country that became known as North Macedonia. The agreement was unpopular and opposition to it resulted in the largest protests in Greece since the early 2010s.

The end of the 2010s saw the return to power of one of Greece's traditional parties, New Democracy, and a narrative of a return to 'growth' — one that was interrupted by the COVID-19 pandemic and a scandal where the government subsidized COVID-related public service announcements to the tune of 40 million euros (Kalogeropoulos, 2022).

The ‘growth’ narrative returned as the pandemic waned, and New Democracy was re-elected in 2023.

That year was marked, however, by the collision of two trains in the Tempi region — resulting in 57 deaths, mostly university students — and subsequent mass outrage amid allegations of gross negligence and, subsequently, accusations that the government was interfering in the investigation.

Objectives/Research Questions

The Tempi incident fueled a new wave of protest and popular discontent in Greece, and much as SYRIZA leader Alexis Tsipras emerged as an ‘oppositional’ leader early during the Greek financial crisis, Maria Karystianou, mother of one of the Tempi deceased, emerged as the emblematic ‘face’ of the ‘Justice for Tempi’ movement. In late 2025, she announced the forthcoming launch of a new political party that would challenge the entirety of Greece’s political establishment. The Tempi movement has been accompanied by a series of other movements questioning the current government and, in some cases, the legitimacy of the political system, including the OKEPEPE scandal involving the mismanagement of EU farm subsidies, the “Movement of the Towel,” demanding that public access to Greek beaches be preserved in the face of encroaching commercialization, and the Greek farmers’ protests in late 2025 and early 2026, during which they blockaded Greek highways for over 50 days in a direct confrontation with the government. During this period, “Mitsotaki Gamiesai” (“Mitsotakis go f*** yourself,” referring to center-right Greek Prime Minister Kyriakos Mitsotakis) went ‘viral’ as a rallying cry against his government.

The early-to-mid 2010s saw social media challenge the status quo in the Greek mediascape, public sphere and civil society (Nevradakis, 2024). Several grassroots movements, including those opposing gold mining operations in the environmentally fragile region of Skouries and the government shutdown of public broadcaster ERT in 2013, heavily used social media to coordinate opposition. Several new online- and social media-based media outlets produced ‘citizen journalism’ and used social media to collect and disseminate news to an audience increasingly disenchanted with incumbent media outlets. This paper seeks to explore social’s media role in the Greek public sphere and Greek civil activism in the post-2023 period by addressing the following research questions: Has the post-2023 period seen a rebirth of the popular mood of the early 2010s? Have social media and other alternative online channels of communication again been used to challenge establishment narratives and have they helped inject new and oppositional voices into the Greek public sphere? Have social media contributed to the rise of the ‘Justice for Tempi’ movement and the popular “Mitsotaki gamiesai” sentiment?

Methods

This paper draws upon semi-structured interviews with key societal influencers, a content analysis of content circulating on Greek social media in the post-2023 period and a review of the findings of popular opinion polls conducted during this period. This research serves as a follow-up to longitudinal research conducted in the field in Greece between 2011 and 2017, with several follow-up interviews of influencers who were first interviewed in the 2011-2017 period.

Results

This paper will argue that while Greece ranks highest among the 50+ countries surveyed in the Reuters Institute/Oxford University Digital News Report in the consumption of online news and the use of social media as sources of news — and also ranks lowest in news trust (Kalogeropoulos, 2025), ‘traditional’ media and traditional methods of collective action (Bennett & Segerberg, 2012) have been significantly more instrumental in galvanizing what remains a disparate and fractured opposition to the current government and the political establishment more broadly despite the ongoing credibility crisis (DiaNEOsis, 2024). Television has helped disseminate, and perhaps even fuel, popular dissatisfaction over the Tempi and farm subsidy scandals, with Greece’s Parliamentary TV channel attaining the largest viewership in its history during its live coverage of the parliamentary inquiry into the OKEPEPE scandal (Documentonews.gr, 2025). In turn, the Greek farmers’ protests and political developments such as the return of former left-wing Greek prime minister Alexis Tsipras to the political arena, have largely been offline media spectacles marked by traditional strategies of protest or politicking, indicating that not only were the civil activist initiatives launched with the help of social media in the 2010s ephemeral in nature, but so was social media’s role as a driver of political development in Greece.

Instead, social media has adopted a secondary role of retransmitting content from establishment and quasi-establishment media sources — albeit frequently accompanied by criticism and commentary — and as a conduit for transmitting memes such as “Mitsotaki Gamiesai” and those making light of the sometimes ludicrous revelations stemming from scandals such as that involving OKEPEPE. In turn, legacy media continues to play an outsized role in the Greek mediascape, which in turn contributes to the continued entry of wealthy oligarchs into Greece’s already saturated and unprofitable Greek news market.

Future Work

This research is being conducted toward the publication of a monograph on social media’s influence within the Greek public sphere and civil society during and after the Greek financial crisis.

References

- Bennett, W. L., & Segerberg, A. (2012). The logic of connective action: Digital media and the personalization of contentious politics. *Information, Communication & Society*, 15(5), 739-768.
- DiaNEOsis. (2024). Ti pistevoun e Ellines – Meros A' [What Greeks believe – Part 1] [Data file]. https://www.dianeosis.org/wp-content/uploads/2024/04/TPE2024_Part_A.pdf
- Documentonews.gr (2025, December 12). Para to 'dikaïoma tis siopis', o 'Frappes' efere rekor tiletheasis sto kanali tis Voulis [Despite the 'right to remain silent', 'Frappes' brought record viewership to the Parliamentary channel]. Documento. <https://www.documentonews.gr/article/para-to-dikaioma-tis-siopis-o-frappes-efere-rekor-tiletheasis-sto-kanali-tis-voulis/>
- Kalogeropoulos, A. (2022, July 6). Greece. In Reuters Institute digital news report 2022 [data file]. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2022-06/Digital_News-Report_2022.pdf
- Kalogeropoulos, A. (2025, June 17). Greece. In Reuters Institute digital news report 2025. <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2025/greece>
- Kapa Research. (2016, October 25). I krisi ton thesmon kai i apohi [The crisis of the institutions and abstention] [Data file]. https://kaparesearch.com/wp-content/uploads/2017/12/KapaVima_Oct16_GR.pdf
- Nevradakis, M. (2018). From the polis to Facebook: Social media and the development of a new Greek public sphere (Ph.D. thesis). The University of Texas
- Nevradakis, M. (2024). "Can social media expand public discourse in a 'captured' mediascape? The case of Greece." *Central European Journal of Communication*, 17(2, 36), 268-284.
- Sims, J.R. (2003). Politicians and media owners in Greek radio: Pluralism as diaploki. *Journal of Radio Studies*, 10(2), 202-215.

3D.3 A Multimodal Discourse Analysis of the Iranian #MeToo movement on the Islamic Republic Media

Pardis Yarahmadi (QUT)

Introduction

The #MeToo hashtag was brought to global prominence by Alyssa Milano's 2017 tweet on Twitter against the famous Hollywood producer Harvey Weinstein, speaking about his sexual violence against her (Gill & Imran Rahman-Jones, 2020). While the movement originated over a decade earlier, when survivor and activist Tarana Burke created the "MeToo" community in 2006 to support survivors of sexual assault (Burke, 2026) Milano's tweet inspired a variety of hashtags and campaigns worldwide. These (often national) hashtags and campaigns brought together victims of sexual harassment, encouraging them to speak up and to stand for their rights, and had significant impacts on media, policy and other platforms (e.g. Noetzel et al., 2023; O'Halloran & Cook, 2024; Serisier, 2020, 2020; Starkey et al., 2019).

In Iran, the movement shared global features but also developed unique characteristics influenced by cultural concepts. Iran-specific themes include breaking the silence (Lee, 2018; O'Halloran & Cook, 2024; Serisier, 2020; Starkey et al., 2019), victim-blaming (Abraham, 2021; Bongiorno et al., 2020; Lucarini et al., 2020; Shoaf et al., 2024), creating a more open discourse to discuss sexually-related concerns (Ghosh et al., 2022; Noetzel et al., 2023), and public support, empathy and solidarity (Manikonda et al., 2018; Peleg-Koriat & Klar-Chalamish, 2020). There are more cultural, patriarchal and religious-related themes, such as 'haya' or modesty (cultural shame and blame), and 'aberoo', honour or family reputation (pressure to remain silent to protect family reputation) (Basmechi, 2023; Foomani & Reimer, 2025; Yaghoobi et al., 2024).

The Iranian #MeToo conversation gained momentum on Twitter in 2020 following allegations against Keyvan Imam Verdi, a student rather than a celebrity, prompting thousands to speak out on social media and sparking public debate about sexual assault. Subsequent accusations against prominent artists, such as Mohsen Namjoo (2021) and Pejman Jamshidi (2025), have continued to this day. The Islamic Republic's media coverage of these cases highlights how journalism addresses gender, power, and public values. However, limited research exists on how Iranian news outlets report or overlook #MeToo stories.

Media censorship remains a significant concern in the Islamic Republic of Iran (Rahimi, 2015) and is frequently used to serve the regime's propaganda objectives (Goodrich, 2025). Despite well-documented reports of misogyny and gender inequality in Iran (Kamrava, 2024), the response of state media to sexual harassment is not well studied.

This research aims to clarify how news outlets and the state address these issues in the current social and political context.

Objectives/Research Questions

This study analyses three sexual assault cases: Imam Verdi (2020), Namjoo (2021), and Pejman Jamshidi (2025). These cases were selected for their significance and diversity. Imam Verdi was accused of drugging and assaulting women. The allegations surfaced on social media, where victims used the Metoo hashtag to bring the case to court (Rezaei, 2022). When the hashtag gained popularity, many actresses supported the movement by sharing their stories of harassment by other actors, directors, and other men, which created a more open atmosphere for such topics.

Following their steps, non-celebrity victims began sharing accounts of sexual abuse by the singer Mohsen Namjoo. Namjoo, an Iranian singer with anti-regime views, had already left the country due to his religious and political beliefs (Namjoo, 2025) when six victims spoke up about their experience. Pejman Jamshidi (The Week News Desk, 2025), a former soccer player and actor known for his apolitical stance, was accused of rape by a young woman in October 2025, who brought the case to court. Subsequently, additional individuals alleged sexual abuse by Jamshidi (Gholami, 2025). While details remain unclear due to the current situation in Iran, the case has received significant media attention. This case is important as it offers insight into the evolving media response to sexual abuse allegations.

This diverse sample helps minimise bias. To assess domestic media coverage of these cases in Iran, I use qualitative critical discourse analysis (CDA) to examine texts, images, and videos. The study addresses the following questions:

1. How have Iranian news outlets presented the cases involving Keyvan Imam Verdi (2020), Mohsen Namjoo (2021), and Pejman Jamshidi (2025)?
2. What language, images, and themes do Iranian media use to shape stories about victims, perpetrators, and the wider #MeToo movement?
3. What does the news coverage reveal about public discussion and gender politics in Iran over these five years?

The main goal is to explore how institutional reporting, rather than social media's grassroots stories, shapes public views of sexual violence, accountability, and gender justice. This study aims to provide a fuller picture of how the Islamic Republic media influences, limits, or reshapes feminist conversations under political and social constraints.

Methods

This study employs a qualitative, comparative Critical Discourse Analysis (CDA) framework, incorporating elements of Multimodal Critical Discourse Analysis (MCDA)

as outlined by Ledin and Machin (2018) and Van Dijk's (2014) Discourse–Cognition–Society model. This approach reveals underlying ideological, cultural, and gendered assumptions in news coverage.

Telegram is the most popular platform in Iran (Akbari & Gabdulhakov, 2019) where all the official media outlets own channels and are the primary news sources for many people. The analysis examines Persian-language Telegram news outlets affiliated with official Iranian news channels or agencies that have at least 1000 members. The time frame is from August 2020 to January 2026. Sources are 12 major Iranian news Telegram channels, including: ISNA, IRNA, Tasnim, Fars, Mashregh, Etemad Online, Ham Mihan, Mehr, Vatan Emrooz, Alef, Asre Iran, and Bashgahe Khabarnegarane Javan (YJC). Due to the current situation in Iran and the Islamic Republic's humanitarian concerns, websites ending in ir are not accessible from most countries outside Iran, including via VPNs; therefore, studying Telegram is the best research design.

I have used case-based tracking through targeted searches of the channel archives, combined with keywords such as “victim”, “sexual harassment”, and “accusation”. After omitting irrelevant news, I now have a sample of 280 posts. The sample will be examined for:

- Lexical choice and evaluative language surrounding victims and perpetrators, including the presence of any global or region-specific themes
- Narrative structures (e.g., heroisation, moralisation, silencing)
- Visual framing (where relevant, including images, captions, and headlines)
- Any connections to the global #MeToo discourse.

By comparing these cases, the study assesses whether and how Islamic Republic media coverage evolved over time, either progressing toward greater openness or maintaining repression in reporting on gendered violence and accountability. The data-gathering timespan for each channel starts with the first mention of each case study, which differs across channels.

Results

Preliminary analysis suggests that Iranian media discourse around #MeToo remains deeply entangled with patriarchal and moralistic frameworks, though subtle discursive shifts are observable over time. Coverage of Keyvan Imam Verdi's reports was initially limited and episodic, largely avoiding systemic critiques. The case was covered very late in many channels (late 2021-2022) with 31 posts across 12 Telegram channels. The focus remained on individual deviance, while victims' testimonies were filtered through moral caution.

Many media outlets largely ignored Mohsen Namjoo's case in 2021, while Western media outlets covered it widely. Only 4 news channels covered his story, while his name was mentioned for other purposes, such as his perceived unpopularity among

the Iranian diasporas. In contrast, the Pejman Jamshidi case (2025) received unbelievable attention, with increased discussion of its details despite continued constraints in state media. From the total of 280, 236 posts covered Jamshidi's story. Although the case remains unclear and multiple narratives persist, most media either supported Jamshidi through Himpathy (Manne, 2017) – sympathising with the male perpetrator-- or claiming that this is only an allegation to taint a respected person's reputation.

Jamshidi's case is also intertwined with a very interesting occurrence: it was used as an excuse for an official warning to Ham Mihan Newspaper, which published an article on the feminist aspect of the case. The website was shut down by the government after what was claimed to be an inappropriate image (DeFFI, 2025). Ham Mihan was also the newspaper that reported on the death of Mahsa Amini in 2022, which created the Woman, Life, Freedom movement. The broad coverage of the case suggests a gradual normalisation of #MeToo discourse and its growing acceptance in the media. However, the treatment of more feminist outsources reveals the persistence of systematic misogyny and oppression.

Future Work

This is the first study to systematically examine how Iranian news media report on #MeToo-related cases, providing new insights into gender, media, and culture in restrictive environments. Previous research has focused on #MeToo content from Persian social media users, but institutional news coverage remains unexamined. This gap is important because Iranian media operate under strict regulations shaped by state ideology, cultural norms, and self-censorship (RSF, 2023). Analysing coverage of cases such as Keyvan Imam Verdi, Mohsen Namjoo, and Pejman Jamshidi deepens our understanding of the relationship between power, gender, and media responsibility in a semi-authoritarian context.

Future research can analyse all Iranian-based news websites to compare their coverage with that of diaspora and international Persian-language outlets. I have previously compiled a list of active Telegram channels with more than 1000 members across different political orientations, both inside and outside Iran. So far, 105 channels have been identified. Doing a Network analysis of the case studies throughout all Persian-language Telegram News resources provides a better understanding of the dynamics of how each actor and the news outlet operate. This approach will strengthen research by incorporating authentic local perspectives and can provide a very clear picture of the overall news algorithms, censorship, and other possible themes.

Comparing Iran's media coverage with that of other Muslim-majority countries, such as Pakistan, Bangladesh, or Indonesia, would place Iran's approach in a broader Islamic and cultural context. This can help identify similar patterns, clarifying whether Iran's

response is unique or part of a regional or religious trend. Additionally, analysing discussions on platforms like Telegram would show how audiences engage with and reinterpret #MeToo narratives, connecting official reporting to everyday conversations.

References

- Abraham, S. (2021). Crafting a pedagogical third space in a transnational teacher education project. *Teaching and Teacher Education*, 97, 103207. <https://doi.org/10.1016/j.tate.2020.103207>
- Akbari, A., & Gabdulhakov, R. (2019). Platform surveillance and resistance in Iran and Russia: The case of Telegram. *Surveillance & Society*, 17(1/2), 223–231.
- Basmechi, F. (2023). *The Iranian #MeToo Movement* (I. Shagufta Cheema, Ed.). Oxford University Press. <https://doi.org/10.1093/oso/9780197619872.003.0015>
- Bongiorno, R., Langbroek, C., Bain, P. G., Ting, M., & Ryan, M. K. (2020). Why Women Are Blamed for Being Sexually Harassed: The Effects of Empathy for Female Victims and Male Perpetrators. *Psychology of Women Quarterly*, 44(1), 11–27. <https://doi.org/10.1177/0361684319868730>
- Burke, T. (2026). Tarana Burke and chief vision officer. Me Too. <https://metoomvmt.org/get-to-know-us/tarana-burke-founder/>
- DeFFI. (2025). The Website of Ham-Mihan Newspaper Has Been Shut Down. Defending Free Flow of Information in Iran (DeFFI).
- Foomani, E. (Ellie) M., & Reimer, K. E. (2025). Beyond #MeToo in Iran: Change through informal feminist educational leadership. *Journal of Educational Administration and History*, 57(2), 176–193. <https://doi.org/10.1080/00220620.2024.2383414>
- Gholami, N. (2025, October 25). Pejman Jamshidi's case: Charges of "rape" and pressure on victims. DW.COM. <https://amp.dw.com/fa-ir/-پرونده-پژمان-جمشیدی-بحث-اتهام-تجاوز-و-فشار-بر-قربانیان/a-74461978>
- Ghosh, S., Su, M.-H., Abhishek, A., Suk, J., Tong, C., Kamath, K., Hills, O., Correa, T., Garlough, C., Borah, P., & Shah, D. (2022). Covering #MeToo across the News Spectrum: Political Accusation and Public Events as Drivers of Press Attention. *The International Journal of Press/Politics*, 27(1), 158–185. <https://doi.org/10.1177/1940161220968081>
- Gill, G. & Imran Rahman-Jones. (2020, July 9). Me Too founder Tarana Burke: Movement is not over. BBC. <https://www.bbc.com/news/newsbeat-53269751>
- Goodrich, N. H. (2025). Propaganda in the Islamic Republic of. Propaganda in the Digital Age: Global Conflicts, Media, Politics, and Fake News.

- Kamrava, M. (2024). From misogyny to security: Women and the state in Iran. *Third World Quarterly*, 45(13), 1947–1962.
- Lee, B.-H. (2018). #Me Too Movement; It Is Time That We All Act and Participate in Transformation. *Psychiatry Investigation*, 15(5), 433–433.
<https://doi.org/10.30773/pi.2018.04.30>
- Lucarini, A., Suitner, C., Brown, R., Craig, M. A., Knowles, E. D., & Salvador Casara, B. G. (2020). The #MeTooLate Effect: Victim blame and trust denial for sexual harassment not immediately reported. *Personality and Individual Differences*, 167, 110240. <https://doi.org/10.1016/j.paid.2020.110240>
- Manikonda, L., Beigi, G., Kambhampati, S., Liu, H., Hyder, A., Dancy, C., Bisgin, H., & Thomson, R. (2018). Metoo Through the Lens of Social Media. In *LECT NOTES COMPUT SC* (Vol. 10899, pp. 104–110). Springer International Publishing AG.
https://doi.org/10.1007/978-3-319-93372-6_13
- Manne, K. (2017). *Down Girl: The Logic of Misogyny*. Oxford University Press, Incorporated.
<http://ebookcentral.proquest.com/lib/qut/detail.action?docID=5061698>
- Namjoo, M. (2025). Mohsen Namjoo. Mohsen Namjoo.
<https://www.mohsennamjoo.com/bio>
- Noetzel, S., Mussalem Gentile, M. F., Lowery, G., Zemanova, S., Lecheler, S., & Peter, C. (2023). Social campaigns to social change? Sexual violence framing in U.S. news before and after #metoo. *Journalism* (London, England), 24(6), 1232–1262.
<https://doi.org/10.1177/14648849211056386>
- O'Halloran, O., & Cook, N. (2024). Breaking the silence: Exploring women's experiences of participating in the #MeToo movement. *Feminist Media Studies*, 24(6), 1371–1385. <https://doi.org/10.1080/14680777.2023.2231656>
- Peleg-Koriat, I., & Klar-Chalamish, C. (2020). The #MeToo movement and restorative justice: Exploring the views of the public. *Contemporary Justice Review : CJR*, 23(3), 239–260. <https://doi.org/10.1080/10282580.2020.1783257>
- Rahimi, B. (2015). Censorship and the Islamic Republic: Two modes of regulatory measures for media in Iran. *The Middle East Journal*, 69(3), 358–378.
- Rezaei, R. (2022). #MeToo Branded a “Zionist Movement” in High-Profile Tehran Rape Trial. *Iranwire*. <https://iranwire.com/en/women/71035/>
- RSF. (2023). Iran has arrested 79 journalists in a staggering crackdown since Mahsa Amini's death a year ago. *Reporters Without Borders (RSF)*. <https://rsf.org/en/iran-has-arrested-79-journalists-staggering-crackdown-mahsa-amini-s-death-year-ago>

- Serisier, T. (2020). #MeToo and the Politics of Social Change; #MeToo, Weinstein and Feminism; #MeToo: Stories From the Australian Movement: edited by Bianca Fileborn and Rachel Loney-Howes, Switzerland, Palgrave Macmillan, 2019, 350 pp., €29.11, ISBN 978-3-030-15213-0. *Australian Feminist Studies*, 35(105), 310–313. <https://doi.org/10.1080/08164649.2020.1843135>
- Shoaf, H., Pierce, H., & López, J. L. (2024). Witches, Victims, and Villains: #MeToo and the Political Polarization of Sexual Violence. *Violence Against Women*, 30(8), 1910–1933. <https://doi.org/10.1177/10778012241239949>
- Starkey, J. C., Koerber, A., Sternadori, M., & Pitchford, B. (2019). #MeToo Goes Global: Media Framing of Silence Breakers in Four National Settings. *Journal of Communication Inquiry*, 43(4), 437–461. <https://doi.org/10.1177/0196859919865254>
- The Week News Desk. (2025, October 23). Who is Pejman Jamshidi? Iranian footballer-turned-actor arrested for rape and kidnapping. <https://www.theweek.in/news/middle-east/2025/10/23/who-is-pejman-jamshidi-iranian-footballer-turned-actor-arrested-for-rape-and-kidnapping.html>
- Van Dijk, T. A. (2014). 'Critical Discourse Studies: A Sociocognitive Approach.' In *Methods for Critical Discourse Analysis*, ed. By Ruth Wodak & Michael Meyer, 62–86, selected 62–67, 75–80. London: Sage. In *The Discourse Studies Reader: Main currents in theory and analysis* (pp. 389–399). John Benjamins Publishing Company.
- Yaghoobi, C., Mincheva, D., Hooman, N., Hamidi, Y., Rezai, Y., & Moradian, M. (2024). The #MeToo Movement in Iran: Reporting Sexual Violence and Harassment. *Feminist Formations*, 36(1), 106–119. <https://doi.org/10.1353/ff.2024.a928332>
-

Session 3E - **PANEL: Meta Content Library** **Community Roundtable: Where to for Social** **Media Research?**

Day 1 · Mon 13 July · 14:00–15:00 · Seminar Suite B, ARC

Axel Bruns, Laura Vodden (Queensland University of Technology), Fabio Giglietto (Università di Urbino Carlo Bo), Jo Lukito (Digital Democracy Center), Katrin Weller (Heinrich-Heine-Universität Düsseldorf)

This panel brings together a number of leading social media researchers to discuss their experiences and frustrations in using the Meta Content Library (MCL). It forms part of an ongoing effort to build an academic MCL user community and facilitate the sharing of information, knowledge, and collaborative research across institutions – independent of Meta and oriented towards public-interest research.

Research access to social media data from Meta’s platforms has changed substantially in recent years. Meta decommissioned the previous data access service CrowdTangle in August 2024, replacing it with the Meta Content Library (MCL). In its first iteration, provided through the Inter-University Consortium for Political and Social Research (ICPSR) at the University of Michigan, users conducted their analysis within a so-called ‘Virtual Digital Enclave’ (VDE), accessed as a virtual machine within a virtual machine. In a second iteration, introduced in November 2025, accredited researchers can now access a ‘Secure Research Environment’ (SRE), hosted directly by Meta, via their Web browser. In either version, they can then run Python or R code within a Jupyter Notebook that connects to the MCL databases; function calls for doing so differ, remain subject to change, and are insufficiently documented.

The MCL’s impermanence, unreliability, intransparency, and lack of documentation create a clear and pressing need for accredited academic researchers to organise, in order to share information, knowledge, experience, and code as well as develop a shared and distributed research agenda. The early and leading MCL users on this roundtable will present and discuss their own experiences with this service, and highlight key workarounds for its central flaws. We invite other MCL users to join our discussion and community, and hope to continue this scholarly community engagement across future conferences.

Following the decommissioning of CrowdTangle, the Meta Content Library is now the central source of digital trace data for Meta platforms including Facebook, Instagram, and Threads: platforms which – despite the growing number of alternatives that have emerged in recent years – remain amongst the largest and most influential social media

spaces in many countries. Further, in addition to its role as a data source for research on these platforms, the MCL model may also serve as a template for similar data access schemes to be offered by other platforms.

It is therefore important for the social media research community both to critique the overall MCL model – its overall accreditation and access frameworks, built-in assumptions about the research interests it claims to cater to, and operationalisation of user privacy concerns as a convenient reason for limiting research functionality – and to offer more communal support and knowledge-sharing to those researchers who strive to make sense of the often ill-documented and inconsistent functionality of the MCL and create robust and replicable workflows for its use in day-to-day research projects.

This panel continues this process, by building a collective understanding of how to understand the MCL's features and limitations, how to access and use the MCL, and how to navigate Meta's rules for generating scholarly presentations and publications from this work. Our panellists will present their research teams' experience with MCL accreditation and access, discuss the relative benefits and shortcomings of both the VDE and SRE versions of the MCL, outline how they work around common obstacles in MCL functionality (such as the monthly deletion of all data from the SRE), and present a further outlook on how they intend to shape their future research to incorporate the research opportunities that MCL access provides.

Finally, beyond its specific focus on the Meta Content Library as an important but flawed solution for authorised access to digital trace data from a single platform provider, and looking instead towards possible future developments, a broader question for and from the panel is what shape research data access frameworks that are less dependent on the whims and goodwill of individual platform providers might take, and what political and technical structures are needed to support them.

Session 4A - Platform Governance and Regulation (1)

Day 1 · Mon 13 July · 15:15–16:15 RD Suite, ARC

4A.1 Decentered Platform Governance: Learning trust and Safety Practices with Communities on Decentralised Social Media Platforms

Ashwin Nagappa, Daniel Angus, Kateryna Kasianenko (Queensland University of Technology)

Introduction

Although the World Wide Web has made digital communication seamless and fast, governance mechanisms for such communication are highly complex. The 'democratic utopia' (Papacharissi, 2004) hyped by early web advocates never materialised. Instead, a rise in online harm has threatened social, political, and financial institutions globally (Vasist et al., 2023; Zahidi, 2023). Trust and safety debates among various stakeholders mostly focus on a few major platforms, including Google, Amazon, Facebook, YouTube, X, and Reddit (Roth & Lai, 2024; Gorwa, 2019).

Addressing online harm on mainstream platforms remains complex. Regulatory measures, like the EU Digital Services Act (DSA), are important steps towards making platform companies accountable (Bartolo & Matamoros-Fernández, 2023). Yet, achieving consensus among global stakeholders on definitions for platform governance is challenging (Thorpe, 2023). While Fukuyama (2021) suggests moving governance away from platform companies, Faris & Donovan (2021) contend that large platforms are essential to sustain large-scale trust and safety measures.

Outside these dominant platforms, the web enables non-institutional actors to foster healthy participation and safer communication. For over a decade, collectives have created alternative social media (ASM) platforms to address the drawbacks of big platforms. Facebook's 2014 'real name' policy, which limited pseudonymity and impacted marginalised communities, and concerns about data extraction from personalised ads, inspired platforms like Ello, Diaspora, GNU Social, and Twister (Gehl, 2017; Niedt, 2016; Salisbury & Pooley, 2017; van Dijck et al., 2018).

Historically, these alternative solutions began as marginal experiments. After brief popularity, they could not manage financially or prevent problematic content as they

scaled up (Gehl, 2017). Some newer platforms, like Gab, have grown substantial user bases but also fostered misinformation and polarisation, creating associations between ‘alternative social media’ and ‘alt-right social media’ (Gehl & Zulli, 2022). Despite these challenges, many communities continue working toward prosocial alternatives, with renewed interest in decentralised (Web3) and AI-driven governance. Recently, Twitter's drastic change to X drove users to ASM platforms like Mastodon, seeking to shift governance from corporations to user communities.

In January 2026, Mastodon had over 12 million users (Fedidb.Com/Stats_, n.d.), while Bluesky had about 42 million accounts (Stats for Bluesky by Jaz (Jaz.Bsky.Social), n.d.). These projects aim for polycentric governance—balancing free speech, trust, and safety through multiple autonomous decision centres (Carlisle & Gruby, 2019). However, these smaller platforms remain understudied, lack transparency requirements, and are often overshadowed by larger platforms.

This work-in-progress paper shares early findings gathered with diverse stakeholders on decentralised social media platforms.

Objectives/Research Questions

The decline of Twitter and the rise of decentralised networks have underscored the need for research into alternative governance structures (Roth & Lai, 2024). Recent research around decentralised social media platforms has focused on polycentric governance models in digital communication. Hwang et al. (2025) highlight “governance frictions”—privacy risks arising from incompatibilities in values, security practices, and software across communities. In a workshop with users and administrators of decentralised platforms, Hwang et al. (2025) explored aspirations for sustaining prosocial social media ecosystems. Kissane & Kazemi (2024) found, through qualitative interviews with Fediverse administrators, that governance in the Fediverse is diverse and highly localised. They emphasise that, rather than a universal system, there are many governance models shaped by community goals, server size, norms, and resources. While these studies illuminate critical issues, their small sample sizes (fewer than 30 participants) indicate that further research is needed to examine the key characteristics of prosocial platforms. Moreover, there are means to invite more stakeholders from diverse communities and to capture their perspectives on creating and sustaining prosocial platforms.

This paper identifies the characteristics of polycentric governance and the potential influence of innovations in AI on decentralised platforms. It uncovers how these characteristics shape prosocial digital practices. It does so by gathering insights from diverse stakeholders who own, operate, and use decentralised social media platforms. This paper primarily focuses on prominent community-driven Fediverse platforms such as Mastodon, Lemmy, and Pixelfed, as well as on commercially funded alternatives such as Bluesky. However, the scope is not limited to these examples; it will include

any decentralised social media platform that aligns with the research objectives. The decentralised ecosystem is highly dynamic, with new platforms emerging and older ones evolving or fading over time, making it difficult to compile a comprehensive and definitive list. To maintain the integrity and safety of the research, platforms known for hosting extreme or highly problematic content—such as 4chan or similar spaces—will be screened out and excluded from analysis.

Methods

The ongoing efforts to develop a prosocial, decentralised social network that encourages free and safe participation provide opportunities to understand the strengths and limitations of such an online communication system. Simultaneously, these decentralised platforms are tackling public policy issues which are considered 'wicked problems' (Rittel & Webber, 1973). Therefore, absolute solutions to online harm may not be achievable. Nonetheless, the repeated failures and challenges faced by prosocial alternative approaches indicate that they can be effective only under specific circumstances. Consequently, any solution must involve a careful review of current platform governance strategies and emerging future directions. Addressing these wicked problems requires innovative methods for data collection and analysis that can encompass diverse, sometimes conflicting viewpoints of stakeholders involved in platform governance within the Fediverse.

Therefore, this paper takes a novel approach by applying group concept mapping (GCM) (Novak & Gowin, 1984; Conceição et al., 2017) to digital media, and we are currently engaged in this collaborative phase. While concept mapping is widely used in fields such as health and psychology, its application to Internet and social media research has been limited (see McLinden, 2017; Moreno et al., 2013). As McLinden (2017, p. 293) notes, a 'useful and safer approach is to start in the recent past and explore the maturation of concept mapping as a means of pointing toward future possibilities,' particularly in addressing 'wicked problems in communities.' Using concept mapping as a key method, we invite stakeholders (including, but not limited to, decentralised social media instance owners, instance owners, and users of decentralised social media platforms) to contribute their ideas from February to April 2026. At the time of writing this abstract, the research team is recruiting participants.

In the first part, stakeholders are invited to brainstorm on the GroupWisdom platform by providing open-ended responses about their perspectives on prosocial platforms. From April to June 2026, the research team, alongside stakeholders, will identify and refine these themes through a participatory process that captures diverse perspectives and encourages collaborative problem-solving. Stakeholders will play a central role in ranking and interpreting the concept map. Participants' scores and feedback will guide the development of a revised concept map, which will serve as the basis for creating an online safety and trust framework, policy guidelines, and academic publications. By

sharing their insights, participants help shape practical solutions for safer, more inclusive decentralised platforms, ensuring that the outcomes reflect real-world challenges and community priorities. This collaborative process offers valuable insights that might otherwise be difficult to obtain under logistical and operational pressures.

Results

As indicated in the previous section, this paper shares preliminary findings from the group concept mapping involving stakeholders from prosocial decentralised platforms. The research design and data collection nuances were approved by the University Human Research Ethics Committee (UHREC) in December 2026. The brainstorming, concept development, sorting, and ranking will take place between February 2026 and June 2026. This work-in-progress paper will present insights from applying the Group concept mapping to a social media setting and the preliminary findings from brainstorming, which leads us to develop concepts around the governance of prosocial, decentralised platforms. It will also share early insights from the GCM ranking process. These early findings will also be used to develop questions for semi-structured interviews with willing participants to gain further insight into the key concepts and ideas that emerge in the GCM process.

Future Work

Findings from the GCM will guide a follow-up study using semi-structured interviews to gain deeper qualitative insights into the key themes identified through the mapping process. While GCM is effective for capturing and organising perspectives from diverse stakeholder groups, it provides limited opportunities for in-depth exploration of individual experiences and rationales. The subsequent interviews, therefore, address this by allowing for a more detailed examination of selected issues. Collectively, insights from the GCM and semi-structured interviews will inform community-facing reports, academic publications, and policy guidelines aimed at fostering a prosocial decentralised digital media ecosystem.

References

- Bartolo, L., & Matamoros-Fernández, A. (2023). Online harm. In M. Khnan & B. Yang (Eds.), *ISP-platform governance terminologies essay series*. Yale Law School.
- Carlisle, K., & Gruby, R. L. (2019). Polycentric systems of governance: A theoretical model for the commons. *Policy Studies Journal*, 47(4), 927–952.
<https://doi.org/10.1111/psj.12212>
- Conceição, S. C. O., Samuel, A., & Biniecki, S. M. Y. (2017). Using concept mapping as a tool for conducting research: An analysis of three approaches. *Cogent Social Sciences*, 3(1), Article 1404753. <https://doi.org/10.1080/23311886.2017.1404753>

- Faris, R., & Donovan, J. (2021). The future of platform power: Quarantining misinformation. *Journal of Democracy*, 32(3), 152–156.
- Fedidb.com/stats. (n.d.). Statistics. Retrieved November 17, 2025, from <https://fedidb.com/stats>
- Fukuyama, F. (2021). Making the internet safe for democracy. *Journal of Democracy*, 32(2), 37–44.
- Gehl, R. W. (2017). Alternative social media: From critique to code. In J. Burgess, A. Marwick, & T. Poell (Eds.), *The SAGE handbook of social media* (pp. 330–349). SAGE Publications.
- Gehl, R. W., & Zulli, D. (2022). The digital covenant: Non-centralized platform governance on the Mastodon social network. *Information, Communication & Society*, 0(0), 1–17. <https://doi.org/10.1080/1369118X.2022.2147400>
- Gorwa, R. (2019). What is platform governance? *Information, Communication & Society*, 22(6), 854–871. <https://doi.org/10.1080/1369118X.2019.1574914>
- Hwang, S., Nanayakkara, P., & Shvartzshnaider, Y. (2025). Trust and friction: Negotiating how information flows through decentralized social media. *Proceedings of the ACM on Human-Computer Interaction*, 9(CSCW), Article 335. <https://doi.org/10.1145/3757516>
- Hwang, S., Rollins, S., Nunes, T. A., Liu, Y., Wong, R., Shaw, A., & Monroy-Hernández, A. (2025). Governing together: Toward infrastructure for community-run social media (arXiv:2509.19653). arXiv. <https://doi.org/10.48550/arXiv.2509.19653>
- Kissane, E., & Kazemi, D. (2024). Governance on Fediverse microblogging servers. <https://fediverse-governance.github.io/images/fediverse-governance.pdf>
- Liu, A. (2012, November 4). What happened to Diaspora, the 'Facebook killer'? It's complicated. Motherboard.
- McLinden, D. (2017). And then the internet happened: Thoughts on the future of concept mapping. *Evaluation and Program Planning*, 60, 293–300.
- Moreno, M. A., Kota, R., Schoohs, S., & Whitehill, J. M. (2013). The Facebook influence model: A concept mapping approach. *Cyberpsychology, Behavior, and Social Networking*, 16(7), 504–511. <https://doi.org/10.1089/cyber.2013.0025>
- Niedt, G. (2016). Social media affordances and the capital of queer self-expression: Facebook, Ello, and the Nymwars. In A. Novak & I. J. El-Burki (Eds.), *Defining identity and the changing scope of culture in the digital age* (pp. 99–116). IGI Global.

- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge University Press.
<https://doi.org/10.1017/CBO9781139173469>
- Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, 6(2), 259–283.
- Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4(2), 155–169.
- Roth, Y., & Lai, S. (2024). Securing federated platforms: Collective risks and responses. *Journal of Online Trust and Safety*, 2(2), Article 2.
<https://doi.org/10.54501/jots.v2i2.171>
- Salisbury, M., & Pooley, J. D. (2017). The #nofilter self: The contest for authenticity among social networking sites, 2002–2016. *Social Sciences*, 6(1), Article 10.
<https://doi.org/10.3390/socsci6010010>
- Stats for Bluesky by Jaz. (n.d.). Retrieved November 17, 2025, from
<https://bsky.jazco.dev/stats>
- Thorpe, A. (2023, June 26). Australia vs the internet: What can the government do to fight online misinformation? ABC News. <https://www.abc.net.au/news/2023-06-27/misinformation-disinformation-australia-proposed-laws-crackdown/102524764>
- van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society*. Oxford University Press. <https://doi.org/10.1093/oso/9780190889760.001.0001>
- Vasist, P. N., Chatterjee, D., & Krishnan, S. (2023). The polarizing impact of political disinformation and hate speech: A cross-country configural narrative. *Information Systems Frontiers*, 26, 1–26.
- Zahidi, S. (2023, January 12). We can't solve the world's biggest crises without protecting our basic needs. TIME. <https://time.com/6246774/basic-needs-davos-2023/>

4A.2 Trust Café: A Marxist-feminist approach to digital commons and content moderation

Dayei Oh (University of Strathclyde), Craig Ryder (Ludwig Maximilian University of Munich)

Introduction

Content moderation is a process that underwrites the platform ecosystem (Gillespie, 2018, Roberts, 2019). The opaque mechanisms that sift through, elevate, downplay and remove content are the de facto intellectual property of a platform, shaping its communities, demonstrating its politics and determining its profits (Gorwa, 2019). Major platform content moderation is outsourced to workers in the Global South (Ahmad & Krzywdzinski, 2022), but due to human limitations (Schöpke-Gonzalez, 2022) and financial imperatives, it is increasingly delegated to non-human systems (Oh & Downey, 2025). The opacity of content moderation practices contributes to the phenomenon of “shadowbanning” (Cotter, 2023), where the content removed or downvoted by platforms is interpreted as a political move to silence specific users and communities, creating “algorithmic folklore” (Savolainen, 2022).

The scale and speed of moderation on major platforms raises important questions, not only about the technicalities of moderation systems but also fundamentally about the normative role of moderation. In mainstream discourse, including governmental and corporate justifications, content moderation is largely framed as deontological rigid rule-following to maintain legal compliance or as utilitarian customer-satisfaction practices under digital capitalism (Fuchs, 2021; Oh, 2025; Suzor, 2019). Problems arise when such legal and business practices work against the rights and safety of marginalized communities. For example, platforms may abide by intolerant laws of authoritarian regimes to stay in business and algorithmic moderation may remove marginalized activist content without fully transparent mechanisms for appeal (Rauchberg, 2022; York, 2022). State legislative pressures, such as the UK Online Safety Bill, further incentivize platforms to act faster and at greater scale, where mistakes and harms caused by improper moderation can only be remedied retroactively and without full transparency. This approach resembles Mark Zuckerberg and Facebook’s old motto of “Move fast and break things,” emphasizing speed over caution, while failing to acknowledge that the harms of improper moderation may be unevenly distributed.

Such a conventional framework is at odds with the utopian vision of an early internet landscape as digital commons (Fuchs, 2021), characterized by collaborative communities based on decentralized and nonpropriety models of participation. There are, however, reactionary projects emerging that attempt to redress the uneven power

asymmetries of the systems that moderate through entrenched biases and opaque bureaucracies.

Objectives/Research Questions

To explore an alternative model for content moderation, this article delves into the emerging pro-social platform, Trust Café. Founded in the UK by Jimmy Wales, co-founder of Wikipedia, Trust Café has a highly specific approach to content moderation that foregrounds human collaboration over automated processes of governance. The unique affordance of Trust Café is that every user is a moderator. Thus, rather than “liking” or “retweeting” peer content, Trust Café users award each other “trust points” that contribute to the visibility and reach of the peer’s post. Whilst still in Beta mode, with approximately 100,000 registered users—and far fewer active users—Trust Café is reimagining a platform experience that resists labour outsourcing and algorithmic sorting by making content moderation an intrinsic component of the user experience.

By reestablishing a digital commons via a reimagining of the normative role of content moderation, the article examines the extent to which Trust Café embodies a Marxist-feminist approach to digital ethics (Oh, 2025). A Marxist-feminist approach to digital ethics combines the Marxian emphasis on digitalization that fosters the common good and emancipation rather than profit (Fuchs, 2022) with intersectional-feminism which stresses the importance of situated knowledge and epistemic justice for marginalized communities in digital infrastructures (D’Ignazio & Klein, 2023). From a Marxist-feminist perspective, content moderation is not primarily about complying with state laws or maintaining majoritarian customer satisfaction. Its purpose is to ensure that the digital public sphere is open, inclusive, and safer for all, especially for those lacking alternative means to voice their demands and influence national discourse.

Methods

Based on nine months of augmented ethnography (Ryder, 2024) among Trust Café content moderators, the article investigates the platform’s content moderation processes vis-a-vis the platform’s commitment to inclusive user participation and freedom of speech. After first contact with the Trust Café Community Manager, Bobby, we were invited to a private Discord server to discuss the project goals more in-depth. Once proper terms of engagement were agreed, Bobby, announced on the platform’s main “branch” (a Trust Café technical feature similar to a subreddit) that an “anthropologist was amongst us”—and that we will be observing the platform’s moderation system and the sociality between moderators and users. Users were told the ethnographers may interact with your content and reach out for interviews, and that you have no obligation to participate and can opt out if you choose. As such, an

ethnographic study that complied with the ethical best practice of informed consent was established (AOIR, 2019).

Since March 2025, we have been a consistent presence on both Trust Café and the related Discord server. As an active content moderation volunteer (“volly”), we engaged in moderation discussions, observed decision-making, and conducted interviews with 25 individual users and the core staffers by video calls and text-based platform discussions.

Results

Three recurring themes form the basis of our preliminary findings.

First, content moderation on Trust Café is performed by people that use their diverse intersectionality as an instrument of knowing what is, and what is not, to be permitted on the platform. The two lead community managers of Trust Café both identify as trans/non-binary and actively bring their experience of queer identity to the procedure of content moderation. A Marxist-feminist approach to content moderation suggests that whether algorithmic or human-moderator-centred, moderation processes must incorporate bottom-up voices, particularly from marginalized groups in vulnerable positions under current political climates, including LGBTQ+ individuals and racial, ethnic, and religious minorities. While the moderators and users are overwhelmingly White, being based predominantly in the UK, USA, Northern Europe, there is rich intergenerational and multigender diversity. Ethnographically, we observed multiple occasions when transphobic content or deliberate pronoun slurs resulted in swift user removal, because, as Jess, an American community manager, would say, “screw that”.

Second, human-centered content moderation on Trust Café demands an ongoing debate about digital ethics that reinforces an “us versus them” dictum that situates Trust Café and their ethical approach to content moderation as distinguishable to major platforms. When asked about what makes Trust Café distinct, Skin, the platform’s lead tech-developer, said they think “Zuckerberg should be in Le Hague!” They claim that Trust Café is the only platform that seems to admit that content moderation is an extremely nuanced process that has serious legal, ethical and health ramifications. Skin accuses Meta platforms of criminality whose CEO should be imprisoned for his quest of platform capitalism that has equated to serious negligence around content moderation. Simultaneously, they elevate Trust Café as an ontological panacea to the problem of content moderation.

Third, there is recurring tension between a utopian human process of content moderation versus the adoption of automated systems. Not unlike Wikipedia, which is famously humanly moderated but increasingly prone to “various data-driven algorithmic systems which Wikipedians rely on to help manage and govern the “anyone can edit” encyclopedia at a massive scale” (Geiger, 2017, 1), Trust Café moderators are

prone to adopting AI for seemingly banal and bizarre use cases. For example, when moderating a user suspected of posting conspiratorial content, a Swedish moderator confesses to using ChatGPT to draft her dialogue with the user so she can navigate the situation with extra sensitivity. The use of ChatGPT in the moderation procedure appears to complicate Trust Café's vision for a human moderation at scale, highlighting the inherent contradictions—or almost unavoidable “heteromation” (Ekbia & Nardi, 2025)—at play in contemporary digital and non-digital experiences.

Future Work

This study opens several avenues for future research into pro-social models of platform governance, content moderation, and their implications for the digital public sphere and the digital commons (Fuchs, 2021).

First, while this article focuses on Trust Café as a single case study, comparative research with other small alternative platforms (e.g., fediverse) would help further articulate what a genuinely Marxist-feminist approach to content moderation (Oh, 2025) might look like beyond the logics of legal compliance and profit. Comparative studies can also examine whether Trust Café's approach (slower, deliberative, and collectively governed models) represents a broader shift in social imaginaries of non-AI-assisted moderation.

Second, longitudinal studies are required. Trust Café is still in an early stage, and longitudinal research could trace how moderation norms, ethical commitments, and internal conflicts evolve as the platform stabilizes and expands.

Third, future research would benefit from case studies outside the “West.” While this article foregrounds power inequalities and marginalized identities within moderation decision-making, we acknowledge that the Trust Café's user and moderation team base is predominantly White and from the Global North. Extending ethnographic inquiry to non-Western or multilingual platforms would strengthen a Marxist-feminist analysis of epistemic justice and digital commons.

References

- Ahmad, S, & Krzywdzinski, M. (2022) ‘Moderating in obscurity: How Indian content moderators work in global content moderation value chains’. In M. Graham & F. Ferrari (Eds.), *Digital work in the planetary market* (pp. 77–95). The MIT Press International Development Research Centre
- AoIR. (2019) Internet Research: Ethical Guidelines 3.0 Association of Internet Researchers, [AoIR.org](https://aoir.org/).
- Cotter, K. (2023) ‘Shadowbanning is not a thing’: black box gaslighting and the power to independently know and credibly critique algorithms’, *Information*,

- Communication & Society, 26:6. pp. 1226-1243. doi:
10.1080/1369118X.2021.1994624
- D'Ignazio, C and Klein, L. (2023) Data Feminism. The MIT Press ·
- Ekbja, H & Nardi, B. (2025) Heteromation, and Other Stories of Computation and Capitalism, MIT Press
- Fuchs, C (2021). 'The Digital Commons and the Digital Public Sphere: how to advance digital democracy today'. Westminster Papers in Communication and Culture 16, 1
- Fuchs, C. (2022) Digital ethics: Media, communication and society. Routledge.
- Geiger, S. (2017) 'Beyond opening up the black box: Investigating the role of algorithmic systems in Wikipedian organizational culture', Big Data & Society July–December 2017: 1–14
- Gillespie, T. (2018) Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media. Yale University Press
- Gorwa, R. (2019) 'The platform governance triangle: conceptualizing the informal regulation of online content'. Internet Policy Review, 8(2). doi:
10.31219/osf.io/tgnrj
- Jereza, R. (2024). "'I'm not this person": Racism, content moderators, and protecting and denying voice online'. New Media & Society, 26(8), 4454–4470.
- Oh, D. (2025, August). From civility to parity: Marxist-feminist ethics for context-aware algorithmic content moderation. In Proceedings of the The 9th Workshop on Online Abuse and Harms (WOAH) (pp. 32-40).
- Oh, D & Downey J. (2025). 'Does algorithmic content moderation promote democratic discourse? Radical democratic critique of toxic language AI'. Information, Communication & Society, 28(7), 1157-1176
- Roberts, S. T. (2019). Behind the screen: Content moderation in the shadows of social media. New Haven, CT: Yale University Press
- Ryder, C. (2024) 'Becoming Augmented: The Latent possibilities of research in a pandemic, 'Becoming augmented: The latent possibilities of ethnography in a pandemic'. Possibility Studies and Society, 3(1), 63-77. DOI
10.1177/27538699241264608
- Schöpke-Gonzalez, A., Atreja, S., Shin, H. N., Ahmed, N., & Hemphill, L. (2022). 'Why do volunteer content moderators quit? Burnout, conflict, and harmful behaviours'. New Media & Society, 26(8), 4454–4470.

- Rauchberg, J. (2022). '#Shadowbanned: Queer, Trans, and Disabled creator responses to algorithmic oppression on TikTok'. In LGBTQ digital cultures (pp. 196-209). Routledge.
- Savolainen, L. (2022), 'The shadow banning controversy: perceived governance and algorithmic folklore', *Media, Culture & Society*, 44(6), pp: 1091–1109, doi: 10.1177/0163443722107717
- Suzor, N (2019). *Lawless: The Secret Rules That Govern Our Digital Lives*. Cambridge University Press
- York, J. C. (2022). *Silicon values: The future of free speech under surveillance capitalism*. Verso Books.

4A.3 Social media ‘back in the hands of the people’?

Nuancing assumptions of legitimate content moderation on Mastodon

Klara Matusiewicz, João C. Magalhães (University of Manchester)

Introduction

The decisions platform companies make to determine what content and behaviour ought to be allowed in their services have never been straightforward (Gillespie, 2018). But in the past ten years, this inherent difficulty has been greatly compounded by a series of extraordinary crises – the re-emergence of authoritarian populism, the COVID-19 pandemic, the wars in Ukraine and Gaza – which reinforced global concerns about platforms’ political role, and called into question their early relatively lax moderation approaches. While, over time, platforms increased the transparency of moderation decisions and rules, deepened their relations with civil society and governments and, above all, ramped up their intervention efforts (Douek, 2021), these moves were not able to solve the issues of moderation. Platform content moderation has come to be perceived as either too strict or not strict enough – a political impossibility that attracted endless controversy and regulatory attention. Taken together, such criticisms contributed to a ‘techlash’ that, we argue, is largely underpinned by a crisis of political legitimacy.

Yet, such crisis has not only shaped how dominant social media work. It has also led to a renewed interest in alternative digital spaces, ranging from ostensibly ‘libertarian’ alt-tech platforms to the so-called non-centralised Web. The appeal of these technically fragmented services is importantly political. Scholars have suggested that decentralising control over online spaces can help make online governance more legitimate, as seeking users’ participation would make its process more inclusive and meaningful (Zuckerman and Rajendra-Nicolucci, 2023). This view is at the heart of the Fediverse – the ensemble of ‘protocols, servers, applications and communities’ behind decentralised social media’ (Rozenshtein, 2023, p. 222). Built as a network of communities (known as ‘instances’) which can link with one another – but without a central authority to directly oversee them – the Fediverse embodies the promise of also decentralising speech governance, as it allows instances’ members to choose their servers’ moderation rules and the individual moderators who will be in charge of enforcing them (Gehl and Zulli, 2023). Such a structure leads its largest software, Mastodon, to brand itself as ‘radically different social media, back in the hands of the people’ (Mastodon, n.d.).

Objectives/Research Questions

Despite the software's flatter materiality and the opportunities for participation that this design might potentially allow for, however, Mastodon instances' rules are still enforced by administrators and moderators. These individuals enjoy, in principle, a more privileged position than regular Mastodon users. They are empowered to decide how other members ought to behave and the content they are exposed to, decisions that are likely to be politically-laden and controversial. As any other form of power, this one calls for reasons and justifications. In other words, Mastodon's design cannot by itself do away with the issue of legitimacy. This raises a critical but hardly studied question: How do Mastodon moderators negotiate their political legitimacy?

Addressing it is important to probe the extent to which these services' promise of a more legitimate form of social media, which rely on strong assumptions about the political consequences of technological design, have or not been fulfilled, and why. To do so, we resort to political theory to define legitimacy, and then apply the concept to the issue of moderation in the Fediverse. We argue that legitimacy hinges, centrally, on the production of consent to power, and that such consent exists at the entanglement of participation (the expression of subordinates' values and interests), representation (the embodiment of subordinates' interests and values), and accountability (including transparency, justifications and the possibility of sanctions) (Coicaud, 2002; Beetham, 2012; Schmidt, 2013; Haggart & Keller, 2021). The relationships between these three aforementioned components, we argue, create a conceptual spectrum with democratic legitimacy at one end, and authoritarian legitimacy at the other (Von Soest & Grauvogel, 2017). We propose that a democratic view of legitimacy sees representation as emerging from participation and accountability, and is dependent on the subordinates' acceptance of authority (Scherz, 2021). An authoritarian view of legitimacy, on the other hand, largely dismisses participation and accountability, and takes representation for granted. Therefore, to understand how moderators negotiate their political legitimacy, we investigate the three constituent variables through the following sub-questions:

1. What role does representation play in how moderation happens within Mastodon instances?
2. What role does participation play in how moderation happens within Mastodon instances?
3. What role does accountability play in how moderation happens within Mastodon instances?

Methods

We conducted semi-structured, in-depth interviews with moderators from Mastodon, currently the largest and most popular software of the Fediverse. In-depth semi-structured interviews were chosen as they help explore social actors' practices and

perceptions, allowing us to gain insights into how Mastodon moderators view and negotiate their legitimacy (Charmaz & Belgrave, 2012).

Participants were selected using purposive sampling: they had to be moderators of a Mastodon instance, and were chosen with a view to diversity. The aim was to include moderators of both general interest and topic-specific instances, and of instances with open, closed or approval-based registration. The participants were recruited through a message on the Mastodon profile of one of the researchers, and by using the Fediverse Observer – an online list of Fediverse instances – to find Mastodon instances and contact their moderators directly. Mastodon's historical focus on privacy meant that many users were sceptical about being researched, and interview requests were often met with silence and refusals. Consequently, our data was eventually sourced from fifteen in-depth interviews.

A combination of inductive and deductive thematic analysis was used to examine the data (Braun & Clarke, 2006). This approach combines using data-driven- and theory-driven codes, allowing us to use the sub-concepts of participation, representation and accountability as general themes to inform the analysis. These theory-driven general themes, in turn, guided the development of data-driven initial and second-order codes. To increase transparency and the quality of our conclusions, the findings were sent via email to interviewees, who were offered the chance to comment on them.

Results

The findings suggest that all moderators negotiate their legitimacy but to different extents and in different ways. Some of the interviewees appeared more invested in seeking legitimacy than others, which indicates that Mastodon's legitimacy mechanisms and regimes are more complex than what its decentralised, instance-based nature might suggest. More specifically, moderators' legitimacy negotiation tactics appear to depend mainly on how these individuals politically perceive their instances – understandings that are associated with particular legitimacy-seeking practices.

Building on our conceptualisation of the politics legitimacy seeking, we contend that these perceptions materialise as two broad types: the view of instances as bottom-up participatory spaces, which broadly aligns with a democratic approach to legitimacy; and the view of instances as top-down governable spaces, closer to an authoritarian take on legitimacy and which appears to be enabled by Mastodon's technological design. Moderators holding the former view perceived their instances as spaces designed for users, and sought their opinions and participation, while also being open to being held accountable. This perception, however, was largely overshadowed by a view of servers as top-down governable spaces. In such cases, moderators' perspectives were often shaped by the visions of instances' creators, who appeared to

see their instances as their property. In this view, instances are like political fiefdoms: users are free to join and leave but should not expect to have a substantive say in the community's rules or how it is governed. Importantly, these are just explanatory models. Rather than a fully democratic versus a fully authoritarian instance dichotomy, multiple shades of legitimate and illegitimate speech governance seem to co-exist in most participants' descriptions – and imaginations – of their instances.

Therefore, while it seems intuitive to say that 'decentralizing control over online spaces can bring legitimacy' as it offers users greater input and more choices (Zuckerman and Rajendra-Nicolucci, 2023, p. 7), it does not by itself return social media to 'the hands of the people' (Mastodon, n.d.). On the scale of the Fediverse or of Mastodon, an approach based on giving users the choice between instances with different rules – which appears to help address the impossible conundrum of moderation – only seems to do so because it devolves highly-charged political issues to the more granular level of individual communities. If that is the case, rather than helping solve the challenges of content moderation or necessarily demonstrating how to moderate in a more legitimate way, our research shows that Mastodon moderators' approach might demonstrate the endurance of the difficulties early online spaces faced when dealing with speech governance (Schneider, 2024).

Future Work

Both Mastodon and the Fediverse are very diverse, and any attempt to put a simplistic, all-encompassing conceptual blanket over them would be inaccurate. The above results may differ depending on the sample under observation, and more research is needed to corroborate such findings. Future studies could investigate the perceptions of legitimacy of Mastodon users to supplement and nuance the moderators' insights this work focused on. Moreover, they could research legitimacy in other spaces within the Fediverse – such as Pleroma or Pixelfed – or investigate legitimacy more broadly, focusing on the actual disputes around how to design and tweak the affordances allowed by the ActivityPub protocol which underlies them all.

References

- Beetham, D. (2012). Political legitimacy. In E. Amenta, K. Nash & A. Scott (Eds.), *The Wiley-Blackwell Companion to Political Sociology* (pp. 120–129). John Wiley & Sons.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
<https://doi.org/10.1191/1478088706qp063oa>

- Charmaz, K., & Belgrave, L. (2012). Qualitative interviewing and grounded theory analysis. In J. Gubrium (Ed.), *The SAGE Handbook of interview research: The complexity of the craft* (pp. 347–365). Sage Publications.
- Coicaud, J.-M. (2002). *Legitimacy and politics: A contribution to the study of political right and political responsibility*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511490200>
- Douek, E. (2021). Governing online speech: From "posts-as-trumps" to proportionality and probability. *Columbia Law Review*, 121(3), 759–834.
- Gehl, R. W., & Zulli, D. (2023). The digital covenant: Non-centralized platform governance on the Mastodon social network. *Information, Communication & Society*, 26(16), 3275–3291. <https://doi.org/10.1080/1369118X.2022.2147400>
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
<https://doi.org/10.12987/9780300235029>
- Haggart, B., & Keller, C. I. (2021). Democratic legitimacy in global platform governance. *Telecommunications Policy*, 45(6), 1–17.
<https://doi.org/10.1016/j.telpol.2021.102152>
- Mastodon. (n.d.). Social networking that's not for sale. <https://joinmastodon.org/>
- Rozenshtein, A. Z. (2023). Moderating the Fediverse: Content moderation on distributed social media. *Journal of Free Speech Law*, 3, 217–236.
- Scherz, A. (2021). Tying legitimacy to political power: Graded legitimacy standards for international institutions. *European Journal of Political Theory*, 20(4), 631–653.
<https://doi.org/10.1177/1474885119838137>
- Schmidt, V. A. (2013). Democracy and legitimacy in the European Union revisited: Input, output and 'throughput'. *Political Studies*, 61(1), 2–22.
<https://doi.org/10.1111/j.1467-9248.2012.00962.x>
- Schneider, N. (2024). *Governable spaces: Democratic design for online life*. University of California Press.
- Von Soest, C., & Grauvogel, J. (2017). Identity, procedures and performance: How authoritarian regimes legitimize their rule. *Contemporary Politics*, 23(3), 287–305.
<https://doi.org/10.1080/13569775.2017.1304319>
- Zuckerman, E., & Rajendra-Nicolucci, C. (2023). From community governance to customer service and back again: Re-examining pre-web models of online governance to address platforms' crisis of legitimacy. *Social Media+Society*, 9(3), 1–12. <https://doi.org/10.1177/20563051231196864>

Session 4B - Digital and Data Methods

Day 1 · Mon 13 July · 15:15–16:15 Seminar Suite A, ARC

4B.1 Stance-taking as a technology of power and the self: A linguistic anthropological approach for analyzing social media interaction

Julian Canjura (The Pennsylvania State University)

Introduction

This paper presents an interdisciplinary conceptual and analytical approach to social media interaction by synthesizing sociolinguistic stance-taking (Du Bois, 2007), Foucauldian social theory (e.g., 1977, 1978), and a linguistic anthropological lens on text and ritual (Agha, 2007; Silverstein, 2022) to conceptualize social media interaction as a technology of the self and a technology of power and government. To illustrate how this approach can be operationalized to analyze power, ideology, and subjectivity on social media, this paper presents a case study where issues of gender, which has become an increasingly problematized and debated online (cf. Bulter, 2024), are discussed and contested in a Facebook Group.

Stance-taking is a public, intersubjective act where social actors evaluate objects, position themselves within a sociocultural field, and (dis)align with others (Du Bois, 2007). While stance-taking is a productive concept for studying identity and power online (e.g., Marino, 2023; Zentz, 2021), traditional models are often based on face-to-face, dyadic interaction, not considering the complexities of social media interaction, and have lacked engagement with larger social theories. To address the later, I turn to Foucault's treatment of power and subjectivity. Particularly, I draw on his notion of subjectification: the ways human beings become subjected to another's dominance and become tied to their own identity through self-knowledge (Foucault, 1982). This is accomplished through technologies of power, which determine and enforce the conduct of individuals, and technologies of the self, through which people come to know what kinds of selves are possible (Foucault, 1988). These notably include the use of scientific discourse to "objectify" aspects of the subject. Together, these technologies enable or discourage certain ways of being. Individuals are led to regulate both themselves and others while becoming increasingly individualized yet increasingly grouped.

Previous work applying Foucauldian concepts to social media has often focused on big data surveillance and algorithmic governance (e.g., Waycott et al., 2017) or online-offline relationships (e.g., Møller, 2023). I expand upon this line of inquiry by arguing that all micro-level interactions (i.e., stance-taking) function as a means of corrective disciplinary training. Through public stance-taking, an essentially voluntary action that is encouraged and rewarded by platforms, users move from unobservable “lurkers” into visible entity that other users may observe and interact with. By documenting their opinions, users submit themselves to a virtual panopticon where they are observable and judgeable by both the platform and other users. Users evaluate and discipline each other’s positions, regulating what subjectivities and opinions are considered permissible or “normal.”

Objectives/Research Questions

The study is guided by two core questions: (1) how, and why, does stance-taking occur on social media (i.e., Facebook Groups), and (2) how might stance-taking circulate, manifest, perpetuate, and/or challenge gender ideologies, discourses, and subject positions? A key objective is to move social media studies toward answering not only what people do, but why they engage in digital practice (Brock, 2016). This provides a theoretical grounding for the performativity or “virtue signaling” associated with Web 2.0 (e.g., Zentz, 2021). The project aims to demonstrate how linguistic anthropology and stance analysis enrich micro-level studies of social media discourse, ideology, and subjectivity. Furthermore, it offers a framework for analyzing how digital interactions circulate and contest oppressive gender ideologies, while remaining applicable to a wide range of social media contexts and subject matters.

Methods

The data is comprised of three thematically related posts from the People Incorrectly Correcting Other People Facebook Group (PICOP), a public and heavily trafficked Group with over 2.7 million members. Members share, and discuss, screenshots of instances where they feel someone has been incorrectly corrected. The three posts, posted in late August 2024, pertained to Imane Khelif, the Algerian boxer who became a controversial figure during the 2024 Paris Olympics after her swift victory against Angela Carini. The fight incited discourse surrounding Khelif’s sex and gender and the role of sex testing in sports. In the PICOP posts, participants discussed sex and gender, Khelif, transphobia, and intersexuality. The data set totals 2,041 comments, although the presentation focuses on a sequential analysis of an individual thread reply chain with 44 comments.

The data is qualitatively analyzed primarily via Du Bois’s (2007, 2014) and Du Bois and Kärkkäinen’s (2012) stance analysis approach. The qualitative analysis also draws upon

polysemiotic discourse analysis (Wortham & Reyes, 2020) and social media conversation analytic techniques (e.g., Farina, 2018). This involves attention to the full polysemiotic resources utilized by participants, such as turn-by-turn sequential positioning, syntactic design, lexical choice, prosody, reported speech, and dialogic resonance. I combine these approaches an additional conceptual grounding from the linguistic anthropological works of Agha (2007) and Silverstein (2022). This perspective moves away from viewing social media as a collection of "communities" or "cultures," but rather as continually complexifying and growing digital texts mediated by the social media platform. In this view, every post or comment is a contribution to a text that is influenced by previous interactions and influences how subsequent people engage despite not having read every part of the text. These texts are embedded and scalar: a comment contributes to a thread, a post, a Facebook Group feed, an individual's aggregated personal feed, and so on. Silverstein also views all interaction as rituals that entail certain contextual conditions, occur within particular social and spatiotemporal contexts, and accomplish a purpose that shapes its textual form. Rituals position participants in legible roles and require some presumed shared knowledge of normativity to operate successfully. Stance-taking via a Facebook post or comment is an interaction ritual to visibly and publicly (dis)align oneself with other users and become known both to others and to themselves, while contributing to the ongoing text.

Results

The sequential analysis of the thread demonstrates how participants took stances that evoked scientific discourses about sex and gender to problematize (e.g., deny) or legitimize (e.g., affirm) Imane Khelif, intersex, and gendered subjectivities. Throughout the thread, commenters took stances on several constantly shifting stance objects. The thread quickly became a negotiation about scientifically literate subjectivities. The participants bestowed upon "scientists" the epistemic and deontic rights to determine who should and should not be considered a woman (Stevanoic, 2021). The negotiation in the thread therefore became a struggle over who was more "scientific," attaching a scientific and intelligent subjectivity to a socially conservative social figure on the one hand, or a socially progressive figure on the other. Simultaneously, both figures were labeled as bigots, morons, and idiots. Stance-taking was therefore not only a means to settle a debate about whether Khelif is a woman, etc., but also to establish themselves as scientific and thus intelligent individuals with particular sets of gender-related ideological commitments. Participants accomplished this by disciplining (i.e., ridiculing) the participants they disagreed with, labeling them and their beliefs as abnormal and therefore impermissible. Participants argued with individual participants (e.g., "Barry") while simultaneously positioning them as a member of a larger group (e.g., "scientifically illiterate transphobe"). Despite this, both "sides" of the debate

nonetheless relied upon, and reiterated, an ideology of sex and gender that centers on scientific discourses and is bio-essentialist.

However, participants engaged with the thread in a quantitatively limited capacity. The majority left only one comment. Why contribute to a conversation and never return? This is addressed through the lens of ritual. PICOP is a Facebook Group designed to elicit debate: someone is correct and someone is incorrect in every post, and members weigh in on the debate. If the central ritual of social media participation is to make oneself visible, then the PICOP ritual is even more saliently to publicly and recognizably take a stance on an issue. Yet, in this ritual, participants are not necessarily talking to, or arguing with, a singular interlocutor (e.g., the commenter they are “replying” to), but with the omnipresent yet ultimately imagined audience of all possible PICOP users who may subsequently encounter this text. The consequentiality of comments for subjectivity therefore is not only how it is taken up in future comments. Comments are also consequential in their own right as brief public displays or performances of subjectivity and discipline, even if they are never actually seen or responded to. The act of posting, coupled with the knowledge that the comment is potentially viewable by a networked public, is sufficient to fulfill the intersubjective need for recognition.

Future Work

This case study is intended to lay the foundations for how a ritual and text-based approach to stance-taking, coupled with Foucauldian social theory, helps elucidate the micro-level realizations of subjectivity and ideology in social media interactions. This study is part of a larger project with additional case studies. This future work includes further investigation and consideration of the role Facebook, as a platform, sociotechnological infrastructure, and set of affordances, play in stance-taking. Another two case studies will more explicitly connect how such everyday social media stance-taking functions contribute to the circulation and contestations of gender ideologies and gendered subjectivities, particularly transphobia and the intersections of race and gender.

References

- Agha, A. (2007). *Language and social relations*. Cambridge University Press.
- Brock, A. (2016). Critical technocultural discourse analysis. *New Media & Society*, 20(3), 1012–1030. <https://doi.org/10.1177/1461444816677532>
- Butler, J. (2024). *Who's afraid of gender?* First edition. Farrar, Straus and Giroux.
- Du Bois, J. W. (2007). The stance triangle. In *Pragmatics & beyond*. New series (pp. 139–182). <https://doi.org/10.1075/pbns.164.07du>

- Du Bois, J. W. (2014). Towards a dialogic syntax. *Cognitive Linguistics*, 25(3), 359–410.
<https://doi.org/10.1515/cog-2014-0024>
- Du Bois, J. W., & Kärkkäinen, E. (2012). Taking a stance on emotion: affect, sequence, and intersubjectivity in dialogic interaction. *Text and Talk*, 32(4).
<https://doi.org/10.1515/text-2012-0021>
- Farina, M. (2018). *Facebook and conversation analysis: The Structure and Organization of Comment Threads*. Bloomsbury Publishing.
- Foucault, M. (1977). *Discipline and punish: The Birth of the Prison*. Pantheon Books.
- Foucault, M. (1978). *The History of Sexuality, Volume 1: An Introduction: An Introduction*. Pantheon Books.
- Foucault, M. (1982). The subject and power. *Critical Inquiry*, 8(4), 777–795.
<https://doi.org/10.1086/448181>
- Foucault, M. (1988). "Technologies of the Self." In L. Martin, H. Gutman, & P. Hutton (Eds.),
Technologies of the Self: A Seminar with Michel Foucault. University of Massachusetts Press.
- Marino, F. (2023). #Twospirit: Identity construction through stance-taking on TikTok. *Discourse Context & Media*, 54, 100711.
<https://doi.org/10.1016/j.dcm.2023.100711>
- Stevanovic, M. (2021). Deontic authority and the maintenance of lay and expert identities during joint decision making: Balancing resistance and compliance. *Discourse Studies*, 23(5), 670–689. <https://doi.org/10.1177/14614456211016821>
- Waycott, J., Thompson, C., Sheard, J., & Clerehan, R. (2017). A virtual panopticon in the community of practice: Students' experiences of being visible on social media. *The Internet and Higher Education*, 35, 12–20.
<https://doi.org/10.1016/j.iheduc.2017.07.001>
- Wortham, S., & Reyes, A. (2020). *Discourse analysis beyond the speech event*. Routledge.
- Zentz, L. (2021). #LadiesWeGotYou: Stances of moral–political alignment in the formation of group Identity on Facebook. *Journal of Sociolinguistics*, 25(5), 852–870. <https://doi.org/10.1111/josl.12507>

4B.2 Gone But Not Missed? Understanding the Drivers and Implications of Post Deletions on Bluesky

Yining Wang, Johannes Gruber, Yannik Peters, Kartin Weller (GESIS - Leibniz Institute for the Social Sciences)

Introduction

Data deletion is a continuous phenomenon on online platforms, yet it remains underexplored in empirical social media research. While platform-initiated content moderation has received considerable scholarly attention, user-initiated deletion practices have been studied primarily through qualitative interviews (Sleeper et al., 2013), we know less about the systematic patterns that drive ordinary users to remove their own content.

Understanding post deletion matters for social media researchers and platform scholars for two distinct but related reasons. First, deletion reflects how users manage their online presence and negotiate the tension between ephemerality and permanence on social media (Jacobsen, 2022; Ringel & Davidson, 2022). While platforms store content indefinitely by default, users actively curate what remains visible. Deletion is a form of boundary work through which individuals define which expressions belong to their permanent digital identity and which should be withdrawn. Second, deletion creates methodological challenges that affect the entire field of social media research (Salvatore et al., 2021). When posts disappear after data collection, datasets become inconsistent over time. If deletion follows systematic patterns, favoring certain topics, sentiment, or levels of engagement, observable data may become biased. The timing of data collection can determine not just how much data is available, but which kinds of content remain visible at all.

Large-scale quantitative analyses of deletion patterns remain scarce, particularly on emerging decentralized platforms where complete observational coverage is possible. Research has shown substantial deletion rates on Twitter, with approximately 11% of tweets deleted within a four-week observation period (Bhattacharya & Ganguly, 2021) and removal rates of 23 to 29% within a single year (Khan et al., 2025). However, these studies rely primarily on qualitative methods, surveys, or post hoc analyses, leaving gaps in our understanding of systematic deletion drivers.

Bluesky, an emerging decentralized social network built on the open-source AT Protocol, offers an ideal setting for studying deletion patterns. Its public Firehose API provides complete observational coverage of all creation and deletion events (Balduf et al., 2024; Kleppmann et al., 2024). Unlike centralized platforms, content moderation on Bluesky happens through labels rather than deletions, meaning deletions are

predominantly user-initiated. As decentralized platforms gain traction in the post-Twitter landscape, understanding user behavior on these emerging infrastructures becomes increasingly relevant for the social media research community.

Objectives/Research Questions

This study addresses the gap in large-scale quantitative research on user-initiated post deletion by examining deletion behavior on Bluesky. Our analysis identifies which factors drive deletion decisions and quantifies their relative importance. We test whether deletion is explained by audience response (likes, comments, reposts), content characteristics (text length, URLs, post type), author behavior (posting volume, deletion history), temporal patterns (rapid posting), direct error detection (spelling and grammar), or error-correction behavior (deleting and reposting similar content).

Specifically, we aim to determine which post-level and author-level factors predict deletion on Bluesky and assess their relative importance. We further examine whether deletion patterns create systematic bias in observable social media data, with implications for data quality, temporal validity, and reproducibility in platform research. Our findings contribute to ongoing discussions in digital methods about the reliability and representativeness of social media data.

Methods

Data Collection: We collected all creation and deletion events from the Bluesky Firehose between August 1st and August 31st, 2025. During this period, we captured 86,734,657 posts, of which 2,795,268 (3.22%) were subsequently deleted. Original posts had a deletion rate of 3.96%, while comments showed a lower rate of 2.40%.

Sampling Design: We employed a stratified sampling design, drawing a random sample of 10,000 deleted posts (one per author, text-containing only) and collecting all non-deleted posts from the same authors during the same period, resulting in 944,908 comparison posts from 9,553 authors. This within-author design controls for stable individual characteristics that might simultaneously predict both posting behavior and deletion decisions. To address temporal bias, we standardized engagement measurement to 24 hours after post creation for both groups.

Feature Engineering: We constructed 13 features spanning seven categories: (1) engagement features (log-transformed likes, comments, reposts; comment-to-like ratio), (2) content features (text length, URL presence), (3) post type (original vs. reply), (4) temporal features (rapid posting within 5 minutes), (5) author-level features (total posts, deletion ratio), (6) text similarity to next post (capturing error-correction behavior), and (7) direct error detection using LanguageTool.

Analytical Approach: We employed progressive logistic regression with seven models, each adding a new category of predictors to isolate marginal contributions. We used

class-weighted estimation to address class imbalance and bootstrap resampling ($n = 1,000$) for confidence intervals. We validated findings using Random Forest and XGBoost to ensure robustness across different modeling assumptions (Saito & Rehmsmeier, 2015).

Results

Progressive Model Performance: Our progressive modeling reveals which information categories matter most. Adding author behavior features (posting volume and deletion history) increased model performance eightfold (PR-AUC 0.029 to 0.239), indicating that knowing who posted is far more informative than knowing what was posted or how audiences responded. Text similarity features provided the second major gain (52% improvement), while direct error detection contributed minimally.

Key Findings:

Author Characteristics Dominate: The strongest predictor is posting volume ($\beta = -3.79$, OR = 0.02). Each standard deviation increase reduces deletion odds by 98%. This likely reflects multiple mechanisms: experienced users post more carefully, high-volume posters attach less importance to individual posts, or this group includes automated accounts. Users who have deleted before are more likely to delete again ($\beta = +0.48$, OR = 1.61), suggesting deletion is a stable habit.

Engagement Shapes Curation: Posts receiving more likes are less likely to be deleted ($\beta = -1.05$, OR = 0.35), reducing odds by 65% per SD. However, posts attracting more comments face higher deletion risk ($\beta = +0.35$, OR = 1.42). Posts with high comment-to-like ratios show a protective effect ($\beta = -0.54$, OR = 0.58), indicating balanced engagement supports retention.

Error Correction Follows Strategic Logic: Text similarity to the next post strongly predicts deletion ($\beta = +0.49$, OR = 1.64), indicating users delete and repost corrected content. This effect intensifies with low engagement (interaction $\beta = +0.19$, OR = 1.21), suggesting strategic correction before posts spread. However, actual spelling errors barely predict deletion ($\beta = +0.07$, OR = 1.07), ranking last. Deletion is driven by awareness and intent, not error presence.

Original Posts Face Higher Risk: Standalone posts are more likely deleted than replies ($\beta = +0.35$, OR = 1.42), as they are more visible and represent the user's independent voice.

Cross-Method Validation: Random Forest and XGBoost validate findings. All methods converge: author posting volume, deletion history, likes, and text similarity consistently rank in the top five.

Implications for Social Media Research: Deletion creates systematic, not random, selection bias through three mechanisms: (1) user-level bias, where prolific users

delete far less, causing archived data to overrepresent frequent posters; (2) engagement-based survivorship bias, where low-engagement posts are 2.8 times more likely to disappear; and (3) content-type bias, where original posts show higher deletion than replies. These patterns compromise temporal validity and reproducibility. For researchers relying on social media data, these findings underscore the importance of accounting for deletion-induced bias in study design and recognizing that observable data represents what users chose to preserve.

Future Work

Our analysis captures primarily short-term deletion behavior during one month. Future work will extend observation windows to examine long-term deletion patterns and lifetime deletion probability. We plan to analyze content characteristics not yet examined, including topics, sentiment, emotional tone, and personal disclosure, which may reveal additional systematic deletion patterns relevant to discourse and public opinion research.

We also plan comparative analysis across platforms (e.g., Mastodon, X/Twitter) to examine whether deletion patterns observed on Bluesky generalize to other social media environments or reflect platform-specific dynamics shaped by different infrastructures and user communities. This comparative approach will help establish whether our findings represent broader principles of user behavior or are specific to decentralized platforms.

References

- Balduf, L., Sokoto, S., Ascigil, O., Tyson, G., Scheuermann, B., Korczyński, M., Castro, I., & Król, M. (2024). Looking AT the Blue Skies of Bluesky. Proceedings of the 2024 ACM on Internet Measurement Conference, 76–91.
<https://doi.org/10.1145/3646547.3688407>
- Bhattacharya, P., & Ganguly, N. (2021). Characterizing deleted tweets and their authors. Proceedings of the International AAAI Conference on Web and Social Media, 10, 547–550. <https://doi.org/10.1609/icwsm.v10i1.14803>
- Jacobsen, B. N. (2022). 'You can't delete a memory': Managing the data past on social media in everyday life. Sociological Research Online, 27(4), 1003–1019.
<https://doi.org/10.1177/13607804221110237>
- Khan, M. T., Dimitrov, D., & Dietze, S. (2025). Characterization of tweet deletion patterns in the context of COVID-19 discourse and polarization. Proceedings of the 36th ACM Conference on Hypertext and Social Media, 43–47.
<https://doi.org/10.1145/3720553.3746663>
- Kleppmann, M., Frazee, P., Gold, J., Graber, J., Holmgren, D., Ivy, D., Johnson, J., Newbold, B., & Volpert, J. (2024). Bluesky and the AT Protocol: Usable

decentralized social media. Proceedings of the ACM CoNEXT-2024 Workshop on the Decentralization of the Internet, 1–7.

Ringel, S., & Davidson, R. (2022). Proactive ephemerality: How journalists use automated and manual tweet deletion to minimize risk and its consequences for social media as a public archive. *New Media & Society*, 24(5), 1216–1233.

<https://doi.org/10.1177/1461444820972389>

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>

Salvatore, C., Biffignandi, S., & Bianchi, A. (2021). Social media and Twitter data quality for new social indicators. *Social Indicators Research*, 156(2), 601–630.

<https://doi.org/10.1007/s11205-020-02296-w>

Sleeper, M., Cranshaw, J., Kelley, P., Ur, B., Acquisti, A., Cranor, L., & Sadeh, N. (2013). "I read my Twitter the next morning and was astonished": A conversational perspective on Twitter regrets. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 3277–3286.

<https://doi.org/10.1145/2470654.2466448>

Session 4C - Social Media Culture and Everyday Life (4)

Day 1 · Mon 13 July · 15:15–16:15 Seminar Suite C, ARC

4C.1 Navigating Social Media in the Remote High Mountains: Aging, Literacy, and Technological Obsolescence in Nubri Valley, Nepal

Kunsang Choden (University of Washington)

Introduction

As social media platforms and smartphones expand into geographically remote regions, digital participation is seen as inevitable, progressive, or at times empowering. However, this narrative obscures how platform design, language hierarchies, and rapid technological turnover shape uneven capacities for participation. In Nubri Valley, an ethnically Tibetan community in the high mountains of northwestern Nepal, residents speak Tibetan dialects, practice Tibetan Buddhism, and sustain livelihoods through subsistence farming with occasional tourism-related income (Childs & Choedup, 2018; Childs, 2004). Here, social media is encountered not as seamless connectivity, but as an unreliable yet necessary infrastructure embedded within aging populations, limited formal education, and accelerating device obsolescence.

Participants described adopting smartphones to maintain family ties, particularly with relatives living abroad. A 25 year old youth reflects on the aging population's digital adoption, "When you are old and you have a smartphone in your hand, you have to use it... your son is calling from abroad, you have to be able to receive it". Yet simply adoption does not guarantee meaningful engagement. Instead, older residents often experience confusion, anxiety, and partial participation.

Frequent hardware turnover further disrupts learning processes, especially among elders who are not literate in English or Nepali and primarily communicate in local Tibetan dialects. A senior Tibetan monk and scholar (age 46) notes, "If mother is using one phone this year, the next year, you can't find this phone in the market at all. That's the problem. If the pace of technology keeps changing, their [elders] speed will not be enough at all. " Accumulated knowledge is lost with each replacement cycle, reinforcing dependence on younger family members and intensifying digital inequality. Language barriers compound these challenges. Social media interfaces

primarily operate in English or Nepali languages. Majority older Nubri residents above age 35 cannot read or write. A Tibetan literate Amchi (Tibetan doctor, 73 years) notes, “Everything inside is in English... If you know how to read and write in English, then you can open anything on that phone”. The Tibetan scholar and monk adds, “In the near future, all the shops will go without physical cash. 100%, they’ll make this...In terms of opportunities, they’ll miss a lot, a lot of opportunities. Basic and daily life will also be made difficult.”

This research positions Nubri Valley as a place-based research, examining how social media participation is shaped not only by access, but by the intersection of aging, literacy stratification, and platform-driven technological change. By foregrounding ethnically Tibetan highlanders perspectives, the research challenges dominant assumptions of universal usability embedded within global platform expansion narratives.

Objectives/Research Questions

This study addresses the following research question: How do technological obsolescence, literacy levels, and age shape social media navigation practices in Nubri Valley?

Specifically, the research examines three interconnected dimensions: 1) How rapid hardware replacement and platform updates disrupt learning continuity and erode accumulated digital knowledge among older users. 2) How linguistic and educational barriers embedded within platform interfaces restrict navigation, exploration, and participation. 3) How age-based social roles and labor responsibilities influence perceptions of relevance, motivation, and risk associated with social media use. Participants emphasized that foundational literacy precedes digital inclusion. One resident stated, “The first preliminary thing to ask is, ‘Are people ABCD literate?’... Most people above age 35 must be illiterate... there’s no need to question digital literacy”.

Rather than framing social media navigation challenges as an individual skills deficit, this research conceptualizes digital capabilities as a socio-technical condition produced through market logics, interface design, language dominance, and generational learning ecologies.

Methods

This IRB approved study draws on 41 semi-structured interviews, 17 culturally embedded participatory “tea conversations,” and long-term observational data, including reflective journaling. Participants represented multiple generations: 29% young adults (20–39 years), 44% middle-aged adults (40–59 years), and 27% elders (60–85 years) as well as diverse occupational backgrounds, including farmers, monks, retired elders, teahouse owners, NGO staff, village leaders, and youth.

Interviews were conducted in Tibetan and Nepali. These were later translated and transcribed into English. Study data were inductively thematically coded using qualitative software MaxQda. Analysis focused on recurring narratives related to technology adoption practices, fear of device damage, language barriers, obsolescence, and everyday social media engagement.

Rather than relying on standardized digital literacy metrics, this research adopts a situated use perspective, examining how participants navigate interfaces, rely on visual cues, depend on younger intermediaries, or selectively disengage from platforms. This methodological approach foregrounds lived experience and everyday interaction with social media infrastructures, aligning with critical social computing and ICT ethnography traditions.

Results

Findings reveal that social media navigation in Nubri Valley is shaped by three intersecting dynamics, alongside persistent challenges related to sporadic connectivity caused by remote geography, limited maintenance capacity, and frequent infrastructural breakdowns (Gomez & Camacho, 2012; Jang et al., 2019; Star, 1999; Young, 2019).

First, technological obsolescence undermines learning stability. While recounting the process of buying a phone for his 75-year-old mother, a senior monk (46 years) noted that elders like his mother often lose hard-earned digital skills each time a device is replaced: “Once the old phone is no longer in use... full phone knowledge is also lost”. Fear of damaging costly devices discourages experimentation: “They feel scared of pressing on the phone” (lodge owner, age 50), limiting exploratory learning that social media platforms often assume.

Second, literacy and language exclusion structurally constrain participation. Participants emphasized that English-dominated interfaces restrict access: “If you don’t know good English... you can only use a few features” (village spiritual leader, age 57). Many users rely on visual memory, recognizing contact photos instead of reading names (farmer, age 31) or scroll passively through content without targeted search, “I scroll around and watch whatever is shown on the screen” (Tibetan doctor, 73 years).

Third, age shapes motivation and engagement patterns. Older residents frequently described social media as “for youths”, while simultaneously relying on phones for transnational communication. This produces selective usage focused on calling rather than active social media participation.

Together, these dynamics demonstrate that social media navigation and accessibility challenges in Nubri are produced through platform design, market-driven

obsolescence, and generational learning conditions, not simply through access disparities.

Future Work

From a policy and design perspective, this research calls for rethinking digital inclusion beyond connectivity metrics. Sustainable social media participation requires longer device life cycles, stable interface design, and community-based training models that recognize memory limitations, labor constraints, and economic risk.

By centering high-mountain communities at the margins of platform expansion, this work contributes to Social Media & Society debates on digital inequality, platform governance, and inclusive design, arguing that meaningful participation depends on technological stability, linguistic accessibility, and culturally situated learning ecologies.

References

- Childs, G. (2004). *Tibetan diary: From birth to death and beyond in a Himalayan valley of Nepal*. Univ of California Press.
- Childs, G., & Choedup, N. (2018). *From a trickle to a torrent: Education, migration, and social change in a Himalayan valley of Nepal*. University of California Press.
- Gómez, R., & Camacho, K. (2012). Who Uses Public Access Venues?. In *Libraries, Telecentres, Cybercafes and Public Access to ICT: International Comparisons* (pp. 11-23). IGI Global.
- Jang, E. H. B., Garrison, P., Vistal, R. V., Cunanan, M. T. D., Perez, M. T., Martinez, P., ... & Heimerl, K. (2019). Trust and technology repair infrastructures in the remote rural Philippines: Navigating urban-rural seams. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 1-25.
- Star, S. L. (1999). The ethnography of infrastructure. *American behavioral scientist*, 43(3), 377-391.
- Young, J. C. (2019). Rural digital geographies and new landscapes of social resilience. *Journal of Rural Studies*, 70, 66-74.

4C.2 Distinguishing suicide attempts from suicidal ideation in young people: The role of addictive digital behaviours and negative life events

Marianne Etherson (University of Glasgow)

Introduction

Though contentious, research suggests that the use of digital technologies contributes to mental health problems, such as suicidal thoughts and behaviors in young people in the US (Twenge et al., 2018). In line with this, various systematic reviews and meta-analyses have found digital engagement (e.g., internet addiction, frequency of smartphone use, and problematic screen use) to increase the risk of suicidal thoughts and behaviours (Chen et al., 2024; Macrynika et al., 2021; Marchant et al., 2017). While important, there is a lack of research examining whether digital activity distinguishes suicide attempts from suicidal ideation (see Etherson et al., 2025). Liu et al. (2020), for instance found suicide-related social media behaviours to distinguish between suicide attempts and ideation in young people. Yet, research examining other types of digital activity is warranted.

Likewise, in the Adolescent Brain Cognitive Development (ABCD) dataset, only one study has examined the role of digital activity (e.g., screen time; van Velzen et al., 2022). Van Velzen et al. (2021), for instance, found that screen time did not distinguish suicide attempts from ideation in young people. Notably, however, screen time has been heavily criticised as it does not distinguish between different types of risky digital activity (e.g., Kostyrka-Allchorne et al., 2025). Research examining more specific digital risk markers (e.g., social media addiction) is vital to understand whether they distinguish suicide attempts from suicide ideation in young people. Research also suggests addictive digital behaviours may be amplified in response to negative life events (e.g., may act as a coping mechanism; Wolfers & Utz, 2022). Stressful life events for instance, have been linked to internet addiction and social media addiction (Brailovskaia et al., 2019; Wang et al., 2022), suggesting a possible interaction between these variables (see Wolfers & Utz, 2022).

Objectives/Research Questions

The current study aims to build upon our understanding by examining whether addictive digital behaviours and negative life events distinguish suicide attempts from suicidal ideation in young people in the Adolescent Brain Cognitive Development (ABCD) study.

RQ1: Does social media addiction, mobile phone involvement and video game addiction distinguish the suicidal ideation and behaviour group from the control group and the suicidal behaviour group from the suicidal ideation group?

RQ2: Do negative life events distinguish the suicidal ideation and behaviour group from the control group and the suicidal behaviour group from the suicidal ideation group.

RQ3: Does the interaction between addictive digital behaviours and negative life events distinguish the suicidal ideation and behaviour group from the control group and the suicidal behaviour group from the suicidal ideation group?

Methods

Participant data was from the ABCD Study curated Annual Release 5.0. The ABCD study is a large, prospective cohort study of brain development and health in adolescents from 21 recruitment sites across the US. The ABCD study includes approx. 11,800 children aged 9 to 10 years at baseline through US school systems who are followed annually for up to 10 years. Centralized institutional ethical approval was obtained from the University of California, San Diego.

Cross-sectional analyses were conducted using 2-year follow-up data (young people aged 11-12 years) from the Adolescent Brain Cognitive Development study.

We used data reported by parents on child age, sex, and race/ ethnicity. Young people responded to questionnaires assessing social media addiction (Social Media Addiction Questionnaire), addictive mobile phone use (Mobile Phone Involvement Questionnaire), video game addiction (Video Game Addiction Questionnaire), total negative life events (Adverse Life Events Scale), and suicidal thoughts and behaviours (Kiddie Schedule for Affective Disorders and Schizophrenia for the Diagnostic and Statistical Manual of Mental Disorders).

Multinomial regression analysis was then conducted in R using group membership of suicide status (i.e., a suicide behaviour group, a suicide ideation only group, and a control group) as the dependent variable.

Results

Participants were classified into three mutually exclusive groups: suicidal behaviour (N = 167), suicidal ideation only (N = 691), and controls (N = 10001).

Preliminary results revealed that social media addiction (OR = 1.05; 95% CI 1.03-1.07), video game addiction (OR = 1.02 ; 95% CI 1.01-1.04) and negative life events (OR = 1.23; 95% CI 1.16-1.30) distinguished the suicidal ideation group from the control group, however mobile phone involvement did not (OR = 1.00; 95% CI 0.99-1.01). The interaction between social media addiction and negative life events distinguished the suicidal ideation group from the control group, however no other interactions were significant.

Similarly, social media addiction (OR = 1.08; 95% CI 1.04-1.12), video game addiction (OR = 1.03 ; 95% CI 1.00-1.07) and negative life events (OR = 1.45; 95% CI 1.32-1.59) distinguished the suicidal ideation group from the control group, but no interactions were significant. In contrast, only negative life events distinguished the suicidal behaviour group from the suicidal ideation group (OR = 1.18; 95% CI 1.07-1.31).

Findings indicate that suicidal ideation is more strongly linked to problematic digital behaviours, particularly in response to negative life events. However, suicidal behaviour appears to be more strongly driven by negative life events.

Future Work

This study is the first to examine whether addictive digital behaviours distinguish suicide attempts from suicidal ideation in a large US dataset of young people. Findings suggest that problematic digital behaviours may signal risk for suicidal ideation (particularly following negative life events), highlighting the value of screening specific digital risk factors (e.g., social media addiction) for early identification. In contrast, the assessment of negative life events appears critical for preventing escalation from suicidal thoughts to suicidal behaviour in young people.

While this study provides important contributions, findings should be considered in light of the limitations of the study. For instance, despite the large sample size of the ABCD study, young people exhibited low-base rates of suicidal ideation and behaviours, hence the study may have lacked sufficient power. Given that incidence rates of suicidal ideation and behaviours rise sharply from childhood to adolescence (Hawton et al., 2012), research is required to replicate whether various risk factors distinguish between suicide attempts and suicidal ideation in a larger and older sample of young people. In line with this, given the cross-sectional study design, future work examining addictive digital behaviours should utilise follow-up data from the ABCD study.

In addition, while our study examined important markers of addictive digital behaviours, future work should examine other important markers of digital activity in this context (e.g., exposure to suicide/ self-harm on social media). The current study also relied on self-report measures of digital behaviours. Future work should move beyond self-reported digital behaviours (e.g., parental reports or experimental studies).

References

- Brailovskaia, J., Rohmann, E., Bierhoff, H. W., Schillack, H., & Margraf, J. (2019). The relationship between daily stress, social support and Facebook Addiction Disorder. *Psychiatry Research*, 276, 167-174.
<https://doi.org/10.1016/j.psychres.2019.05.014>

- Chen, Z., Liao, X., Yang, J., Tian, Y., Peng, K., Liu, X., & Li, Y. (2024). Association of screen-based activities and risk of self-harm and suicidal behaviors among young people: A systematic review and meta-analysis of longitudinal studies. *Psychiatry Research*, 338, 115991. <https://doi.org/10.1016/j.psychres.2024.115991>
- Etherson, M. E., Lee, S., Loney, K. J., Steward, I. P., Ward, J., McClelland, H., Kandola, A., De Alcantara Mendes, J. A., Hollis, C., Townsend, E. Auer, D. P., & O'Connor, R. C. On behalf of the Digital Youth Research Programme. (2025). Exploring risk and protective factors which distinguish suicidal and self-harm behaviours from suicidal and self-harm ideation in young people: A systematic review. *PloS One*, 20(9), e0326381. <https://doi.org/10.1371/journal.pone.0326381>
- Hawton, K., Saunders, K. E., & O'Connor, R. C. (2012). Self-harm and suicide in adolescents. *The Lancet*, 379(9834): 2373-2382. [https://doi.org/10.1016/S0140-6736\(12\)60322-5](https://doi.org/10.1016/S0140-6736(12)60322-5)
- Kostyrka-Allchorne, K., Bourgaize, J., Murray, A., Stoilova, M., Abbas, I., Bridgwood, A., ... & Digital Youth Research Programme. (2025). Understanding youth online experiences and mental health: development and validation of the digital activity and feelings inventory (DAFI). *International Journal of Methods in Psychiatric Research*, 34(2), e70028. <https://doi.org/10.1002/mpr.70028>
- Liu, X., Huang, J., Yu, N. X., Li, Q., & Zhu, T. (2020). Mediation effect of suicide-related social media use behaviors on the association between suicidal ideation and suicide attempt: cross-sectional questionnaire study. *Journal of Medical Internet Research*, 22(4), e14940. <https://doi.org/10.2196/14940>
- Macrynika, N., Auad, E., Menjivar, J., & Miranda, R. (2021). Does social media use confer suicide risk? A systematic review of the evidence. *Computers in Human Behavior Reports*, 3, 100094. <https://doi.org/10.1016/j.chbr.2021.100094>
- Marchant, A., Hawton, K., Stewart, A., Montgomery, P., Singaravelu, V., Lloyd, K., ... & John, A. (2017). A systematic review of the relationship between internet use, self-harm and suicidal behaviour in young people: The good, the bad and the unknown. *PloS One*, 12(8), e0181722. <https://doi.org/10.1371/journal.pone.0181722>
- Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among US adolescents after 2010 and links to increased new media screen time. *Clinical Psychological Science*, 6(1), 3-17. <https://doi.org/10.1177/2167702617723376>
- van Velzen, L. S., Toenders, Y. J., Avila-Parcet, A., Dinga, R., Rabinowitz, J. A., Campos, A. I., ... & Schmaal, L. (2022). Predictors of suicidal thoughts and behavior in children: results from penalized logistic regression analyses in the ABCD study.

The British Journal of Psychiatry, 220, 210–218.

<https://doi.org/10.1192/bjp.2022.7>

Wang, J., Hao, Q. H., Tu, Y., Wang, Y., Peng, W., Li, H., & Zhu, T. M. (2022). The relationship between negative life events and internet addiction disorder among adolescents and college students in China: a systematic review and meta-analysis. *Frontiers in Psychiatry*, 13, 799128.

<https://doi.org/10.3389/fpsy.2022.799128>

Wolfers, L. N., & Utz, S. (2022). Social media use, stress, and coping. *Current Opinion in Psychology*, 45, 101305. <https://doi.org/10.1016/j.copsyc.2022.101305>

4C.3 Who gets to speak the truth? Russian invasion of Ukraine and construction of authenticity in social media communication

Marie Heřmanová (University College London)

Introduction

Shortly after the Russian invasion of Ukraine in February 2022, the media started to call the conflict “the first TikTok war” (Bosch, 2022). Online platforms were almost immediately flooded with audiovisual content from the site of the conflict, both authentic and misleading. This was further enhanced by the fact that prominent personalities on social media, including professional content creators, journalists, and politicians (such as the Ukrainian president Volodymyr Zelensky), took to social media to talk about the war and support their political messaging (Serafin, 2022). At the same time, content creators, sometimes also labelled social media influencers, pivoted in their content from lifestyle to showing the reality of the invasion directly from Ukraine. Immediately, starting on the first day of the conflict, Ukrainian content creators acted as important sources of news from the site of the conflict, with many documenting their everyday lives during the war or their attempts to flee to Europe (Divon & Eriksson Krutrök, 2024). Professional journalists from Ukrainian media, including the platform United24 operated by the Ukrainian government, quickly developed content strategies for social media in order to reach broad international audiences and sustain attention over time (Węglińska et al., 2024).

In this sense, the full-scale Russian invasion of Ukraine presents an important case study of the shifting dynamics of online political communication and the building of trust and authority in online spaces. As Divon and Eriksson Krutrök note, the war in Ukraine marks a new era of conflict documentation, with “war influencers” emerging as a category of social media users who “transform war zones into consumable units of content by using platform-specific vernaculars to convey their lived experiences” (Divon & Eriksson Krutrök, 2025). Professional journalists, on the other hand, are under increasing pressure to adapt to the widespread use of social media (Mellado and Hermilla, 2021) and to perform as content creators to promote their reporting, relying on practices such as vlogging, presenting behind-the-scenes content, adapting to the speed of social media platforms, and producing content that can be perceived as authentic rather than professional. The war in Ukraine has exacerbated these issues for journalists reporting from the site of the conflict (Węglińska et al., 2024).

Objectives/Research Questions

The presented paper investigates how content creators and journalists based in Ukraine employ social media in their everyday work and how social media functions as a tool for documenting everyday life in a war zone. Rather than treating social media merely as a channel for dissemination, the paper approaches platforms as sociotechnical environments that actively shape practices of witnessing, storytelling, and authority. Based on long-term ethnographic fieldwork with content creators and journalists based in Ukraine, the research pursues the following objectives:

On an empirical level, the paper explores everyday practices of using social media, with a specific focus on TikTok and Instagram as the most widely used platforms among the research respondents, to create content about the war and to document everyday reality under conditions of ongoing conflict. The analysis attends to how personal experiences are rendered visible through short-form video and platform-native storytelling formats. Particular attention is paid to the process of “translating the war” from local to global contexts using platform affordances such as viral trends, hashtags, templates, sounds, and algorithmic logics that privilege certain aesthetics, temporalities, and narrative structures. This includes examining how creators and journalists anticipate imagined international audiences and adapt their content accordingly. Related to this, the paper explores how people who label themselves as content creators approach these practices differently than people who work as professional journalists, even when the concrete practices they employ—such as filming on location, addressing the camera directly, or narrating lived experience—are often very similar. In this sense, the paper examines how both groups reflect on, justify, and negotiate their positionality, professional identity, and ethical responsibilities while navigating the increasingly blurred boundaries between journalism and content creation as professions.

On a theoretical level, the paper explores the links between platform vernaculars and processes of building authority based on the authenticity of “being there.” It asks how authority is constructed not only through professional credentials or institutional affiliation, but through claims of proximity, presence, and embodied experience that are made legible through platform-specific forms of communication. Drawing from research on authenticity in social media communication (Abidin, 2017, among others), the role of content creators in conflict coverage (Divon & Eriksson Krutrök, 2025), and epistemic hierarchies in conflict reporting (Kotíšová, 2023), the paper proposes the concept of “authenticity of being there” as a key analytical lens. This concept helps to capture how research respondents mobilize physical presence in a war zone, personal risk, emotional exposure, and everyday ordinariness as resources for constructing credibility and trust.

Methods

Methodologically, the paper is grounded in digital anthropology and the anthropology of media and employs a combination of online and offline ethnography. The ongoing offline ethnographic fieldwork started in November 2025 and takes place mostly in Kyiv, with occasional interviews and observations in other Ukrainian cities such as Lviv and Uzhhorod, depending on the situation of the research respondents. The respondents are content creators and journalists who create content on Instagram and TikTok in English or in both English and Ukrainian. Content creators are defined as people who use social media professionally and monetize their content either directly (via subscriptions, ads, etc.) or indirectly (for example, through fundraising calls). Similarly, the journalists included in the research create social media-specific content, rather than only reposting reporting published elsewhere. They mainly work for two English-language Ukrainian media platforms, the Kyiv Independent and United24. In several cases, the two categories overlap, as some content creators previously worked for these media outlets and later transitioned to full-time content creation, while some content creators were hired by media organizations based on their social media presence. So far, the research has engaged around 15 respondents; however, given the ethnographic nature of the project, the focus is on long-term observation and deep understanding rather than representativeness. The offline fieldwork is complemented by online ethnography, including observation of respondents' social media accounts and ethnographic content analysis (Altheide & Schneider, 2013). Fieldwork is conducted in English and Ukrainian.

Results

While the research is ongoing and the paper presents a work in progress, the preliminary analysis points to two key findings that illuminate the tensions inherent in contemporary conflict-related social media practices. First, both content creators and journalists express deep ambivalence about their use of social media, even as they rely on it as a central and often indispensable working tool. Social media is simultaneously described as enabling visibility, reach, and immediacy, and as constraining the kinds of stories that can be told. For example, Yulia, a former journalist and now full-time content creator, notes that “you do what you have to do to make the platform work for you. It’s not necessarily the best work, but it is what works, if you want to get the message across.” Similarly, Jon, a journalist, reflects that “what I do on Instagram and TikTok is not journalism. It is seen by way more people than my journalism, but it is mostly like ranting at the camera.” Such reflections highlight a perceived disconnect between professional ideals and platform-driven practices. This ambivalence is most often articulated in relation to the algorithmic affordances of platforms, which are seen as limiting because they require simplification, emotional appeal, and speed at the expense of depth, nuance, and context. At the same time, these affordances are widely perceived as unavoidable and even inevitable, often described as a “necessary evil” in order to remain visible and relevant within highly competitive platform environments.

Second, all respondents acknowledge that the boundaries between social media content creation and journalism are increasingly blurred, both in terms of everyday practices and in how roles are publicly perceived. Several content creators explicitly question whether they are, in effect, performing journalistic work, albeit without the institutional protection, resources, ethical frameworks, and training associated with professional media organizations. At the same time, there exists an intricate and not always overtly articulated hierarchy between different types of content and different types of creators. As Jon puts it, “we are all here, that’s what gives us credibility,” pointing to the authenticity of being there as a central dimension of this hierarchy. However, this authenticity is not distributed evenly. Content produced in English is valued differently from content produced in Ukrainian, and content created by Ukrainians—even when addressed to international audiences in English—is widely perceived as more authentic than content created by foreigners based in or temporarily visiting Ukraine. These distinctions reveal how authority in online war reporting is shaped by intersecting factors such as language, nationality, proximity, and claims to lived experience.

Future Work

As the paper presents work in progress, future research will focus on a deeper and more systematic analysis of the ambivalences and hierarchies identified in the preliminary findings. In particular, greater attention will be paid to content produced in the Ukrainian language, both as a form of domestic communication and as a site of meaning-making that is often overshadowed by English-language content aimed at international audiences. This will potentially be approached through a comparative analysis of Ukrainian- and English-language content covering the same events and time periods on the same social media platforms, allowing for closer examination of differences in narrative strategies, platform vernaculars, and audience engagement. Additionally, gender emerges as a crucial dimension of content creation practices and professional positioning, with more content creators being women, while journalists—especially in editorial and managerial positions—are more often men. Future analysis will therefore examine how gender intersects with platform visibility, authority, and perceptions of credibility within the evolving online media ecosystem shaped by the war in Ukraine.

References

- Abidin, C. (2017). #Familygoals: Family influencers, calibrated amateurism, and justifying young digital labor. *Social Media and Society*, 3(2).
<https://doi.org/10.1177/2056305117707191>

- Altheide, D. L., & Schneider, C. J. (2013). Ethnographic Content Analysis. In *Qualitative Media Analysis* (s. 23–37). SAGE Publications, Ltd.
<https://doi.org/10.4135/9781452270043>
- Bösch, M. (2022). WARTOK – TikTok und der Krieg in der Ukraine. *Televizion*, 35, 4-8.
- Divon, T., & Eriksson Krutrök, M. (2024). Playful Trauma: TikTok Creators and the Use of the Platformed Body in Times of War. *Social Media + Society*, 10(3), 20563051241269281. <https://doi.org/10.1177/20563051241269281>
- Divon, T., & Eriksson Krutrök, M. (2025). The rise of war influencers: Creators, platforms, and the visibility of conflict zones. *Platforms & Society*, 2, 29768624251325721. <https://doi.org/10.1177/29768624251325721>
- Kotišová, J. (2023). The epistemic injustice in conflict reporting: Reporters and ‘fixers’ covering Ukraine, Israel, and Palestine. *Journalism*, 14648849231171019. <https://doi.org/10.1177/14648849231171019>
- Mellado, C., & Hermida, A. (2021). The Promoter, Celebrity, and Joker Roles in Journalists’ Social Media Performance. *Social Media + Society*, 7(1), 2056305121990643. <https://doi.org/10.1177/2056305121990643>
- Serafin, T. (2022). Ukraine’s President Zelensky Takes the Russia/Ukraine War Viral. *Orbis*, 66(4), 460–476. <https://doi.org/10.1016/j.orbis.2022.08.002>
- Węglińska, A., Seklecka, A., Peszyński, W., & Węgliński, B. (2024). Transformation of a war journalist profession: The case of the full-scale Russo–Ukrainian conflict. *Media, War & Conflict*, 17506352241307013. <https://doi.org/10.1177/17506352241307013>
-

Session 4D - **PANEL: When the trolls find our research: A workshop in online harassment of scholars and how to move forward**

Day 1 · Mon 13 July · 15:15–16:15 · Seminar Suite B, ARC

*Anatoliy Gruzdt (Ted Rogers School of Management at Toronto Metropolitan University),
Esteban Morales (University of Groningen), Yimin Chen (RMIT University), Jaigris Hodson
(Royal Roads University)*

The possibilities for advancing relevant and impactful research lie in connecting academic endeavours with networks of other scholars and the general public (Brennan et al., 2004). In this scenario, digital platforms such as X/Twitter, TikTok, YouTube, and Instagram have increasingly become places where researchers can find such opportunities to share their research and engage in fruitful discussions about their work (Greenhow & Gleason, 2014). Nevertheless, fostering these conversations hinges on having a secure and supportive platform to do so. Unfortunately, scholars who rely on digital platforms such as social media are often targets of online harassment, trolling, and bullying (hereafter grouped as online abuse) (Veletsianos et al., 2018). Those who bear the brunt of these harms are frequently marginalized researchers such as black, indigenous, and LGBTQ+ researchers (Galpin & Vernon, 2024). These practices of online abuse have critical consequences for those attending SMSociety, silencing important science and research communication efforts. These problems are only getting worse, as recent political and policy changes around the world, as well as technical and moderation changes to platforms, have exacerbated these problems.

The body of evidence suggests that online abuse has become a common consequence of internet use. In addition, experiences of online abuse are on the rise. In 2014, Pew Research Center found that 35% of American adults had experienced online abuse (Duggan, 2014). This number grew to 41% in a 2017 study (Duggan, 2017b). In 2021, researchers at Pew found evidence to suggest that online abuse has become more extreme (Desilver, 2021). Pew Research Center's 2021 study found that although the number of people that report experiencing online abuse has remained fairly stable since 2017, the number of individuals that experience severe forms of abuse as well as multiple forms of harassing behaviors has dramatically increased.

Online abuse can take many forms, including: 1) inappropriate or harassing messages (Jane, 2014), 2) the non-consensual distribution images intended to be private (McGlynn and Rackley, 2017a, 2017b), 3) other privacy-related offences, such as doxing which is the disclosure of personal or private information, known as doxing (Eckert and

Metzger-Riftkin, 2020), 4) disruption to people's online profiles or websites through hacking (Massanari, 2015), 5) disruption of people's offline lives through stalking or visiting a person's place of employment (Jane, 2014) and 6) technical disruptions such as DDoS attacks (Traer and Bednar, 2021). As can be seen in these examples, online abuse acts exist on a spectrum of severity and not all instances of abuse cause the same degree of harm, though all disrupt the lives of those who experience abuse. One of the difficulties of defining online abuse comes from how quickly new forms of abuse emerge: in the last few years, for example, the role of generative AI in online abuse has quickly risen as a concern (Wood, 2024). As technologies change, so too do methods of abuse.

In 2014, Emma Jane noted a lack of research on the topic of online abuse, this has changed dramatically. There is now a substantial body of research that underscores the prevalence of online abuse across varying online contexts, some of which was authored by the proposed panelists (see for example, Hodson et. al, 2018. Soares et. al., 2023, Chen et. al, 2025, Morales et. al, 2024, O'Meara et. al, 2024). Some of the literature even examines online abuse specifically in the context of higher education. For example, a study of 121 faculty at a Canadian university, Cassidy, Faucher, and Jackson (2014) found that over 17% of respondents had experienced some form of online abuse. Following this important early work, Veletsianos et. al (2018), showed the strain placed on the women scholars who experience online abuse when they use digital tools to communicate their work. More recently, Hodson et. al(2023), showed how scholars from marginalized backgrounds often feel they do not get the support they need from their institutions.

The significant role online spaces play in the professional and personal lives of researchers cannot be overstated. As digital technology becomes more deeply embedded into modern life and the digital economy subsequently expands, staying offline to avoid abuse is untenable. A growing portion of the population cannot afford to simply withdraw from online spaces (Gosse, 2021). Given this relationship, there is a strong imperative to direct our attention to understanding how better to support and protect researchers, particularly digital researchers, from online abuse. The proposed panel contributes to this goal, by providing both a well researched understanding of the issue as situated in the extant literature, and also providing some strategies for researchers to mitigate the harms of online abuse.

The objectives of this proposed session are as follows:

1. Convene leading researchers of online harms and toxicity;
2. Create a space to collaboratively and creatively discuss the online abuse of researchers, and its impacts on scholarly endeavours;

3. Explore possible collaborative interventions among participants to mitigate its harms;
 4. Initiate a community of practice for researchers to support each other, share resources and advocate at their own institutions.
-

Session 4E - **PANEL: Social Media in Flux: Embracing uncertainty and experimentation in research methods**

Day 1 · Mon 13 July · 15:15–16:15 · Studio 2, ARC

Anna Feigenbaum (University of Glasgow), Ozlem Demirkol Tonnesen (University of Bergen) Doug Specht (University of Westminster) Sweta Baniya (Virginia Tech) Taylor Annabell (Cardiff University)

This panel features researchers from the Social Media Methods Workbook. The workbook is our attempt to create an accessible, honest and creative intervention into the teaching and practice of social media research methods. Drawing on the principles that shaped the workbook, the session has three core objectives:

First, we aim to make more visible the inherent messiness and complexity of social media methods. Research design in this field rarely follows a linear path: platforms evolve during projects, data access shifts, and research questions require refinement as scholars encounter unexpected affordances and constraints. Rather than presenting methodology as a mechanical process of “method selection,” the workbook frames research design as iterative and adaptive. The panel will explore strategies for navigating this complexity, including combining methods thoughtfully, scoping projects realistically, and building strong conceptual foundations that can navigate shifting digital environments.

Second, we confront the difficulty of working within institutional ethical guidelines that are often ambiguous, platform-dependent and inconsistently enforced. Institutional ethics frameworks were largely developed for offline or clearly bounded digital contexts, and do not always account for the blurred boundaries between public and private spaces, evolving terms of service, shifting data access policies, etc.

Researchers must frequently make judgement calls about consent, anonymity, data scraping, platform governance and the redistribution of user-generated content. Panel participants will discuss how researchers can navigate ethical frameworks that are not fit for purpose, and move beyond minimal compliance toward reflexive, context-sensitive ethical decision-making that accounts for power, privacy, vulnerability and platform governance.

Third, we consider the challenge of sustaining methodological innovation amid rapid technological and policy change. As platforms update affordances, APIs, moderation systems and monetisation models, established research practices can quickly become outdated. Our panel of workbook contributors will present our future-oriented approach

that emphasises transferable skills, critical platform literacy and creative experimentation.

Session 5A - Platform Governance and Regulation (2)

Day 1 · Mon 13 July · 16:30–17:30 RD Suite, ARC

5A.1 Call the (Bot-)Police: User-Led Platform Governance of “(In)Authenticity” on Instagram

Nathalie Schäfer (Bauhaus-University Weimar)

Introduction

Meta’s 2025 decision to end third-party fact-checking—justified by Chief Global Affairs Officer Joel Kaplan’s claim that “too much harmless content gets censored”—exposes a profound contradiction in platform governance: while Meta admits to over-moderation, it simultaneously fails to act against rampant inauthentic behavior. On Instagram, this imbalance manifests in the unchecked proliferation of fame-enhancing bots and porn bots, which evade automated detection and platform intervention. In response, users have taken matters into their own hands, forming self-organized “bot police” profiles to identify, expose, and report inauthentic accounts. These users deploy a dual strategy: detecting coordinated bot behavior—such as repetitive posting and engagement—and publicly shaming offenders through flagging, documentation, and digital vigilantism (Favarel-Garrigues, Tanner & Trottier 2020).

This study investigates how user-led bot policing reflects and reshapes Instagram’s authenticity governance. Drawing on Tarleton Gillespie’s (2019) concept of platforms as “custodians of the Internet,” I analyze how these users position themselves as de facto enforcers of platform norms—yet operate beyond institutional oversight. Through qualitative visual and textual analysis (Laestadius 2016) of two case studies, I reveal that bot policing is not merely compliance with platform rules, but a complex practice shaped by power asymmetries between Meta and users, and among users themselves. The bot police’s actions—while often effective—raise ethical concerns about digital vigilantism, misidentification, and the privatization of public accountability.

This research contributes to growing scholarship on user agency in platform governance by uncovering the hidden labor and moral authority embedded in user-led policing. It highlights how Instagram’s design—through features like flagging and community notes—encourages peer surveillance (Leistert 2016) while deflecting responsibility from the platform. Ultimately, the bot police exemplify both the potential

and perils of decentralized governance in a capitalist, algorithmically mediated public sphere.

Objectives/Research Questions

This study examines how user-led bot policing on Instagram functions as a form of peer surveillance and mutual moderation within the platform's broader authenticity governance (Lindquist & Weltevrede 2024). It investigates the dual role of users as both enforcers and subjects of platform power, particularly in the context of Meta's shifting moderation strategies—from third-party fact-checking to community notes (McMahon 2025)—and the resulting transfer of responsibility to users.

The research focuses on two distinct bot policing practices: one targeting fame-enhancing bots (automated accounts that boost engagement metrics), and another addressing porn bots (accounts using sexual solicitation to exploit users). These cases reveal how users interpret and enforce “authenticity” in ways that reflect both platform norms and their own agency.

The study asks:

How do user-led bot policing practices contribute to Instagram's authenticity governance?

In what ways do these practices reflect, reproduce, or challenge power asymmetries between the platform, users, and other users?

How do the bot police position themselves as custodians of authenticity, and what are the implications of their self-appointed authority?

By analyzing public profiles through qualitative visual and textual analysis and applying thick description, this research uncovers the complex interplay between platform design, user agency, and digital vigilantism—shedding light on how mutual moderation shapes the visibility, legitimacy, and boundaries of online identity.

Methods

This study examines two user-led bot policing profiles on Instagram as case studies, representing distinct approaches to combating fame-enhancing bots and porn bots. The first profile focuses exclusively on detecting and reporting automated engagement, offering insight into botting practices. The second includes a German group (active 2019–2020) and a US-based profile (active as of August 2024), both using the hashtag #nobotsby2020. The German group, which ceased activity after 2020, may have abandoned its mission due to a self-imposed deadline. In contrast, the US profile remains active and has gained legitimacy through high follower counts and public endorsements from influencers.

Data were collected manually via Instagram's public interface in October 2023, March 2024, and August 2024, using screen captures to preserve content. All data were

publicly available and re-verified prior to analysis. To protect privacy, identifiable information of targeted accounts was anonymized or excluded. The study's limitations include a small sample size and the inherently obfuscated nature of bot activity, which restricts full observability.

A qualitative visual and textual analysis (Laestadius 2016) was conducted, examining posts, captions, hashtags, comments, and interaction metrics. Hashtags served as entry points for data tracing. Clifford Geertz's (1973) "thick description" was applied to interpret the bot police's actions, intentions, and social context, revealing how they construct and enforce authenticity. This approach enables a nuanced understanding of user-led moderation as a form of peer surveillance and digital governance.

Results

This study investigated the practice of user-led bot policing against botting and porn bots on Instagram. Two case studies showed various strategies for user participation in Instagram's authenticity governance. Users' scope of action was limited to flagging and reporting profiles or content. There was no feature for governing "inauthentic" liking or following behavior. The key findings showed that authenticity governance on Instagram is crucial for some users, motivating those who saw botting as moral deviance to become active. The user-led bot policing profiles developed pillorying and coordinated flagging strategies to participate in Instagram's authenticity governance. They considered themselves necessary to protect "authenticity" and the platform values and perceived themselves as its custodians.

Their perspective of "authentic" behavior as "genuine" or "real" behavior does not account for the hybrid and the entangled nature of human and non-human agents in social media environments. The bot police acted autonomously without having authority from the platform or an institutionalized code of conduct to execute their approaches to authenticity governance. Their digital vigilantism potentially caused harm to other Instagrammers. The pursuit of visibility was not only the reason for botting and deploying porn bots but also a strategy used by bot police to establish their own legitimacy and a motivation for their bot-policing activities. However, any individual user engaging in the practice had limited agency, and the case studies demonstrated the imbalanced power relations between users and the platform.

In this respect, the findings revealed complex power asymmetries regarding user-led authenticity governance. Instagram is controlled by its parent company, Meta, which pursues its economic interests and is under political-economic pressure from politicians and customers. To avoid responsibility, authenticity governance has been shifted to users. The flagging feature (Crawford & Gillespie 2016) is critical to consider, as it grants limited power and agency to users and enforces mechanisms of peer surveillance and mutual moderation, placing certain users above other users.

However, “authenticity” itself is a vague concept and a double-standard marketing term, as “inauthentic” behavior is promoted by algorithmic ranking while simultaneously prohibited by Instagram’s community guidelines. A bot police force empowered by the platform would have to be trained and act by a code of conduct. This form of authenticity governance would, however, be paid labor, which Instagram already abolished in the case of the independent third-party fact-checkers. As long as Instagram is a privatized capitalist platform, we will have contradictory authenticity governance, prioritizing political-economic interests over user interests and service for a good and safe societal public space.

Future Work

Future research should explore the evolving governance of “authenticity” in AI-driven social media environments, particularly as advanced AI-generated bots—capable of mimicking human behavior with increasing sophistication—challenge existing detection and moderation systems. The rise of AI-run fame-enhancing and porn bots may further blur the boundaries between human and non-human agency, complicating user-led bot policing and undermining the very notion of “authenticity” as a stable category. Additionally, the impact of Instagram’s community notes on user-led bot policing remains underexplored. As a platform-driven alternative to third-party fact-checking, community notes may reshape how users identify and respond to inauthentic behavior—potentially reinforcing or disrupting existing power dynamics. Research should investigate whether community notes empower users or deepen asymmetries by distributing responsibility without accountability. Longitudinal studies could track how bot policing practices adapt to algorithmic changes, platform policies, and emerging technologies. Comparative analyses across platforms and cultural contexts would also reveal how authenticity governance varies in non-Western, non-democratic, or non-capitalist digital environments. Ultimately, future work must critically examine whether user-led moderation can be reimagined as a transparent, equitable, and ethically grounded practice—rather than a privatized, unregulated form of digital vigilantism.

References

- Crawford, K., & Gillespie, T. (2016). What is a flag for? Social media reporting tools and the vocabulary of complaint. *New Media & Society*, 18(3), 410–428.
<https://doi.org/10.1177/1461444814543163>.
- Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In *The interpretation of culture: Selected essays* (pp. 3–30). Basic Books.
- Gillespie, T. (2019). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
<https://doi.org/10.12987/9780300235029>.

- Favarel-Garrigues, G., Tanner, S., & Trottier, D. (2020). Introducing digital vigilantism. *Global Crime*, 21(3–4), 189–195. <https://doi.org/10.1080/17440572.2020.1750789>
- Kaplan, J. (2025, January 7). More speech and fewer mistakes. *Meta*. <https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/>.
- Laestadius, L. (2016). Instagram. In L. Sloan & A. Quan-Haase (Eds.), *The SAGE handbook of social media research methods* (pp. 571–592). SAGE Publications Ltd. <https://doi.org/10.4135/9781473983847>.
- Leistert, O. (2016). Soziale Medien als Technologien der Überwachung und Kontrolle. In J. Schmidt & M. Taddicken (Eds.), *Handbuch soziale Medien* (pp. 571–592). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-03895-3_13-1.
- Lindquist, J., & Weltevrede, E. (2024). Authenticity governance and the market for social media engagements: The shaping of disinformation at the peripheries of platform ecosystems. *Social Media + Society*, 10(1), 1–14. <https://doi.org/10.1177/20563051231224721>.
- McMahon, L., Kleinman, Z., & Subramanian, C. (2025, January 7). Meta to replace ‘biased’ fact-checkers with moderation by users. *BBC News*. <https://www.bbc.com/news/articles/cly74mpy8klo>.

5A.2 The Epistemic Burden of AI Labels: Interface Governance and User Responsibility on Social Media

Piotr Sochoń (Shanghai Jiaotong University), Nadia Urban (East China Normal University)

Introduction

Recent regulatory initiatives and platform policies increasingly require the labeling of generative AI content. These labels are often framed as mechanisms of transparency, safety, or accountability. However, little attention has been paid to how such labels function in practice at the interface level, or how they redistribute interpretive responsibility between institutions, platforms, and users.

This paper examines AI-generated content labels as a form of sociotechnical mediation rather than a neutral informational cue. We argue that labeling does not resolve ambiguity about AI content, but instead shifts the burden of interpretation onto users, who must decide what an AI label means, how risky a piece of content is, and how they should respond to it. In doing so, labeling normalizes uncertainty and produces new expectations of user competence, vigilance, and responsibility.

By focusing on labeling as an interface-level practice, this study contributes to debates on platform governance, sociotechnical imaginaries, and AI regulation by showing how governance is enacted not only through law or policy, but through everyday interactions with platform design.

Objectives/Research Questions

This paper introduces the concept of the epistemic burden of AI labels to capture how interface-level labeling shifts interpretive responsibility from institutions to users.

This study addresses the following research questions:

How do AI-generated content labels function as interface-level governance mechanisms on social media platforms?

How do labeling practices redistribute epistemic responsibility and uncertainty between regulators, platforms, and users?

How do different governance regimes shape the meanings, expectations, and affective responses associated with AI labels?

Methods

The study adopts a comparative qualitative and computational approach. It examines AI labeling practices on Instagram and Xiaohongshu, two platforms embedded in distinct regulatory and governance contexts.

The analysis draws on platform documentation, regulatory texts, interface analysis, and a dataset of labeled posts and user comments. Computational methods are used to identify patterns in labeled content and user engagement, while qualitative analysis focuses on how creators and audiences interpret, contest, and respond to AI labels in practice.

This mixed-methods approach allows the study to trace how labeling practices are operationalized by platforms and enacted by users in different sociotechnical contexts.

Results

The findings show that AI labels do not function as straightforward indicators of authenticity or risk. Instead, they act as ambiguous signals that require users to perform additional interpretive work. Across both platforms, labels shift responsibility onto users, who must decide how to assess credibility, legitimacy, and intent without clear guidance. This produces what we conceptualize as an epistemic burden, where users are expected to make judgments without adequate informational or institutional support.

The comparative analysis reveals that different governance configurations shape how this burden is experienced. Despite differences, in both contexts users become the primary site where uncertainty is managed. This demonstrates that transparency measures can paradoxically intensify, rather than resolve, epistemic uncertainty.

Future Work

This paper represents an initial step in theorizing AI labeling as a mediating practice rather than a purely technical intervention. Future work will extend this analysis by examining additional platforms and labeling regimes, as well as other interface-level signals that organize user interpretation and responsibility, such as content warnings, credibility indicators, or ranking cues. The concept of epistemic burden will be further developed to analyze how different governance configurations redistribute uncertainty and responsibility across platforms, institutions, and users.

5A.3 Governing Social Media Harms: A Critical Realist Extension of the Theory of The Firm

Robert Pavlovich, Bertrand Guillotin (Fox School of Business, Temple University)

Introduction

Minimally regulated social media platforms (Lubin et al., 2024) host an estimated 5.6 billion individual users (68.7 percent of global population) while generating market capitalizations that exceed the GDP of many sovereign nations (Statista, 2025; Dyvik, 2025). As public companies, their governance structures and principal-agent relationships create strong incentives to prioritize shareholder interests and market value (Friedman, 1970; Jensen & Meckling, 1976). However, this profit maximization mandate also contributes to systemic digital harm, magnified by artificial intelligence, that shapes the economic, social, and political spheres of society. These digital harms manifest themselves through the dissemination of hate speech, disinformation, discrimination, and (foreign) interference in democratic processes (Besen & Verveer, 2021).

Neoclassical economic theories (Table 1) treat negative externalities (digital harm) as accidental and negligible uncharged disservices rendered to third parties (Pigou, 1920; Coase, 1937, 1960; Kapp 1950; Williamson, 1971). We argue that this dominant and rival explanation (Figure 1) for digital harm as “accidental and negligible” is no longer adequate. We propose a conceptual model with a series of testable propositions (Figure 2) that extend the theory of the firm through a critical realist lens by conceptualizing digital harm as an ex-ante design element of the business model. Our model offers a theoretical foundation for academics and policymakers with insights into policy levers and corrective actions when traditional Pigouvian and Coasean solutions are rendered inadequate in today’s digital economy.

Digital harm manifests itself through four foundational components of the social media business model. The inputs in the form of user data, the optimization of user behavior through algorithmic driven engagement, corporate governance with a fiduciary mandate of shareholder primacy (Friedman, 1962), and the statutory frameworks and societal norms in which the platform operates. The problem consists of three key principles associated with 1) Transaction Cost Economics (TCE), 2) shareholder theory, and 3) regulations. Firstly, TCE fails to explain the social media business model where the platform is incentivized to monetize social harms to drive engagement and scalability. Secondly, the foundation of shareholder theory argues that “there is one and only one social responsibility of business—to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game”

(Friedman, 1962, p. 112). Thirdly, the current regulatory and judicial understandings of negative externalities rely on traditional economic theory, from Pigouvian taxes to Coasean bargaining. These are no longer adequate to address social costs generated as structurally incentivized outputs of engagement created by today's omnipotent and growing social media platforms.

Objectives/Research Questions

The primary aim of this qualitative research is to present a framework for a theoretical extension of the economic theory of the firm that addresses the structurally incentivized production of social costs within the social media platform business model.

This goal encompasses four objectives:

1. To establish the empirical basis for digital harm (Study 1) through directed content analysis (DCA). This pilot study seeks to systematically map the observable social costs created and shifted to the political, economic, and civil spheres of society. These represent the actual events requiring causal explanation.
2. To analyze the generative mechanisms of algorithmic engagement (Study 2, Part A) through deductive qualitative analysis (DQA) to isolate the underlying mechanisms driving user engagement. This objective focuses on the prioritization of engagement metrics leading to the coproduction of harmful content that supports value creation.
3. To analyze the perceived fiduciary mandate of corporate executives (Study 2, Part B) where the principal-agent relationship creates a path dependency that disincentivizes content moderation creating a barrier to a civic responsibility to mitigate social costs.
4. To synthesize the findings from Study 1 and Study 2 (Discussion) to propose an extension of the theory of the firm that explains social costs as structural and functional outputs. The objective is to provide a testable framework for regulatory interventions that target the generative mechanisms rather than the symptomatic results of digital harm on society.

Research Questions

To achieve the objectives noted above, this paper addresses the following four research questions covering observable harm, platform operations, managerial implications, and theory extension:

1. (Study 1 – Descriptive Phase): How do digital harms associated with social media platforms manifest across the political, economic, and civil spheres of society?
2. (Study 2 – Causal Phase A): By what mechanisms does the algorithmic-driven engagement retention engine incentivize the prioritization of toxic content to increase advertising revenue?

3. (Study 2 – Causal Phase B): How does the principal-agent relationship and the fiduciary duty to shareholder value create a financial liability for content moderation, and what testable propositions does this framework offer for evidence-based policymaking?
4. (Discussion/Synthesis): How can the theory of the firm be extended to model a business structure where the systemic production of negative social costs is fundamentally linked to the perceived profit-maximization mandate?

Methods

This methodology supports accuracy of the findings and validity of the conclusion by delineating the scope and boundaries of the research (Cuevo-Cazurra et.al., 2016). It establishes the philosophical stance, data collection strategy, and the analytical strategy. It defines the research parameters for a two-phase qualitative critical case study on a globally dominant, engagement-driven, multinational digital platform (Facebook). It maps the theoretical and conceptual frameworks to Bhaskar's (1975/2008) critical realist ontology as shown in Figure 3, where "[m]echanisms, events and experiences [...] constitute three overlapping domains of reality [...] the domains of the real, the actual and the empirical" (p. 56). Bhaskar's ontology posits that the domain of the real cannot be observed directly. Epistemologically, our understanding of the real causal mechanisms of a phenomenon are inferred from our experiences and observations that are shaped by actual events whether we can observe them or not. Employing a critical realist epistemology enables deductive qualitative analysis by making use of retrodution to infer "[w]hat must exist in order for these data and these patterns to have been observed" (Olsen, 2010, p. xx). This approach explores the gap between the firm's socially constructed narratives presented in its public facing documents, and the underlying proprietary mechanisms that are hidden behind trade secrets and intellectual property and actively managed as part of the fiduciary mandate of shareholder primacy.

The data collection strategy is intentional and selective through multi-source purposive sampling of documents that are most relevant to the research questions (Yin, 2018). Documentary evidence has several strengths related to stability and unobtrusiveness. It contains many specific details of events or information and can cover a broad span of time. However, it is understood that unknown reporting biases may exist based on the intent of the author, and researcher reflexivity is critical as bias may impact construct validity if document collection is incomplete and narrowly focused (Yin, 2018). For the initial pilot study, we sourced 16 public documents from 6 source categories during the period of 2017-2025. This serves as a feasibility check and instrument validation to identify flaws in the research design. The documents are assessed for adequacy against the a priori coding framework by reviewing code productivity (Figure 4). The "purpose of the analytic strategy is to link case study data to important concepts of

interest, and then have the concepts provide a sense of direction in analyzing the data” (Yin, 2018, p. 223). This is accomplished by reflecting on the research questions, propositions, theoretical and conceptual frameworks to guide the case study analysis. This strategy allows preexisting and rival theories to be explored to highlight the anomalies and then conjecture about opportunities to extend the theory to avoid them.

Results

The preliminary findings from the code co-occurrence analysis (Figure 5) and thematic co-occurrence analysis (Figure 6) in Study 1 indicates that social media generates broad tensions across all three spheres of society—economic, social, and political.

Regarding the problem statement, where social media platforms don’t produce any physical products but seem to generate global reach & economic power while creating a broad threat not only to democratic institutions but to society itself, the findings suggest:

1. The U.S. government sees Facebook as a “monopoly” in the market and is concerned with data privacy and the potential for misuse of user information.
2. U.S. regulations provide social media platforms immunity for third-party content, yet lawmakers are concerned that platforms enable foreign and domestic actors to spread misinformation and disinformation and influence elections.
3. The platform views U.S. government efforts to limit the capture and targeting of user data and/or moderate content as having an adverse effect on its ability to generate advertising revenue.

Regarding the theoretical gaps in the literature related to social media firm boundaries, externalities, and social costs, the findings suggest:

1. The porous nature of platform boundaries, driven by algorithms that mediate information exchange, extend the platform’s influence beyond its contractually defined perimeter.
2. Digital harms are produced by the platform and amplified by its algorithms requiring a broader theoretical framework that supports practical solutions for corrective actions.
3. The study supports the institutionalist concept of social costs as systemic to the platform business model that creates harms to mental health and democratic foundations.

Regarding the knowledge gaps in the literature related to how government regulatory frameworks mitigate the social costs of social media, the findings suggest:

1. An obsolete regulatory environment in the U.S. enables platforms to host harmful content without being liable due to federal immunity protections as a non-

publisher.

2. Negative externalities are not accidental and negligible but incentivized by the platform to increase engagement and maximize profit.

Regarding the practical gaps in the literature related to the best approaches to bring the social costs of social media into equilibrium with the social benefits, the finding suggest:

1. A new normative framework would inform the development of design-level interventions, governance and policy strategies, and user-focused initiatives, that drive platform accountability and address ethical & civic responsibilities to address harms while protecting social benefits.

Future Work

The proposed critical realist extension of the theory of the firm invites academic discourse and supports further research opportunities in the field of platform regulation and strategy. Planned future work includes a second study, currently in progress, that employs deductive qualitative analysis in a causal explanatory phase to identify the generative mechanisms through which digital harms emerge as functional outputs of the social media business model. This supports arguments for moving the debate from controls over content management to solutions that directly address the economic drivers of social costs that equate the interdependence of technology and society with the civic responsibility of social media platforms as user data fiduciaries. Additional opportunities exist for cross-disciplinary research on the power of very large online platforms including integrating perspectives from psychology, sociology, economics, and political science, particularly in areas related to geopolitical power, corporate social responsibility and corporate citizenship efforts that influence firm boundary decisions. The topic of data governance, where data is considered a critical and strategic asset for digital platforms, would benefit from further exploration as it becomes central to the platform business model, shapes the firm boundary, and enables the accumulation of geopolitical power and control across world.

While this research focuses on descriptive and causal analysis of the phenomenon, a normative element could address ethical principles and support arguments on interpretation of social costs, the impact on society, and the appropriate actions or policy decisions necessary to mitigate negative outcomes (Viens, 2019). In terms of economic ethics, should profit maximization be limited by social responsibility to mitigate digital harm? For social ethics, do social media platforms intentionally target younger populations, and do they shoulder a higher percentage of digital harm? When considering political ethics, how do governments enact controls over harmful content and foreign influence campaigns while respecting universal rights on freedom of expression? Normative analysis is needed to integrate these ethical spheres of society,

support and enhance qualitative document analysis and align them with the principles of polyvocality that provide multiple interpretations of the data from different perspectives (Thimm et al., 2017). This would provide a more nuanced understanding of these complex techno-social issues.

References

- Bhaskar, R. (2008). A realist theory of science. Routledge. (Original work published 1975). <https://doi.org/10.4324/9780203090732>
- Beknazar-Yuzbashev, G., Jimenez-Duran, R., McCrosky, J., & Stalinski, M. (2025). Toxic content and user engagement on social media: Evidence from a field experiment (Working Paper No. 741). Centre for Competitive Advantage in the Global Economy, Department of Economics, University of Warwick. <https://warwick.ac.uk/fac/soc/economics/research/centres/cage/manage/publications/wp741.2025.pdf>
- Besen, S. M., & Verveer, P. L. (2021). Section 230 and the problem of social cost. *Journal of Law and Policy*, 30(1), 68–122. <https://brooklynworks.brooklaw.edu/jlp/vol30/iss1/2>
- Boudreaux, D. J., & Meiners, R. (2019). Externality: Origins and classifications. *Natural Resources Journal*, 59(1), 1–34.
- Coase, R. H. (1937). The nature of the firm. *Economica*, 4(16), 386–405. <https://doi.org/10.1111/j.1468-0335.1937.tb00002.x>
- Coase, R. H. (1988). The problem of social cost. In *The firm, the market and the law* (pp. 95–157). University of Chicago Press. (Original work published 1960). <https://www.law.uchicago.edu/sites/default/files/file/coase-problem.pdf>
- Cuervo-Cazurra, A., Andersson, U., Brannen, M. Y., Nielsen, B. B., & Reuber, A. R. (2016). Can I trust your findings? Ruling out alternative explanations in international business research. *Journal of International Business Studies*, 47(8), 881–897. <https://doi.org/10.1057/jibs.2016.31>
- Dyvik, E. H. (2025). Biggest companies in the world by market value 2024. Statista. <https://www.statista.com/statistics/263264/top-companies-in-the-world-by-market-capitalization/>
- Federal Trade Commission. (2024). A look behind the screens: Examining the data practices of social media and video streaming services. https://www.ftc.gov/system/files/ftc_gov/pdf/Social-Media-6b-Report-9-11-2024.pdf

- Friedman, M. (1962). *Capitalism and freedom*. University of Chicago Press.
- Friedman, M. (1970, September 13). A Friedman doctrine: The social responsibility of business is to increase its profits. *The New York Times*.
<https://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html>
- Golder, S. A., & Macy, M. W. (2014). Digital footprints: Opportunities and challenges for online social research. *Annual Review of Sociology*, 40(1), 129–152.
<https://doi.org/10.1146/annurev-soc-071913-043145>
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305–360.
[https://doi.org/10.1016/0304-405X\(76\)90026-X](https://doi.org/10.1016/0304-405X(76)90026-X)
- Jullien, B., & Sand-Zantman, W. (2021). The economics of platforms: A theory guide for competition policy. *Information Economics and Policy*, 54, Article 100880.
<https://doi.org/10.1016/j.infoecopol.2020.100880>
- Lubin, N., Mayberry, K., Moses, D., Revel, M., Thorburn, L., & West, A. (2024). Mapping the space of social media regulation. *MIT Science Policy Review*, 5, 1–12.
<https://doi.org/10.38105/spr.22y8h1p7>
- Meade, J. E. (1952). External economies and diseconomies in a competitive situation. *The Economic Journal*, 62(245), 54–67. <https://doi.org/10.2307/2227173>
- Meta Platforms Inc. (2023). Form 10-K for the fiscal year ended December 31, 2022. U.S. Securities and Exchange Commission.
<https://d18rn0p25nwr6d.cloudfront.net/CIK-0001326801/5f2c456b-c658-4f08-9ce1-46b5d46ab4da.pdf>
- Olsen, W. K. (2010). Editor's introduction: Realist methodology: A review. In W. K. Olsen (Ed.), *Realist methodology* (Vol. 1, pp. xix–xlvi). Sage.
- Pigou, A. C. (1920). *The economics of welfare*. Macmillan and Co.
<https://archive.org/details/cu31924073868113/>
- Rochet, J.-C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4), 990–1029.
<http://www.jstor.org/stable/40005175>
- Statista Research Department. (2025). Number of internet and social media users worldwide 2025. Statista. <https://www.statista.com/statistics/617136/digital-population-worldwide/>
- U.S. Senate, Committee on Commerce, Science, and Transportation. (2021). Protecting kids online: Testimony from a Facebook whistleblower (S. Hrg. 117-769). U.S.

Government Publishing Office. <https://www.congress.gov/117/chrg/CHRG-117shrg54110/CHRG-117shrg54110.pdf>

U.S. Surgeon General. (2023). Social media and youth mental health: The U.S. Surgeon General's Advisory. U.S. Department of Health & Human Services. <https://www.hhs.gov/sites/default/files/sg-youth-mental-health-social-media-advisory.pdf>

Williamson, O. E. (1971). The vertical integration of production: Market failure considerations. *The American Economic Review*, 61(2), 112–123.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.

Session 5B - Use and Users

Day 1 · Mon 13 July · 16:30–17:30, Seminar Suite A, ARC

5B.1 To understand and to be understood: Young people's perceptions of peer support for mental health in social media

Essi Holopainen, Meri Kulmala, Anu Katainen, Anna-Maija Multas (University of Helsinki)

Introduction

Young people increasingly seek support for their mental health challenges in online environments, where daily struggles and experiences of living with mental health issues become shared and negotiated with peers. In the context of mental health, the “peerness” can include similar experiences, symptoms or diagnoses, or a shared understanding of what it means to live with mental illness (Rayland & Andrews 2023; Fortuna et al. 2022). Supportive interaction among peers have been studied to contain informational, emotional, and esteem-boosting elements (Dennis 2003; Chang 2009; Lin & Kishore 2021; Yan & Tan 2014), which can vary from the expressions of empathy and trust (Lin & Kishore 2021) to providing valuable information about help-seeking practices and everyday coping strategies (Rayland & Andrews 2023; Naslund et al. 2016) and to decreasing the stigma around various issues through personal empowerment (ibid.; Prescott et al. 2020). What has remained less studied is how peer support is formed on currently popular social media platforms.

While social media platforms have developed rapidly towards algorithmic-driven flow of short-form videos (Meier & Krause 2023), the interactions with other users and their content do not only emerge from the social connections within already existing networks, but often through the algorithm itself (Gerbaudo 2024; Sujon et al. 2025). Previous studies have often classified social media use as active or passive activity (Ellison et al. 2020; Masciantonio et al. 2021; Frison & Eggermont 2020), separating activities such as content creation, commenting and messaging from browsing content without social interaction (Beyens et al. 2024; Verduyn et al. 2017; Godard & Holtzman 2024; Kaye et al. 2021). Yet, passive browsing has been also challenged (Meier & Krause 2023, 174), because “digital listening” requires leaning forward, paying attention, and selective exposure to the content, including a diverse range of motivations (ibid., also Ellison et al. 2020, 417). Studying young people's perceptions on supportive interaction with the content and other users on social media, as well as the support-creating

features of the platforms (see Beyens et al. 2024), offers therefore a great opportunity to understand better, how peer support is both received and given on social media.

Objectives/Research Questions

In this study, we examine how young people use social media to support their own wellbeing, both as receivers and providers of peer support. Our research questions are:

- 1) How do young people themselves define and perceive peer support in social media?
- 2) How, according to young people, is peer support formed within these platforms?

We propose a conceptual model where we combine the theory-informed forms of peer support with the reactive and productive practices of young people in social media (see Dennis 2003; Chang 2009; Beyens et al. 2024; Kaye 2021). Our model also considers the public and private aspects of peer support (Beyens et al. 2024; Valkenburg et al. 2022). Our contribution to the conceptual discussion is both demonstrating the interconnections between the types of online practices and the forms of peer support among young people, and in challenging the notion of “passive” social media use.

Methods

To allow a youth-centered perspective to social media based mental health peer support, this study was carried out as co-research, drawing on qualitative datasets co-produced together with young people through several types of interviews, and involving young people as co-researchers.

Data co-production was conducted by the co-research team (including professional researchers and young co-researchers) from September 2023 to February 2024, resulting in various types of data. Co-researchers and group interview participants were recruited based on the criteria of being aged 18–29 and demonstrating an interest in mental health and social media issues. The data used in this study consists of two focus group interview transcripts from the recorded sessions (GI1-2, N=6) and three sets of peer interviews (PI1-3, both individual and group interviews) conducted by five young co-researchers from the topic of their choice (N=11). The produced data and findings were discussed in the workshops with the research team, which enabled a collective dialogical approach in the analysis.

We analyzed the data using abductive content analysis, as it combines both inductive and deductive features and aims at generating novel theoretical insights (Timmermans & Tavory 2012). This method allowed us to build the chosen theoretical concepts further and thus to form our conceptual proposal about the formation of peer support in social media. Building first on prior classifications of peer support and online activity, we later critically re-examined the data to interpret anomalous cases through a synthesis of theoretical models (ibid.).

Results

According to our results, peer support of young people on social media is a combination of different elements and practices, which previous classifications of peer and social support in both offline and online contexts were able to capture only partly. We found that the key forms of peer support in social media are emotional, informational, and affirmational support, which all require a certain level of relatedness perceived by young users, because social media content can be helpful and supportive without being perceived as peer support. These forms of peer support are also interconnected, affecting one another, and situational, perceived as peer support when needed.

We also found that peer support is formed in social media through reactive and productive practices, which have relatable informational, emotional, and affirmational elements. These practices include both public and private ways to communicate in social media, and the various ways to receive and give peer support to others should be considered equally important, as well as seen intertwining. Reactive practices are the personal experiences of receiving peer support from others on social media — the moments when young people feel supported and react to the support with likes or comments. Productive practices in turn refer to young people's intentional interaction to give peer support to others on social media, either by sharing others' content forward or creating their own supportive content.

Summarizing our results, we see that young people use social media to receive and provide peer support in personal and situational ways, according to their needs and what they consider relatable. Relatable content in our data is something which pops up from the content flow, makes a young person to pause and watch, and creates a feeling of recognition through similarities, either in oneself (receiving peer support) or in others (giving peer support). Despite individual differences and preferences, relatable content often includes shared identity (gender, age), situational aspects (similar struggles or life situations) and aspects related to style (humor, communication style). Hence, we argue that relatability is the element that turns knowledge into know-how, empathy into understanding, and motivation into empowerment.

Our study challenges the notion of passive social media use in relation to peer support even when this use is not always publicly visible to larger audiences. On the contrary, our results confirm that a large part of young people's activities in social media, especially related to private and reactive use, remain hidden unless the research questions are framed in a youth-centered way (see Beyens et al. 2024). Our conceptual contribution may help future research to recognize young people's engagement with social media content without falling into the non-nuanced distinction of active and passive forms of use.

Future Work

Social media as multifaceted operational environment offers various features for social interaction and community building and allows young people to learn from each other, participate in normalizing talk about mental health, and feel less alone. Our research has examined young people's perceptions of peer support on social media without focusing on any specific social media platform, which is both an advantage and disadvantage. We managed to name the common denominators of peer support on different kinds of platforms, as perceived by young people, but simultaneously we were not able to focus on the specific affordances of different content formats, such as videos, posts with text and pictures, podcasts, or group-chat messages. Yet we found that young people consider it supportive when content creators discuss difficult topics "with their own faces". Therefore, we encourage future research to investigate these potential distinctions through participatory co-research methodologies and focus on platform-specific practices of young people.

In our future work, we will focus thematically on online help-seeking, but we will expand the context from social media platforms to online chat counselling services.

References

- Beyens, I., Pouwels, J. L., van Driel, I. I. et al. (2024). Social Media Use and Adolescents' Well-Being: Developing a Typology of Person-Specific Effect Patterns. *Communication Research*, 51(6), 691-716.
<https://doi.org/10.1177/00936502211038196>
- Chang, H-J. (2009). Online supportive interactions: Using a network approach to examine communication patterns within a psychosis social support group in Taiwan. *Journal of the American Society for Information Science and Technology*, 60(7), 1504-1517.
- Dennis, C-L. (2003). Peer support within a health care context: a concept analysis. *International Journal of Nursing Studies*, 40(3), 321-332.
[https://doi.org/10.1016/S0020-7489\(02\)00092-5](https://doi.org/10.1016/S0020-7489(02)00092-5).
- Ellison, N. B., Triêu, P., Schoenebeck, S. et al. (2020). Why we don't click: Interrogating the relationship between viewing and clicking in social media contexts by exploring the "non-click". *Journal of Computer-Mediated Communication*, 25(6), 402-426.
- Frison E. & Eggermont S. (2020). Toward an integrated and differential approach to the relationships between loneliness, different types of Facebook use, and adolescents' depressed mood. *Communication Research*, 47(5), 701-728.

- Fortuna, K.L., Solomon, P. & Rivera, J. (2022). An Update of Peer Support/Peer Provided Services Underlying Processes, Benefits, and Critical Ingredients. *Psychiatr Q* 93, 571–586. <https://doi.org/10.1007/s11126-022-09971-w>.
- Gerbaudo, P. (2024). TikTok and algorithmic transformation of social media publics: From social networks to social interest clusters. *New Media & Society*. Advance online publication. <https://doi.org/10.1177/14614448241304106>.
- Gliniecka, Martyna (2024). *Youth Digital Health and Online Platforms*. Routledge Studies in Science, Technology and Society. New York, Routledge.
- Godard, R. & Holtzman, S. (2024) Are active and passive social media use related to mental health, wellbeing, and social support outcomes? A meta-analysis of 141 studies. *Journal of Computer-Mediated Communication*, 29(1).
- Kaye, L. (2021) Exploring the “socialness” of social media. *Computers in Human Behavior Reports*, (3)2021, 100083.
- Kulmala, M., Venäläinen, S., Hietala, O., Nikula, K., & Koskivirta, I. (2024). Lived Experience as the Basis of Collaborative Knowing. Inclusivity and Resistance to Stigma in Co-Research. *International Journal of Qualitative Methods*, 23. <https://doi.org/10.1177/16094069241236271>.
- Lin, X. & Kishore, R. (2021). Social media-enabled healthcare: A conceptual model of social media affordances, online social support, and health behaviors and outcomes. *Technological Forecasting and Social Change*, 166, 120574.
- Masciantonio, A., Bourguignon, D., Bouchat, P. Et al. (2021). Don't put all social network sites in one basket: Facebook, Instagram, Twitter, TikTok, and their relations with well-being during the COVID-19 pandemic. *PLoS ONE* 16(3): e0248384. <https://doi.org/10.1371/journal.pone.0248384>.
- Meier, A. & Krause, H.-V. (2023) “Does Passive Social Media Use Harm Well-Being?: An Adversarial Review”. *Journal of Media Psychology*, 35(3), 169–180. Doi: 10.1027/1864-1105/a000358.
- Naslund, J. A., Aschbrenner, K. A., Marsch L. A. & Bartels, S. J. (2016). The future of mental health care: peer-to-peer support and social media. *Epidemiol Psychiatr Sci.* 25(2):113-22. Doi: 10.1017/S2045796015001067.
- Naslund, J. A., Bondre, A., Torous, J. et al. (2020). Social Media and Mental Health: Benefits, Risks, and Opportunities for Research and Practice. *J. technol. behav. sci.* 5, 245–257. <https://doi.org/10.1007/s41347-020-00134-x>.
- Prescott, J., Rathbone, A. L. & Brown, G. (2020). Online peer to peer support: Qualitative analysis of UK and US open mental health Facebook groups. *DIGITAL HEALTH*, 6. doi:10.1177/2055207620979209

- Rayland, A. & Andrews, J. (2023). From Social Network to Peer Support Network: Opportunities to Explore Mechanisms of Online Peer Support for Mental Health. *JMIR Ment Health* 2023;10:e41855.
- Sujon, Z., Dyer, H. T., & Bonow Soares, F. (2025). Social Media and Society: Platforms, Publics, and Anti-Publics. *Social Media + Society*, 11(3).
<https://doi.org/10.1177/20563051251368234>
- Timmermans, S., & Tavory, I. (2012). Theory Construction in Qualitative Research: From Grounded Theory to Abductive Analysis. *Sociological Theory*, 30(3), 167-186.
<https://doi.org/10.1177/0735275112457914>
- Valkenburg, P.M., Meier, A. & Beyens, I. (2022) Social media use and its impact on adolescent mental health: An umbrella review of the evidence. *Current Opinion in Psychology*, 44, 58-68. <https://doi.org/10.1016/j.copsyc.2021.08.017>.
- Verduyn, P., Ybarra, O., Résibois, M. et al. (2017). Do Social Network Sites Enhance or Undermine Subjective Well-Being? A Critical Review. *Social Issues and Policy Review*. 2017;11(1):274–302.
- Yan, L. & Tan, Y. (2014). Feeling Blue? Go Online: An Empirical Study of Social Support Among Patients. *Information Systems Research*, 25(4), 690–709.
doi:10.1287/isre.2014.0538.

5B.2 The AI Citizenship Framework: Developing interactive AI literacies for critical and proactive citizens

*Elinor Carmi (City, University of London), Anna Feigenbaum (Glasgow University),
Photini Vrikki (UCL)*

Introduction

As AI rapidly reshapes everyday life, a wave of AI literacy frameworks has emerged to help people navigate and use these systems more effectively (Lee et al., 2021; Laupichler et al., 2022; Ng et al., 2021). However, the majority of these frameworks emerge from computer science fields, with far fewer contributions from the social sciences fields, where issues of media and data literacy have historically been central. The predominantly utilitarian orientation of these existing AI literacy frameworks limits them to improving skills and creating better workers, without equipping people to act as citizens that can challenge, participate and propose alternatives to existing and future AI practices and technological developments. By neglecting the skills and knowledge needed to be able to actually critically evaluate, make decisions about, or challenge existing social power structures, these frameworks strip people of their autonomy and citizenship capacity. Our HumAlne project was conceived to address this omission and advance an alternative approach to AI literacy that foregrounds citizenship, rights, and collective empowerment.

Objectives/Research Questions

The objective of the HumAlne project was to address this literacy gap by developing a novel framework for AI literacy from the perspective of AI citizenship. Practically, our aim was to deliver an interactive toolkit of resources to create citizen led campaigns that prioritise people's rights and wellbeing.

Methods

Current AI frameworks, as mentioned above, mainly emerge from human computer interaction design and behavioural science fields, with little to no engagement with the extensive scholarship on media, digital, data, or algorithmic literacies. To address this, as a first step, we developed a rough translation of the Data Citizenship framework (Carmi and Yates, 2023) into AI Citizenship. The Data Citizenship framework views data literacies as the capability (Nussbaum, 2002; Sen, 2009) of citizens to do, think and participate with their data socially and politically. From the Data Citizenship framework, we translated the three different components into AI use and established the first version of our AI Citizenship framework that looked like this:

1. AI Doing - Incorporates practical skills that involve access, ethical use and interpretation of AI. These are a set of core AI literacy skills essential for thoughtful and informed engagement with AI in the AI driven society. Actions that are part of this dimension are:

- Accessing - The ability to search for, identify and access AI services.
- Assessing - The ability to identify the quality and credibility of the AI output.
- Ethical use - The ability to use AI ethically, e.g. not manipulating data for illegal or harmful purposes.

2. AI Thinking - Incorporates mindful, critical and reflective skills and engagement as people view and analyse the world through AI. Actions that are part of this dimension are:

- Awareness of data and AI protection rights - Being aware of local (UK) or regional (The European Union) AI regulations what are the repercussions of them.
- Understanding the AI society - The ability to understand the way the AI economy works, including specifics of how AI companies are funded, how their computation practices affect the environment, and how people's data and IP work is affected by this.
- AI Safety - The ability to consider and implement protective steps when using AI outputs.

3. AI Participating - People's ability to investigate, challenge, negotiate, protest, act upon and transform the structure of AI systems and services for themselves and/or their communities. Actions that are part of this dimension are:

- Participating in society using AI - The ability to use AI for societal participation and civic action.
- AI activism - The ability to proactively protect individual and communities wellbeing in the AI society. This includes the ability to collectively promote and exercise digital/data/AI rights, as well as the ability to object/resist/modify the hegemonic way of using AI services.
- Supporting others with their AI literacy - The ability to help others with their AI literacy.

Results

To test this framework, we applied it to five real-world use cases reflecting challenges citizens encounter with AI. These were developed as interactive scenarios powered by small-scale language models, enabling users to curate their own repositories through Notebook LM. While acknowledging the exploitative and extractivist nature of AI systems and services, our AI Citizenship framework aims to use this technology in order to design a society-centred, community-oriented set of AI Open Educational Resources which will challenge social power structures.

Based on our AI Citizenship framework, we then built a prototype called Play-Aible Ethics, a game-based, hands-on learning experience and related toolkit resources. The

platform operationalises our three-dimensional framework for AI citizenship:

- 1) AI Doing (practical skills): Users practice prompt engineering, source checking, and ethical use through conversational tasks, not lectures.
- 2) AI Thinking (critical analysis): Users examine power structures around AI—who benefits, who bears risks, and how decisions are made—developing critical awareness (“conscientização”) through guided inquiry and real-world cases.
- 3) AI Participating (civic action): Learning becomes action as users produce locally relevant materials (e.g., parent guides, community briefings) that help others navigate AI responsibly.

Instead of lectures or technical training, Play-Alble Ethics embeds real world cases in interactive stories, with missions that people play through at their own pace. Participants are guided by a set of expert characters to help them develop more informed and ethical, real-world, task-based outputs using generative AI. Along the way, they build practical skills for working with AI, learn how to question and challenge AI systems, and explore how they can take part in shaping how these technologies are used.

To operationalise this, ‘citizen players’ input a task or topic that they want to use AI to help them write about (i.e. a report, a campaign briefing, a news article) and progress through a clear journey: Mission Briefing → First Contact → Verification → Critical Analysis → Taking Action → Mission Complete. Along the way, they earn XP, unlock characters, and advance a narrative that builds motivation and confidence. Through each mission AI literacy is embedded into users' reflections and decision-making, culminating in a real output.

Future Work

We have a functional prototype with all six levels operational. It has been tested internally and refined with feedback from colleagues and a small group of educators. The technical infrastructure is working, and the pedagogical framework is established. The next step is to open playtesting up to the public, through a live play-and-learn labs run in partnership with Third Sector Lab (Glasgow) and PIER (Bournemouth). Third Sector Lab is a pioneering provider of digital upskilling and AI skill development for a wide range of organisations spanning the health sector to city infrastructure. PIER is a public and community involvement group that facilitates interactive sessions with a wide diversity of groups and organisations.

Feedback from these play-and-learn labs will go into a final design round for Play-Alble Ethics, before the toolkit and software can be developed on a larger scale.

References

- Carmi, E., & Yates, S. (2023). Civic participation in the datafied society: Data citizenship and data literacies to challenge power imbalance between society and big tech. *International Journal of Communication*, 17, 3619- 3637.
- Laupichler, M. C., Aster, A., Schirch, J., & Raupach, T. (2022). Artificial intelligence literacy in higher and adult education: A scoping literature review. *Computers and Education: Artificial Intelligence*, 3, 100101.
- Long, D., & Magerko, B. (2020, April). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-16).
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041.
- Nussbaum, M. C. (2002). Capabilities and social justice. *International Studies Review*, 4(2), 123–135.
- Sen, A. K. (2009). *The idea of justice*. Harvard University Press.

5B.3 The Dynamics of Use, Motivations, Mindsets on Well-Being for Young Adults on TikTok

Kate Mays, Ellen Kozelka (University of Vermont)

Introduction

In the decades since social media has become a fixture in our civic, social, and personal lives, public discourse has oscillated between celebration and alarm. Platforms have been heralded for expanding social connection and democratizing content creation, while also being criticized for fostering distraction, dependency, and psychosocial distress (Berry et al., 2017; Keles et al., 2020). Consensus is still elusive, and perceptions of social media as “good” or “bad” often depend on who is asked and which outcomes are being considered (Sanchez and Jenkins, 2024). For instance, Pew Research’s recent report found that, while 48% of teens said that social media are bad for their peers, only 14% said that it was bad for them, personally; parents are twice as likely as teens to name social media as the primarily threat to teen mental health (Faverio et al., 2025). More teens say that social media helps, rather than hurts, their friendships and confidence, though the reverse is true for other well-being dimensions (Faverio et al., 2025).

Scrutiny of social media’s influence on wellbeing has intensified in recent years (Meier & Reinecke, 2021), especially as youth mental health concerns continue to rise: an estimated one in seven children aged 10-19 will experience some form of mental illness (WHO, 2025). Yet empirical findings are mixed, with meta-analyses showing that outcomes vary based on individual differences, contexts of use, and users’ beliefs about their relationships with technology (Ellison et al., 2022; Meier & Reinecke, 2021; Valkenburg et al., 2022).

Therefore, platform-agnostic assessments may be insufficient for capturing social media’s impact on well-being. Platforms’ diverse affordances enable different behaviors, norms, and experiences (Bucher & Helmond, 2018). While “older” social media sites like Facebook and Instagram are primarily structured around individuals’ social networks, newer platforms like TikTok rely more on individual interests and in-app behavior to shape user experience (Gerbaudo, 2024). This shift in platform logics, from social networks (via relationships) to social clustering (via interests) has implications for the way online communities are formed and experienced (Gerbaudo, 2024).

This study explores how young adults use TikTok and how use is related to their self-reported well-being. TikTok has grown rapidly in the last decade, already eclipsing Instagram and Snapchat in popularity among young adults (Gottfried & Park, 2025). A

short-video sharing platform, TikTok has distinguished itself based on its effective recommendation algorithm, which drives content sharing and user engagement. TikTok is so good at capturing people's attention that some politicians have described the platform as "digital fentanyl" (Hendrix, 2022). While hyperbolic, the metaphor speaks to TikTok's core mechanics: a personalized "For You Page," frictionless scrolling, and algorithmic reinforcement loops designed to keep users engaged

Objectives/Research Questions

Increasingly, it appears that social networks sites' (SNS) effects depend on what people do with social media, what their motivations are for using it, and how they interpret their experiences (Kross et al., 2021). Research has expanded beyond measures like frequency of use, which do not fully capture the subjective, individualized experiences on social media (Meier et al., 2024). This study draws from three complementary conceptual approaches to explore young adults' TikTok use and well-being: the active/passive model of SNS (what people do), uses and gratifications theory (what motivates use), and social media mindsets (how experiences are interpreted).

The active/passive model of SNS was developed in response to the limitations of broad frequency measures (Verduyn et al., 2017). Distinguishing active vs. passive use adds more explanatory power: passive SNS use (i.e., content consumption) corresponds to reduced well-being (Verduyn et al., 2015), while active SNS use (i.e., content engagement, creation) sometimes relates to improved well-being (Verduyn et al., 2015; Verduyn et al., 2017). The benefits of active use, however, depend on the type of content interaction (social, warm, targeted) and individual differences (Verduyn et al., 2022).

Uses and gratifications theory (UGT) holds that people actively and purposively select their media to fulfill social and psychological needs, such as escapism, information seeking, social connection, and identity expression (Katz et al., 1973; Ruggiero, 2000). The active/passive dichotomy might map some SNS motivations, like escapism or distraction, as more passive acts, while social connection and identity expression align more closely with active use. Some have critiqued these categorizations as too reductive: for example, using social media for escapism could be an intentional act of relaxation or a maladaptive procrastination response (A. Lee & Hancock, 2024; Meier et al., 2024). The effects of such use, therefore, depends more on the way in which the behavior is perceived by the actant.

Social media mindsets conceptualize the beliefs that people have about the way they use social media and the role it plays in their lives (A. Lee & Hancock, 2024). They comprise two dimensions: agency (perceived control over their social media use) and valence (perceived enhancements vs. harms of social media use) (A. Lee et al., 2023; A.

Lee & Hancock, 2024). Compared to other predictors, social media mindsets have stronger explanatory power for well-being (A. Lee & Hancock, 2024).

RQ1: What are the most salient motivations for using TikTok?

RQ2a-b: Are TikTok motivations related to (a) active/passive use and (b) social media mindsets?

H1a-b: Well-being will be (a) positively predicted by active and (b) negatively predicted by passive TikTok use.

H2a-b: Well-being will be positively predicted by (a) agentic and (b) enhancement social media mindsets.

RQ3: Is well-being influenced by TikTok motivations?

Methods

To explore young adults' well-being and TikTok motivations, use, and mindsets, we conducted a survey of students at a medium-sized state university in the northeast United States (N = 442). The online questionnaire was distributed via the survey platform Qualtrics. Participants were recruited through convenience and snowball sampling in August – November 2025 using flyers posted around campus, visits to classes, tabling, word-of-mouth, and e-mail listservs.

All variables were measured using multiple-item, 5-point, Likert-type scales. See Table 1 in the supplementary materials for full scale descriptives. TikTok motivations were adapted from E. Lee et al.'s (2015) Motivations for using Instagram scale and augmented based on Omar and Dequan's (2020) exploratory work on users' TikTok motivations, comprising six motivations: social interaction ($\alpha = .79$), self-expression ($\alpha = .75$), information-seeking ($\alpha = .74$), habit ($\alpha = .79$), escapism ($\alpha = .78$), and distraction (4 items, $\alpha = .83$). Active (6 items, $\alpha = .83$) and passive ($\alpha = .79$) use was adapted for TikTok from Gerson et al.'s (2017) Passive and Active Facebook Use Measure. A. Lee and Hancock's (2024) Social Media Mindsets Scale was used to measure agency ($\alpha = .74$) and enhancement ($\alpha = .73$) mindsets.

Well-being includes three separate measures. Problematic TikTok use ($\alpha = .91$) captures dependence on and addiction to TikTok (Günlü et al., 2023). The loneliness scale (Russell et al., 1978) asked about belongingness, social connection, and feeling understood ($\alpha = .92$). Self-esteem asked about feelings of worth, respect for, satisfaction with, and pride in oneself ($\alpha = .85$), measured using Rosenberg's (1965) self-esteem scale.

A descriptive analysis was conducted first to explore the salience of TikTok motivations in the sample and the relationships between motivations, use, and mindsets. To test the hypotheses, a series of multivariate linear regression models were constructed for

examining the extent to which motivations, use, and mindsets predicted the three well-being outcomes. All analyses were completed using IBM SPSS Statistics.

Results

RQ1 asked about young adults' most prevalent TikTok motivations (see Figure 1 in supplementary materials). The descriptive analyses show that TikTok is used most often for distraction ($M = 3.89$, $SD = .82$), followed by habitual use ($M = 3.46$, $SD = 1.08$), information-seeking ($M = 3.36$, $SD = .83$), and as a form of escapism ($M = 3.32$, $SD = .88$). The least salient motivations were using TikTok for social interaction ($M = 2.79$, $SD = .82$) and self-expression ($M = 1.92$, $SD = .92$).

RQ2 asked whether motivations, use, and mindsets were related to one another (see Table 2 in supplementary materials). Both active and passive use were positively correlated with all TikTok motivations. Active use was most strongly correlated with social interaction ($r = .63$, $p < .01$) and self-expression ($r = .68$, $p < .01$) motivations. Passive use was most strongly correlated with distraction ($r = .63$, $p < .01$) and information-seeking ($r = .60$, $p < .01$) motivations. A different dynamic emerged between social media mindsets and motivations. Those with more of an enhancement mindset were more likely to use TikTok for social interaction ($r = .29$, $p < .001$), self-expression ($r = .23$, $p < .001$), and information-seeking ($r = .31$, $p < .001$); there was no relationship between an enhancement mindset and habitual, distraction, or escapism motivations. Those with an agentic mindset were less likely to use TikTok out of habit ($r = -.38$, $p < .001$), as a distraction ($r = -.28$, $p < .001$), and for escapism ($r = -.28$, $p < .001$); there was no relationship between agentic mindsets and social interaction, self-expression, or information-seeking motivations.

To examine the relative influence of use (H1a-b), mindsets (H2a-b), and motivations (RQ3) on well-being, three multivariate linear regressions were run (see Table 3 in supplementary materials). Overall, the models explained 56% of the variance in problematic TikTok use ($F[10, 253] = 34.37$, $p < .001$); 15% of the variance in loneliness ($F[10, 215] = 5.00$, $p < .001$); and 9% of the variance in self-esteem ($F[10, 214] = 3.17$, $p < .001$). More problematic TikTok use was predicted by passive use ($\beta = .15$, $p < .050$), habit motivation ($\beta = .25$, $p < .001$), and escapism ($\beta = .32$, $p < .001$), while having a more agentic social media mindset predicted less problematic TikTok use ($\beta = -.18$, $p < .05$). Loneliness was reduced for those with a more agentic mindset ($\beta = -.24$, $p < .01$) and who used TikTok out of habit ($\beta = -.27$, $p < .01$). Using TikTok for self-expression ($\beta = .21$, $p < .05$) and escapism ($\beta = .39$, $p < .01$) predicted higher loneliness. Self-esteem was positively predicted for those with an enhanced social media mindset ($\beta = .19$, $p < .01$) and who used TikTok out of habit ($\beta = .27$, $p < .01$). Using TikTok for self-expression ($\beta = -.19$, $p < .05$) and escapism ($\beta = -.27$, $p < .05$) predicted lower self-esteem.

Future Work

How do young adults use TikTok and what effects does that use have on them? Extant research on social media's effects on well-being suggest a nuanced, complex picture. Drawing from three conceptual frameworks—the active/passive model of SNS, uses and gratifications theory, and social media mindsets—this study explored how young adults use TikTok and the extent to which different uses, motivations, and mindsets influence well-being. Descriptively, young adults' TikTok motivations align with what might traditionally be conceived as more passive uses: distraction, habit, escapism. These uses did not necessarily correspond with lower well-being outcomes, however. Those who used TikTok for self-expression and escapism had lower well-being outcomes. Using TikTok out of habit predicted more problematic TikTok use but also corresponded with higher self-esteem. Findings also indicate that mindsets matter somewhat: feeling more control over social media use corresponded with less problematic TikTok use and less loneliness; and feeling like social media use enhanced one's life predicted higher self-esteem. Interestingly, more agentic mindsets were related to less passive-type use, while enhancement mindsets were related to more active-type use. Future analyses might incorporate more advanced modeling techniques to further tease out the pathways between use, motivations, and mindsets and their effects on well-being.

Additionally, it is important to consider an individual's entire social media ecosystem to better understand the effects on well-being. The distraction, habit, and escapist needs fulfilled by TikTok's platform may be complemented by another platform's self-expression and social connection affordances. Future work could investigate how users manage their social media use across platforms and explore practices for navigating platforms' affordances in ways that enhance users' agency and experience.

Attending to young adults' attitudes, mindsets, and beliefs about TikTok and social media more broadly will offer clearer insight into people's expectations and desires in relation their social media use. Context and use-centered design will improve the ways in which researchers understand and address social media use, conceptualization, and associated health conditions like mental illness. Future research might also consider how young adults engage with well-being and mental health on TikTok itself. This would allow scholars to examine how and in what ways attitudes about TikTok use may be changing young people's ideas and beliefs about mental illness and what effect those changes may be having on their identity and health-seeking practices.

References

- Berry, N., Lobban, F., Belousov, M., Emsley, R., Nenadic, G., & Bucci, S. (2017). #WhyWeTweetMH: Understanding Why People Use Twitter to Discuss Mental Health Problems. *J Med Internet Res*, 19(4), e107. <https://doi.org/10.2196/jmir.6173>.

- Bucher, T., & Helmond, A. (2018). The affordances of social media platforms. *The SAGE Handbook of Social Media*, 1(1), 233-253.
- Ellison, N. B., Pyle, C., & Vitak, J. (2022). Scholarship on well-being and social media: A sociotechnical perspective. *Current Opinion in Psychology*, 46, 101340.
- Faverio, M., Anderson, M., Park, E. (Apr 2025). Teens, Social Media and Mental Health. Pew Research Center. <https://www.pewresearch.org/internet/2025/04/22/teens-social-media-and-mental-health/>
- Gerbaudo, P. (2024). TikTok and the algorithmic transformation of social media publics: From social networks to social interest clusters. *New media & society*, 14614448241304106.
- Gerson, J., Plagnol, A. C., & Corr, P. J. (2017). Passive and active Facebook use measure (PAUM): Validation and relationship to the reinforcement sensitivity theory. *Personality and Individual Differences*, 117, 81-90.
- Gottfried, J. and Park, E. (Nov 2025). Americans' Social Media Use 2025. Pew Research Center. <https://www.pewresearch.org/internet/2025/11/20/americans-social-media-use-2025/>
- Günlü, A., Oral, T., Yoo, S., & Chung, S. (2023). Reliability and validity of the problematic TikTok use scale among the general population. *Frontiers in Psychiatry*, 14, 1068431.
- Hendrix, J. (Dec 2022). Is TikTok "Digital Fentanyl"? Tech Policy Press. <https://www.techpolicy.press/is-tiktok-digital-fentanyl/>
- Katz, E., Blumler, J. G., & Gurevitch, M. (1973). Uses and gratifications research. *The Public Opinion Quarterly*, 37(4), 509-523.
- Keles, B., McCrae, N., & Grealish, A. (2019). A systematic review: the influence of social media on depression, anxiety and psychological distress in adolescents. *International Journal of Adolescence and Youth*, 25(1), 79-93. <https://doi.org/10.1080/02673843.2019.1590851>
- Kross, E., Verduyn, P., Sheppes, G., Costello, C. K., Jonides, J., & Ybarra, O. (2021). Social media and well-being: Pitfalls, progress, and next steps. *Trends in Cognitive Sciences*, 25(1), 55-66.
- Lee, A. Y., Ellison, N. B., & Hancock, J. T. (2023). To use or be used? The role of agency in social media use and well-being. *Frontiers in Computer Science*, 5, 1123323.
- Lee, A. Y., & Hancock, J. T. (2024). Social media mindsets: A new approach to understanding social media use and psychological well-being. *Journal of Computer-Mediated Communication*, 29(1), zmad048.

- Lee, E., Lee, J. A., Moon, J. H., & Sung, Y. (2015). Pictures speak louder than words: Motivations for using Instagram. *Cyberpsychology, Behavior, and Social Networking*, 18(9), 552-556.
- Meier, A., & Reinecke, L. (2021). Computer-mediated communication, social media, and mental health: A conceptual and empirical meta-review. *Communication Research*, 48(8), 1182-1209.
- Meier, A., Ellison, N., Reinecke, L., & Valkenburg, P. M. (2024, March). Beyond Active-Passive: Towards the Next Stage of Social Media and Mental Health Research. OSF pre-print.
- Omar, B., & Dequan, W. (2020). Watch, share or create: The influence of personality traits and user motivation on TikTok mobile video usage.
- Rosenberg, M. (1965). Rosenberg self-esteem scale (RSE). Acceptance and commitment therapy. *Measures Package*, 61(52), 18.
- Ruggiero, T. E. (2000). Uses and gratifications theory in the 21st century. *Mass Communication & Society*, 3(1), 3-37.
- Russell, D., Peplau, L. A., & Ferguson, M. L. (1978). Developing a measure of loneliness. *Journal of personality assessment*, 42(3), 290-294.
- Sanchez, G., & Jenkins, J. H. (2024). Social media & subjectivity: Adolescent lived experiences with social media in a Southern California middle school. *Soc Sci Med*, 348, 116839. <https://doi.org/10.1016/j.socscimed.2024.116839>
- Valkenburg, P. M., Meier, A., & Beyens, I. (2022). Social media use and its impact on adolescent mental health: An umbrella review of the evidence. *Current Opinion in Psychology*, 44, 58-68.
- Verduyn, P., Lee, D. S., Park, J., Shablack, H., Orvell, A., Bayer, J., ... & Kross, E. (2015). Passive Facebook usage undermines affective well-being: Experimental and longitudinal evidence. *Journal of Experimental Psychology: General*, 144(2), 480.
- Verduyn, P., Ybarra, O., Résibois, M., Jonides, J., & Kross, E. (2017). Do social network sites enhance or undermine subjective well-being? A critical review. *Social Issues and Policy Review*, 11(1), 274-302.
- Verduyn, P., Gugushvili, N., & Kross, E. (2022). Do social networking sites influence well-being? The extended active-passive model. *Current Directions in Psychological Science*, 31(1), 62-68.
- WHO, World Health Organization. (2025). Mental health of adolescents [Fact Sheet]. <https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health> Accessed 26 January 2026.

Session 5C - Discourse and Public Opinion (4)

Day 1 · Mon 13 July · 16:30–17:30, Studio 2, ARC

5C.1 Hot Nerd, Femininity and Gender Relations in Post-socialist China: A Critical Discourse Analysis of Sapiosexuality on RedNote

Qianzao Yang (University of Cambridge)

Introduction

“Brainy is the new sexy”. On September 30th, 2023, a note on RedNote advocating sapiosexuality as her romantic preference immediately went viral. The hashtag #Sapiosexuality has since accumulated over 2.33 million discussions. RedNote is a Chinese female-dominated platform whose users are predominantly urban young women. It emphasises emotional expression and embodied experience (Gu et al., 2024; Zhai, 2025). The popularisation of #Sapiosexuality has reflected a new trend among young Chinese women, where intelligence is romantically imagined as a symbol of desirability. Originating from Western culture, sapiosexuality refers to sexual attraction toward intelligence (Peckham, 2012). On social media, it incorporates reflections on taste, romantic and class aspirations. Online discussions over sapiosexuality exhibit a striking polarisation. Some view it as women’s pursuit of emotional resonance and self-awareness, while others criticise it as reproducing elitism and educational hierarchy.

This phenomenon merits scholarly attention for several reasons. Firstly, it reflects women’s imagination of mate selection criteria, while crystallising their multiple subject positions and self-presentation through algorithmically mediated platforms. Second, it brings private romantic topics into platform visibility, providing a lens through which to understand femininities and gender relations within the complexity and dynamics of post-socialist China. Third, it demonstrates the role of social media in the negotiation of state hegemony and gender order in women’s everyday digital practices, allowing this study to contribute to media studies and gender studies in general.

This study focuses on the sapiosexuality discussions on RedNote to explore how young Chinese women construct their ideal partners, reassert subjectivities and challenge gender relationships. Drawing on critical discourse analysis, it also critically examines

the role of social media in displaying women's self-presentation and femininities. It further reveals the underlying interplay between post-socialist femininities, platform affordances, feminist agency and ingrained gender ideologies. The analysis is highly engaged with individualisation, neoliberalism and postfeminism within the context of post-socialist China.

Objectives/Research Questions

Rooted in traditional Confucian and patriarchal social norms, China's heterosexual marriage system emphasises male dominance and female subordination (Connell, 2020; Zhong & Wilkinson, 2025). Despite state propaganda of "women's work" in the Maoist era, gender inequalities persisted within intimate relationships and marital roles (Chen, 2003). Economic reforms from the 1980s onward fundamentally reconfigured gender relations. The agendas of post-socialist states, market forces and resurgent patriarchal values intertwined, forging the norm of "marriage up" in the marriage market, institutionalising women's sexuality as a means to secure economic stability (Wu & Dong, 2019). Meanwhile, marketisation, privatisation and the one-child policy gave rise to China's middle class (Rofel, 2007). Primarily born in the 1980s and later, this generation of women confronts typical ambivalence. They are encouraged to pursue their own careers (Liu, 2008), while remaining constrained by persistent gender norms. Thus, young Chinese women live as ambivalent desire subjects between self-independence and the heterosexual patriarchy. They exhibit post-feminist sensibility through their recognition of gendered norms within China's society and transnational post-feminism (Ge 2025).

Previous research has investigated RedNote as a gender-based platform for public feminist expression (Liang & Wang, 2025; Zhao, 2025). Wang and Chang (2025) argued that it has been repurposed as a popularised enclave for lifestyle feminism and intimate publics. However, the discourses surrounding intimate relationships and gendered power have not been sufficiently studied. Moreover, existing scholarship on Chinese femininity has focused on the female body and influencers. Chang et al. (2018) explored profile picture editing as self-empowerment, arguing it reflects a moderate and pragmatic feminine discourse. Peng (2019) analysed discursive strategies of Chinese KOLs, revealing how neoliberal feminism intersects with the construction of a feminised male ideal. This leaves a research gap on how young Chinese women negotiate femininity through romantic preferences in digital spaces.

Accordingly, the following research questions are proposed:

RQ.1 How do young Chinese women construct their mate selection criteria, as well as new forms of self-expression and femininity within online sapiosexuality discourse on RedNote?

RQ.2 How does discursive practice present new femininities in digital spaces?

RQ.3 How does the online sapiosexuality discourse reflect, negotiate, or challenge sociocultural and gender norms of China's post-socialist society?

Methods

This study utilises a critical discourse analysis (CDA) to explore the online debate related to sapiosexuality. CDA is a qualitative analytical approach for critically considering discourse as a social action and interpreting its power relations, ideologies, and social inequalities (Fairclough, 2001). This method is widely used to study practices within online discourse and texts. This study follows Fairclough's three-dimensional approach and focuses on the text analysis, discourse practice, and social practice to analyse how discourse is constructed and reflects socio-cultural implications and power relations, while paying more attention to the gender ideologies and gender power relations implicit within the texts. I collected and cleaned the dataset using Python, conducted textual analysis with Rost CM 6.0.

RedNote is a content-sharing platform with over 300 million monthly active users, where 72.56% of users are female and 46% are under the age of 30 (Jiao et al., 2025). It is widely recognised as a feminine space, making it an ideal data source for examining how young women engage in online discussions about love, class and gender. During data collection process, I registered new accounts to mitigate personal bias and the influence of the platform's recommendation algorithm, searched for the “智性恋 (sapiosexuality)” keyword and sorted by popularity, collecting 100 notes. After manually cleaning irrelevant posts and comments, the final dataset comprises 50 notes and their associated 16447 comments published between March 2022 and October 2025.

All data analysed were publicly available on RedNote. All identifying information was anonymised during the analysis and in all research outputs.

Results

Sapiosexuality discourse reshaped young women's mate selection criteria. One notable feature I identify is women's pursuit of constructing a perfect “hot nerd” as their ideal partner. The idealised figure is characterised not only by superior intellectual capital but also by perfection in appearance, physique and social class. In this way, women developed a discourse strategy to articulate their romanticised imaginaries and class-conscious taste. Influenced by consumerism on the platform, “hot nerd” was transformed into a fashionable and consumable label with homogenised partner traits. It emphasises the complementary roles of women and men within Confucian gender values (Peng, 2019).

Expressions like “Attracted by smart people, like you're evenly matched”, further represented how women openly discussed intimacy on social media, reflecting a proactive approach to pursuing love. Specifically, comments like “It doesn't feel good

to just follow someone” also prioritised women’s personal feelings, aspirations and experiences. These arguments reclaimed greater female agency, while challenging the patriarchal gender roles which position women as passive, subservient and self-sacrificing in romantic relationships (Zurndorfer, 2018; Gui, 2020). Moreover, “Sapiosexuality is my gold standard”, naturalised sapiosexuality as a matter of esthetic and personal choice. These discussions aligned with the rise of the individualistic culture and the neoliberal turn within modern China (Yang & Jin, 2025). Arguably, a self-disciplined, rational female persona with risk management awareness is gradually emerging in digital spheres.

The doubts of sapiosexuality demonstrated an increasingly prominent feminist subjectivity. Statements such as “Knowledge is not sexually transmitted” and “You’re just a Chinese with diploma-oriented bias,” employed mockery to critique the dominant ideologies of elitism and education discrimination, situating sapiosexuality within the highly competitive education system and class hierarchy of contemporary China. These humorous criticisms not only reflected the resistance towards mainstream masculinity performance, but also deconstructed gender norms from a female perspective (Luo & Sun, 2014). Yet, such calls largely remain at the level of emotional expression, encoded in the state-sponsored postfeminism, which celebrates women’s active choices while rejecting structural inequalities and gender oppression in realities to avoid self-conflict (Wu, 2025).

In conclusion, RedNote becomes a discursive space that is closely tied to feminist engagement, platform affordances and China’s socio-economic governance (Zhao et al., 2025). In this context, young Chinese women possess more autonomy in love, while inevitably retaining traditional femininity from heteropatriarchy and Confucianism. These hybridised femininities reflect tensions between neoliberal self-optimisation, postfeminism and Confucianism within China’s political-economic processes.

Future Work

I will further examine the platform logics as well as the algorithm in shaping this gendered discourse. Future research will expand the scope of the corpus, incorporating computational methods such as LDA topic modelling to identify larger scale discourses.

References

- Chang, J., Ren, H., & Yang, Q. (2018). A virtual gender asylum? The social media profile picture, young Chinese women’s self-empowerment, and the emergence of a Chinese digital feminism. *International Journal of Cultural Studies*, 21(3), 325–340. <https://doi.org/10.1177/1367877916682085>

- Chen, T. M. (2003). Female Icons, Feminist Iconography? Socialist Rhetoric and Women's Agency in 1950s China. *Gender History*, 15(2), 268–295.
<https://doi.org/10.1111/1468-0424.00303>
- Connell, R. (2020). *Masculinities*. Routledge.
- Fairclough, N. (2001). Critical Discourse Analysis as a method in social Scientific research. In SAGE Publications Ltd eBooks, 121–138.
<https://doi.org/10.4135/9780857028020.n6>
- Ge, L. (2025). Ambivalent desiring subjects: young women, agency and post-(socialist-)feminist sensibilities in China. *Feminist Theory*.
<https://doi.org/10.1177/14647001251334949>
- Gu, Y., Yang, Y., Saiyinjiya, A., Wu, W., Chen, Q., Chen, S., & Ioana Literat. (2024). Commerce Meets Activism: #StopMenstualShaming and the Dynamics of Feminist Advocacy on Xiaohongshu. *International Journal of Communication*, 18(0), 22. <https://ijoc.org/index.php/ijoc/article/view/22324>
- Gui, T. (2020). “Leftover Women” or Single by Choice: Gender Role Negotiation of Single Professional Women in Contemporary China. *Journal of Family Issues*, 41(11), 0192513X2094391. <https://doi.org/10.1177/0192513x20943919>
- Jiao, L., Jin, Y., & Shen, Y. (2025). Polarized views over feminism: gendered reception of Barbie among Chinese audiences on social media. *Feminist Media Studies*, 1–18.
<https://doi.org/10.1080/14680777.2025.2478390>
- Liang, L., & Wang, H. (2025). Breaking the silence: women's resistance to menstrual shame on the Xiaohongshu social media platform. *Culture Health & Sexuality*, 1–18. <https://doi.org/10.1080/13691058.2025.2511059>
- Liu, F. (2008). Constructing the autonomous middle-class self in today's China: the case of young-adult only-children university students. *Journal of Youth Studies*, 11(2), 193–212. <https://doi.org/10.1080/13676260701800746>
- Luo, W., & Sun, Z. (2014). Are You the One? China's TV Dating Shows and the Sheng Nü's Predicament. *Feminist Media Studies*, 15(2), 239–256.
<https://doi.org/10.1080/14680777.2014.913648>
- Peckham, A. (2012). *Urban Dictionary: Freshest Street Slang Defined*. Andrews McMeel Publishing.
- Peng, A. Y. (2019). Neoliberal feminism, Gender relations, and a Feminized Male Ideal in China: a Critical Discourse Analysis of Mimeng's WeChat Posts. *Feminist Media Studies*, 21(1), 115–131. <https://doi.org/10.1080/14680777.2019.1653350>

- Rofel, L. (2007). *Desiring China: Experiments in neoliberalism, sexuality, and Public Culture*. Duke University Press.
- Wang, H., & Chang, J. (2025). Lifestyle Feminism on RedNote: Digital Platforms, Mediated Intimacy, and the Duality of an Enclaved Feminist Public Sphere. *Social Media + Society*, 11(4). <https://doi.org/10.1177/20563051251409617>
- Wu, A. X., & Dong, Y. (2019). What Is Made-in-China Feminism(s)? Gender Discontent and Class Friction in Post-socialist China. *Critical Asian Studies*, 51(4), 471–492. <https://doi.org/10.1080/14672715.2019.1656538>
- Wu, Q. (2025). Exploring Postfeminist Humor as a Collective Form of Resistance to Anti-feminism in China: A Textual Analysis of Radical Feminist Discourse on Weibo. *DiGeSt - Journal of Diversity and Gender Studies*, 12(1). <https://doi.org/10.21825/digest.91541>
- Yang, X., & Jin, Y. (2025). “Zheng Qinwen is the pure-blood ideal of Gu Eileen”: Exploring the sensemaking of elite female athlete identities on Chinese social media. *International Review for the Sociology of Sport*. <https://doi.org/10.1177/10126902251356556>
- Zhai, H. (2025). Hashtags and Hope: Tracing Chinese Digital Spirituality Through #WishMeABabyBoy Trend on Xiaohongshu (Rednote). *Social Media + Society*, 11(3). <https://doi.org/10.1177/20563051251369816>
- Zhao, X. (2025). “Say no to the commodification of our uterus”: A multimodal critical discourse analysis of anti-surrogacy posts on Chinese social media. *Discourse & Society*. <https://doi.org/10.1177/09579265251369117>
- Zhao, Y., Ali Mahfoodh, O. H., Subramaniam, I., & Ge, J. (2025). Pixels, Poses, and Power: A multimodal analysis of Self-Representation and femininity among female fitness influencers on Chinese social media. *Discourse, Context & Media*, 68, 100954. <https://doi.org/10.1016/j.dcm.2025.100954>
- Zhong, C., & Wilkinson, J. (2025). The paradox of women’s marital freedom: nonlinear individualization in post-reform China. *Theory and Society*, 54(1), 119–138. <https://doi.org/10.1007/s11186-025-09591-w>
- Zurndorfer, H. (2018). Escape from the country: the gender politics of Chinese women in pursuit of transnational romance. *Gender, Place & Culture*, 25(4), 489–506. <https://doi.org/10.1080/0966369x.2018.1453488>

5C.2 Not-So-Happy Ending: How failed series finales come to be and the reaction from online fan bases

Tess Arnold (Toronto Metropolitan University)

Introduction

The feeling of being let down by a series finale is one that many television viewers have experienced at least once in their television-viewing lifetime. Dedicating hours, days and sometimes years to a show just for it to have an unsatisfactory conclusion can be perceived as a betrayal. It can feel as if the writers' room is stabbing the audience in the back after hours of loyal viewership; however, a series finale proves to create complex emotions for both viewers and the creative teams behind the shows that fans feel so strongly towards. This paper will examine the phenomenon of the 'failed' series finale from both the side of the creator and the consumer. The paper will explore the screenwriters' point of view during the writing process in contrast to the discourse from online fanbases following the final episode's release. Both the writers and viewers have their own unique perspective on a series coming to an end; however, similarly, fans and writers often share a closeness to the stories and characters, feeling personally attached to the ways in which their stories end.

Objectives/Research Questions

This research unveils the relationship between fandoms, creators, narratives, and the current state of the television industry as a whole. The goal of this paper is to reveal a more nuanced understanding of the relationship between creatives and the fandoms that form from their creations, where they intersect and where they diverge. As online fandoms continue to grow on platforms that allow for thoughts, opinions, and reactions to be shared, creatives must adjust to the new environment in which fan expectations and desires are constantly out in the open to see and interact with.

Methods

From the point of view of the creator, this paper will explore the screenwriters' process in formulating series finales, the influences that social media now has on this process, and the ways in which they attempt to manage creative vision and fan expectation. Using literature on television studies, screenwriting practices, and audience theories, this paper will dive in depth into the behind-the-scenes of the screenwriting process, demonstrating one side of a complex relationship between creator and consumer. As a primary source, the recent series finale of *Stranger Things* (2016-2025) and the companion documentary *One Last Adventure: The Making of Stranger Things 5* (2025) will be used as an exploration and overall case study that will be used throughout this

paper. The documentary provides an in-depth look into the writing process of the show's creators, Matt and Ross Duffer, as they write and direct the final season. This behind-the-scenes look at this large-scale production provides concrete and topical examples from within the writers' room.

The documentary touches on (often unintentionally) the writer's relationship to fan expectations, creative vision, and the reality of the modern industrial television landscape. The following section will focus on an exploration of fandom expectations and how they shape reactions to series finales, particularly when these expectations are not met. Through a mixed-methods analysis, this paper will explore how fans express their discontent, disappointment, and disapproval on online platforms following the conclusions of long-running series whose endings are described as not living up to expectations. Drawing on John Fiske's work within the realm of television studies and fandom theories, this paper will argue that audiences form deep connections with these shows, often leading them to have strong emotional reactions when these connections are felt to be violated. The *Stranger Things* case study will continue in this section, with the finale being used to formulate an understanding of the fans' perspective. Using online discourse following the aftermath of the final episode, this paper will explore the *Stranger Things* fandom's relationship to the series and their reactions to the final episode of the series, specifically focusing on scenarios in which fans claim their expectations were not met.

This section will also allow for further exploration into related fan theory concepts such as canon vs fanon, a phenomenon of fans creating their own canon for the storyline of something they love. Due to fans' strong attachment to these stories and characters, they may develop their own understanding of how the story should unfold or the ways in which certain characters should act and react in direct opposition to the creative decisions of the writers.

Results

As this paper is currently a work in progress, the results and findings are still being determined; however, based on previous research on similar topics, it can be hypothesized that in the modern landscape of television screenwriting, fan expectations play a large role in the writing process. On the other hand, if and when these expectations are not considered, there can be a strong adverse reaction from fans online.

Future Work

As this paper progresses, further theoretical analysis and exploration of online discourse will occur. The goal of this paper is to provide an interesting and entertaining look behind the curtain of the television creative process in contrast to what fans see

and contribute online, all backed by a strong theoretical framework based on academic research into audiences, screenwriting, and fandoms.

References

Fiske, John. *Television Culture*. 2nd ed. Routledge Classics. Routledge, 2011.

Gray, Jonathan, and Amanda D. Lotz. *Television Studies*. Second edition. Short Introductions. Polity, 2019.

Katharine E. McCain. "Canon vs. 'Fanon': Genre Devices in Contemporary Fanfiction." Thesis, Georgetown University, 2015.
<https://ezproxy.lib.torontomu.ca/login?url=https://www.proquest.com/dissertations-theses/canon-vs-fanon-genre-devices-contemporary/docview/1679935900/se-2?accountid=13631>.

Keidra Chaney and Raizel Liebler. "Canon vs. Fanon: Folksonomies of Fan Culture." *Media in Transition 5: Creativity, Ownership and Collaboration in the Digital Age*, n.d. https://web.mit.edu/comm-forum/legacy/mit5/papers/Chaney_Liebler_MIT5.pdf.

Krakowiak, K. Maja. "How It Ends: Exploring How Narrative Closure and Character Outcomes in Series Finales Relate to Viewer Responses and Coping Behaviors." *Mass Communication and Society* 27, no. 5 (2024): 1028–55.
<https://doi.org/10.1080/15205436.2024.2344515>.

Kreuter, Aaron. "The Last River." In *Shifting Baseline Syndrome*. Oskana Poetry & Poetics 13. University of Regina Press, 2022.
<https://ebookcentral.proquest.com/lib/torontomu/detail.action?docID=7083178>.

Lepori, Gabriele M. "Investor Mood and Demand for Stocks: Evidence from Popular TV Series Finales." *Journal of Economic Psychology* 48 (2015): 33–47.
<https://doi.org/10.1016/j.joep.2015.02.003>.

Redvall, Eva N., and John R. Cook. "Television Screenwriting – Continuity and Change." *Journal of Screenwriting* 6, no. 2 (2015): 131–36.
https://doi.org/10.1386/josc.6.2.131_2.

Walter, Richard. *Screenwriting: The Art, Craft and Business of Film and Television Writing*. A Plume Book. Plume, 1988.

5C.3 Sexual violence (in)justice online: Exploring the influence of negative discourses on victim-survivor justice seeking

Ursula Shepherd (University of Glasgow)

Introduction

This research is concerned with how victim-survivors of sexual violence construct their identities, in relation with justice, despite the propagation of rape culture and myths on social media sites. Within the sociopolitical framework of the United Kingdom – whereby the neoliberal and austerity measures have weakened the welfare state and gutted public spending, including funding for domestic violence refuges, women’s organisations and the criminal justice system (Ishkanian, 2014; Sanders-McDonagh, Neville & Nolas, 2016) – victim-survivors frequently receive little to no support from authorities and face secondary victimisation caused by the police and justice system (Brooks & Burman, 2016; Victims Commissioner, 2020). As a result of potentially harmful formal justice options, online spaces have become a way for victim-survivors to discuss their experiences and build supportive communities (Gorissen et al., 2023). Assisted by the nature of online spaces – as they allow for increased interactivity, communication and anonymity – victim-survivors are able to contribute and construct a ‘safe space’, where those with marginalised and stigmatised offline identities can develop and manage their experiences with others in similar positions (Smith, Wickes and Underwood, 2013; Matich, Ashman & Parsons, 2022); O’Neill, 2024).

While literature on victim-survivor use of online spaces in relation to justice is vast, it appears to inadvertently focus on anonymous disclosures or disclosures within closed groups. This research extends the existing corpus to incorporate those who have received less academic focus, specifically victim-survivors who publicly discuss their experiences online, and in doing so, they defy online perpetrated negative discourses. Online spaces can be the site of misogynistic, racist and ableist violence, especially within the ‘post-#MeToo’ culture, where victim-survivors are finding themselves having to prove that their experiences are believable, with social media taking an active role in scrutinising sexual violence accusations, especially those of high-profile celebrities (Banet-Weiser & Higgins, 2023; Boyle, 2024). Thus, engaging with those who openly give voice to their experiences through these online avenues is a necessary advancement, helping to further understand the ways in which victim-survivors envision and seek justice, despite the chance of receiving harmful and threatening responses.

By providing a space for participants to creatively reflect on their experiences, this research may also be beneficial to the participants themselves, cultivating a space

centring around ‘trust, joy, passion and love’ (Waite et al., 2024: 4). Further, the methods chosen means that creative and engaging outputs can be produced, opening this research to those beyond academic circles.

Objectives/Research Questions

This PhD project seeks to answer questions concerning how victim-survivors of sexual violence construct and represent their identities online and whether their navigation of negative online discourses influence this construction. Furthermore, discussions on how justice is envisioned and understood by victim-survivors will be central, providing insights into how justice can be malleable, different to each victim-survivor and not focused on a particular end point such as a prison sentence.

The research questions are:

- How do victim-survivors of sexual violence construct and represent their identities in online public spaces?
- In what ways does victim-survivor online representation of themselves and their experiences engage with and resist negative online narratives?
- How is (in)justice framed and understood by victim-survivors and what forms of justice are being imagined or pursued online?

Each question incorporates a core theme of the project: victim-survivor identities, resistance to negative online discourses and (in)justice.

Methods

To answer the above questions, a multifaceted theoretical approach – incorporating understandings of feminist digital geographies, social media affordances, identity construction and justice – is adopted. This helps to trace how online spaces can be a site of significant identity curation and navigation wherein justice can be felt, while addressing the harms of having a public victim-survivor identity online.

The project utilises a creative qualitative approach, mixing zine making workshops and interviews with victim-survivors. The methods have been chosen for their ability to answer the research questions in an engaging and creative manner, while allowing for the research to use an empowerment-focused framework, that complements and upholds the work of Downes, Kelly and West (2014) that situates the victim-survivor voice as central.

Furthermore, there is a political aspect to using zines. Thanks to their history within activism, they provide a platform for the personal voice while additionally creating a collective and feminist voice that speaks to justice. The zines allow participants to explore their experiences with justice and online discourses and can help to articulate their thoughts beyond standard interviews. Participants will have the autonomy over their selection of images, the final outputs and will be offered interviews to further

reflect on their experiences. These interviews help to understand the meaning and ideas behind participants zines and will consequently aid the visual work through textual data.

Results

While there are still workshops and interviews to take place, themes of disillusionment with traditional and formal state justice, victim-survivor voice and the injustices faced online emerge. State justice is viewed as old-fashioned for sexual violence, where justice is hard to access or feel. The benefits of social media and using the platforms to voice their experiences are discussed. This sheds light on the current use of social media by victim-survivors and provides a greater understanding of how justice is sought. The paper will outline these findings, illuminated through imagery or quotations from the zines and interviews.

Future Work

Future work will include enhanced analysis of the zines and interviews, providing in-depth understanding of the themes and perspectives of the participants. On completion of the project, individual participant zines will be incorporated into one singular project zine, whereby participants will choose aspects of their zines to be included, giving them further autonomy over the final output. From this, the work will be published within academic journals and if possible, the final project zine will be displayed or included as part of an exhibition. This allows the research reach and impact to move beyond academic circles in an engaging manner.

References

- Banet-Weiser, S., & Higgins, K. C. (2023). *Believability: Sexual violence, media, and the politics of doubt*. Polity.
- Boyle, K. (2024). *#MeToo and Feminism: Weinstein and Beyond*. Springer Nature Switzerland. <https://doi.org/10.1007/978-3-031-67314-6>
- Brooks, O., & Burman, M. (2017). Reporting rape: Victim perspectives on advocacy support in the criminal justice process. *Criminology & Criminal Justice*, 17(2), 209–225. <https://doi.org/10.1177/1748895816667996>
- Downes, J., Kelly, L., & Westmarland, N. (2014). Ethics in Violence and Abuse Research—A Positive Empowerment Approach. *Sociological Research Online*, 19(1), 29–41. <https://doi.org/10.5153/sro.3140>
- Gorissen, M., Van Den Berg, C. J. W., Bijleveld, C. C. J. H., Ruiter, S., & Berenblum, T. (2023). Online Disclosure of Sexual Victimization: A Systematic Review. *Trauma, Violence, & Abuse*, 24(2), 828–843. <https://doi.org/10.1177/15248380211043831>

- Ishkanian, A. (2014). Neoliberalism and violence: The Big Society and the changing politics of domestic violence in England. *Critical Social Policy*, 34(3), 333–353. <https://doi.org/10.1177/0261018313515973>
- Match, M., Ashman, R., & Parsons, E. (2022). Seeking safety and solidarity through self-documentation. In P. Maclaran, L. Stevens, & O. Kravets, *The Routledge Companion to Marketing and Feminism* (1st ed., pp. 388–400). Routledge. <https://doi.org/10.4324/9781003042587-30>
- O'Neill, T. (2024). *Disclosing Sexual Violence in a Digital Society: Storytelling, Activism, and Justice* (Vol. 3). Springer Nature Switzerland. <https://doi.org/10.1007/978-3-031-74990-2>
- Sanders-McDonagh, E., Neville, L., & Nolas, S.-M. (2016). From Pillar to Post: Understanding the Victimisation of Women and Children who Experience Domestic Violence in an Age of Austerity. *Feminist Review*, 112(1), 60–76. <https://doi.org/10.1057/fr.2015.51>
- Smith, N., Wickes, R., & Underwood, M. (2015). Managing a marginalised identity in pro-anorexia and fat acceptance cybercommunities. *Journal of Sociology*, 51(4), 950–967. <https://doi.org/10.1177/1440783313486220>
- Victims Commissioner (2020). Survey finds rape victims have lost faith in the justice system. Available at <https://victimscommissioner.org.uk/news/survey-finds-rape-victims-have-lost-faith-in-the-justice-system/>.
- Waite, S., Darley, D., Eden-Barnard, J., Rutter, N., & Tatton, S. (2024). Women, Shame, and Stigma: Responding to (In)justice Through Zine. *Sociological Research Online*, 13607804241292187. <https://doi.org/10.1177/13607804241292187>
-

Session 5D - **PANEL: Towards a Future Research Agenda: TikTok Creators and Digital Economies**

Day 1 · Mon 13 July · 16:30–17:30 · Seminar Suite B, ARC

Zoetanya Sujon (University of the Arts London), Taylor Annabell (Cardiff University), Marie Heřmanová (University College London), Irida Ntalla (London College of Communication)

The objective of this panel is to share and develop a dialogue based on selected contributions of the Special Issue on TikTok Creators and Digital Economies published in *Continuum, Journal of Media & Cultural Studies* (2026). In line with the Social Media and Society 2026 conference key themes, these select papers provide key insights into the state of content creators on TikTok and what this means for changing affective, cultural, and digital economies. The special issue provides key insights important to the broader social media and society community, including concepts and methodological approaches and their future directions.

Collectively, the special issue focuses on two key thematic, examining ‘creators, influencers, and identities’ and ‘TikTok’s digital economies.’ The papers in the panel present a complex negotiation of identities, creative labour and visibilities in TikTok’s digital economies. For example, Heřmanová et al.’s work examines how Ukrainian creators navigate ‘ambivalent visibility’ while Ntalla argues that single mothers employ a ‘resilient-imperfect’ framework in their content strategies on TikTok. Ambivalence and resilient-imperfect highlight tensions between creators and TikTok’s informal / formal economies, where influencers are positioned as precarious and individual traders, with tenuous avenues for content ownership and visibility (see Annabell). Focusing on identities shows that such tensions exacerbate existing racialized disparities in digital economies, as seen in the experiences of Black creators (Taylor). Ultimately, the session aims to translate this research into a conversation to engage the conference audience in considering and identifying emerging research gaps and research directions emerging from the special issue’s individual papers and collective contributions. To do this, the editors have identified key questions which the panel will take up based on their own TikTok research providing both specific and more general reflections.

The panel engages with and extends the central questions raised by the special issue on TikTok Creators and Digital Economies, contributing and critically reflecting on the current state of TikTok and its scholarship. While TikTok shares common features with other platforms, it has also emerged as a unique site for contemporary culture, popular

expression, and with distinct economic structures and opportunities. TikTok is at once a platform, app, company, market, community, and cultural producer of cultural producers (see Mahetaji and Nieborg, 2025). As the introduction of the special issue argues, TikTok's popularity and platform architecture heightens tensions between commodification and creativity, visibility and voice, performativity and authenticity (Sujon and Ntalla 2025). This panel examines these tensions opening up important questions based on the special issue's featured publications as well as newer research, allowing space for broader reflection and future research.

While the key questions are contextualized throughout, they evolve from big questions around the nature of cultural production, creator practices, contingent economics alongside systems of value and exchange. The panel concludes with a discussion of methodologies, interdisciplinarity, and important research gaps – all of which make up a wider TikTok research agenda.

TikTok creators are presented as credible counter-narrators, agents, and community makers, driven by shared identity markers, and lived experiences and ask a set of collective questions.

- 1) How can we shift our analytical focus beyond the creative, relational and visibility labour of individuals who successfully monetise content, and incorporate the significant diversity in creators' motivations, practices, and professional ambition?
- 2) We consider how creators navigate the tension between being visible as conditioned by algorithms, platform vernaculars, and monetisation pathways, and the desire to be heard and recognised. How do the frameworks and conceptual understandings suggested in the contributions help us explore this negotiation?
- 3) What do the experiences of diverse creators collectively reveal about how TikTok and affiliates operate along lines of race, gender, legal status, and geopolitical location?

Heřmanová et al. (2025) introduce 'ambivalent visibility' to capture how Czech and Ukrainian TikTok creators sharing war-related content from Ukraine, a condition shaped by the conflicting demands creators face in balancing algorithmic amplification, audience engagement and ethical responsibility. Ntalla (2025) examines how single mothers navigate imperfect-resilient motherhood through narratives that use platform affordances to reframe stigmatized identities into aspirational stories. These creators' visibilities become part of a social mechanism for performing resilience, situated within the spaces of the neoliberal ideal of a self-sufficient, self-managing mumpreneur. Annabell et al. (2025) explore the tensions between creators as individuals and traders, identifying the discrimination faced by influencers in their treatment by brands and

ultimately leading to limits in sharing narratives of unfairness and racial discrimination. Taylor (2025) examines how the ambiguity of credit on TikTok shapes who gets acknowledged for their creative labour with a focus on Black creators' struggles for recognition. 'Credit' and the process of attributing the origins and influence of content is not just individual recognition but a process fundamentally shaped by the platform's architecture and practices.

The special issue's introduction also explores the imprecise term of 'digital economies' considering how it has been articulated and what these articulations mean for understanding TikTok's systems of value and exchange. This broad concept is built upon historically situated concepts like cultural and media economies, which account for the production and exchange of meaning making and symbolic values which should be called more to the forefront (Sujon and Ntalla 2025). From this perspective, additional questions arise about TikTok ecosystems:

4) What do creator practices' and the diverse value system that they navigate, as discussed in the panellist contributions, tell us about TikTok digital economies?

We recognise that TikTok, like many other social platforms, intentionally embed pathways that encourage users to become creators, and creators to professionalize their creative practice through monetization, rewards, and audience building. Avenues for content monetization, include fan-based and advertising-driven models including TikTok One, the 'Creator Rewards Program,' donations, and TikTok Live and TikTok Shop. Evidently, the platform actively aims to contribute to couching recognition, voice and popularity in economic terms. The panel draws attention on the platforms vested interest in a 'strategic ambiguity' in relation to the 'cultural economy', including nested concepts and neologisms intended to better capture the wider nuances of our economic systems within changing platform contexts (Sujon and Ntalla 2025).

Heřmanová et al. (2025) present three forms of ambivalence, the tension between raw immediacy and platform friendly visibility, the boundaries between storytelling and journalism, and the virality paradox in which tragedy becomes a driver of engagement. These show vital ethical questions that creators on TikTok need to navigate when developing and sharing content to become visible and remain legible for the algorithm. Ntalla (2025) discusses how single mother creators present themselves on the platform framed through both struggle and success, where resilience functions as an educational capital for the community as part of the mediation of maternal identities and entrepreneurial labour within TikTok's digital economy. Anabell et al. (2025) examine public drama on TikTok as a space to negotiate unequal treatment and accountability outside a formal legal framework. More recent collaborative research extending the perspectives in this paper considers the use of TikTok LIVEs for shopping

in the Netherlands where TikTok Shop is not yet available. The analysis addresses how practices and performances can be understood as acts of both self- and platform governance - as seen in how 'shoppability' is staged as a participatory spectacle (Anabell et al. forthcoming). Taylor (2025) considers how informal crediting practices do not easily fit within existing platform affordances, indicating that creator practices are aligned with both platforms and the broader digital economy.

The panel concludes by opening discussion of current and future methodologies. Questioning empirical and conceptual approaches facilitates better understanding of:

5) How to capture creator practices in evolving TikTok digital economies?

6) How can dialogue between disciplines generate new insights?

7) As TikTok's regulatory and monetisation pathways evolve, what new research gaps are emerging?

Based on these questions, the panellists will collaboratively craft an agenda for future TikTok research, drawing collective research insights from creator practices, values, economics, and strategic lines of enquiry.

References

Annabell, T., Kelder, T., van de Kerkhof, J., Gui, H., & Goanta, C. (2025). Influencer as individual and trader: exploring the boundaries of discrimination in influencer marketing from a multidisciplinary perspective. *Continuum*, 1–16.

<https://doi.org/10.1080/10304312.2025.2518967>

Heřmanová, M., Eriksson Krutrök, M., & Divon, T. (2025). "The algorithm loves the war": ambivalent visibility in content creator practices during war. *Continuum*, 1–17.

<https://doi.org/10.1080/10304312.2025.2507777>

Mahetaji, K., & Nieborg, D. B. (2025). TikTok's tools: The politics of platform tools for cultural production. *Continuum*, 1–28.

<https://doi.org/10.1080/10304312.2025.2525519>

Heřmanová, M., Eriksson Krutrök, M., & Divon, T. (2025). "The algorithm loves the war": ambivalent visibility in content creator practices during war. *Continuum*, 1–17.

<https://doi.org/10.1080/10304312.2025.2507777>

Sujon, Z., & Ntalla, I. (2025). Special issue introduction: reflecting on TikTok creators and digital economies. *Continuum*, 1–19.
<https://doi.org/10.1080/10304312.2025.2549003>

Taylor, Z. (2025). Algorithms, affordances and the ambiguity of credit on TikTok. *Continuum*, 1–14. <https://doi.org/10.1080/10304312.2025.2518971>

Session 5E - **PANEL: From Red Dresses to Memory Stones: Multimedia Activism and Gender-Based Violence in Canada**

Day 1 · Mon 13 July · 16:30–17:30 · Seminar Suite C, ARC

Nicolette Little (University of Alberta), Esteban Morales (University of Groningen), Jaigris Hodson (Royal Roads University), Jessica Ringrose (UCL, Institute of Education)

Our session has two key and intertwined objectives:

On one hand, this panel will help introduce and discuss Nicolette Little's new book, *From Red Dresses to Memory Stones: Multimedia Activism and Gender-Based Violence in Canada* – out June 1 with University of British Columbia Press. UBC Press has promised to make the book available at this conference.

In addition to responding to and discussing this full-length publication's social and political significance, each contributor will share knowledge on the topic of gendered/technology-facilitated violence (Ringrose et al., 2022, 2025), how it is changing as technology evolves, and how citizens continue to meet challenges with new mediated anti-violence solutions. This panel promises to offer rich and engaging discourse concerning how citizens are currently intervening in violence and the discourses that normalize it, and how this will evolve going forward, while also stimulating audience participation.

From Red Dresses to Memory Stones has been awarded a Scholarly Book Award by the Federation for the Humanities and Social Sciences, and an Open Access Grant from the Social Sciences and Humanities Research Council of Canada—the latter of which is to ensure it is accessible to activists and everyday citizens in support of equity, diversity and inclusion.

Gender-based violence (GBV) is the only violent crime on the rise in Canada—the home country of the book's author and panel lead--and rates are also increasing internationally (Canadian Women's Foundation, 2025). It is referred to as a “shadow pandemic” and an “epidemic” in its own right by the United Nations—and GBV impacts members of equity-denied groups with particular frequency and intensity (YWCA, 2024). Meanwhile, misogyny has taken new, potent, and particularly hate-fuelled forms as it spreads across the manosphere, influencing younger and younger men and boys—especially in major cities across the UK, Australia, US, and Canada (Bates, 2021, 2025; Ging, 2017). And technology-facilitated violence—perpetrators' use of media and

technology to extend the reach of their harms—affects women and other equity-denied individuals in the form of (as examples) doxing, cyberstalking, surveillance, the non-consensual sharing of intimate images, and more (Ringrose, 2018; Henry & Powell, 2017).

Great challenges call for significant interventions. This roundtable will convene five leading GBV and mediated anti-violence activism scholars to discuss the conceptual and political significance of Nicolette Little's new book, *From Red Dresses to Memory Stones: Multimedia Activism and Gender-Based Violence in Canada* (UBC Press, June 1, 2026). This book—which is geared towards scholarly, activist, not-for-profit, and policymaker audiences—explores five cases of media activism by leading anti-violence activists across Canada. These cases include Mandi Gray and Kelly Showker's indie documentary, *Slut or Nut*, which covers the lead up to Mandi's experience of a rape trial; Farrah Khan's yearly social media memorialization of femicide victims on the anniversary of the "Montreal Massacre"; Leah Parsons' Memory Stones Project, as she establishes digitally enhanced "mini-memorials" around the globe to raise awareness about GBV's numerous consequences for survivors; and Jaime Black's and Destin Running Rabbit's compelling art installations, the REDress Project and Disposable Red Woman, which respond to the Missing and Murdered Indigenous Women and Girls crisis. Each intervention incorporates dynamic multimedia and collaborative components in service of addressing GBV—which is something the book highlights. For example, Black's and Running Rabbit's installations meld red dresses with QR code technologies or videography with so-called "corpse art"; Gray's work interweaves her video journals, taken throughout her trial, into the documentary, to demonstrate sexual assault's true consequences for survivors; Parsons' multi-modal Memory Stones Project engages both prettily painted beach stones with website and social media content, to extend her anti-violence messaging's reach...and so on.

While the book will offer a point of departure for our conversation, roundtable contributors will share insights from their own work on GBV-related media activism across transnational contexts. Some questions that will fuel our reflections include: how are today's activists uniquely responding to gendered violence using media and emerging technologies? Maintaining an intersectional lens, that is sensitive to the ways identity complicates one's experiences of the world, what struggles do these activists face in the process? What impacts will the emergence of AI have on mediated activism? And how can we—as scholars and advocates ourselves—support advocates' meaningful media work, while pushing for greater systemic change?

Joining me in this panel are Drs. Jessical Ringrose, Jaigris Hodson, and Esteban Morales.

Prof Ringrose has worked in London at UCL UK for 22 years, but she is a Canadian and has supported international research, advocacy and activism to challenge tech-facilitated gender-based violence (TFGBV) including in Canada, Australia, New Zealand, Peru, Hong Kong among others. She will discuss the challenges of creating legal, policy and social change in this area and her co-production public engagement work and educational strategies that have helped to raise awareness and support resistance against sexism, misogyny, transphobia, and other forms of inequity.

Dr Jaigris Hodson is an associate professor and Canada Research Chair whose work critically examines how misinformation, polarization, and anti-social behaviours spread through social media platforms and the vulnerabilities these networks create for individuals and communities. She will contribute on topics of technology-facilitated violence, online misogyny, and responses to these societal issues from both activists and policymakers. She will discuss how digital environments, including social media and algorithms, can enable and sustain violent discourses, but also strategies for mitigation and resilience through communication design and civic engagement.

Dr. Esteban Morales, meanwhile, will contribute critical expertise concerning how digital media environments shape user experiences of harm, violence, and power—but also create opportunities for everyday citizens to resist, in particular through enhanced media literacy. All scholars above are recognized experts in the field of online/offline misogyny and the use of media and technologies by activists to address the world's rising tides of gendered and technology-facilitated violence. They are all particularly attuned to issues of identity: how intersecting identities shape users' experiences of media, and opportunities for speaking back.

Dr. Nicolette Little is the author of *From Red Dresses to Memory Stones: Multimedia Activism and Gender-Based Violence in Canada*. As the author of this work, she is able to speak to it, and the 5 cases of unique multimedia activism it explores, with nuance and clarity. As an advisor to national media (e.g., The Toronto Star and CBC), law enforcement (e.g., Toronto Police Services) and the federal government (e.g., Public Safety Canada and the Federal Status of Women Committee) on topics related to GBV, feminist media activism to raise awareness about it, toxic technocultures, and interventions in manosphere-brewed violence, Nicolette is positioned to speak both specifically about mediated interventions in GBV, as well as about the broader media ecology that is making this activism so necessary (and challenging).

Because the included scholars are leading ones in the field of toxic technocultures, technology-facilitated violence, and gender-based violence interventions – and because we work across North American, UK, and/or European contexts – this conversation promises to be as expansive as it is deep and rich. Since each of us has

supported interventions in GBV, and engaged in advocacy ourselves, this also lends a more fulsome background experience to our conversation.

Session 6A - Digital Education

Day 2 · Tues 14 July · 9:30–11:00, RD Suite, ARC

6A.1 Where do UK universities represent themselves online? The institutional social media landscape post-Twitter

Katy Jordan (Lancaster University)

Introduction

Over the past two decades, social media use has become ingrained within the higher education sector in a range of different ways. In the context of higher education, there is an established body of research literature examining its use to support learning and teaching, or research, impact and other scholarly activities. The focus has typically been on understanding its use from an individual academic or student perspectives - which is logical, given that it is often individuals' personal accounts being used. However, Universities also present themselves - as institutions - on social media platforms, but this perspective is rarely examined.

Some existing studies have considered institutional use of social media, but the field is sparse and typically limited to one platform. Jordan (2017) mapped the network of connections between UK higher educational institutions (HEIs) on Twitter; at the time of data collection, all were represented on the platform. Taking an international sample, Shields (2015) showed that HEIs positioning within the Twitter network reflected academic prestige according to QS rankings. Al-Tamini et al. (2024a) present an analysis of the use of Instagram by UK HEIs. Greenwood (2012) reported an early example of a study seeking to map the range of social media referred to by institutions, focusing on the top 100 colleges in the USA. Findings showed that "Facebook, Twitter, YouTube, iTunes, Flickr, LinkedIn, and Foursquare" were most prevalent, although a number of others were also recorded. Although this study stands out as a multi-platform example, the social media landscape has changed quite considerably since the study was undertaken. Al-Tamini et al. (2024b) focus on the example of one UK HEI, Sheffield Hallam University, to take a detailed look at the range of platforms it uses, and the levels of traffic which each generate, which varied substantially according to platforms.

Social media platforms offer HEIs an opportunity to extend their online presence beyond their own website domain, and as such, student recruitment is often a key reason for its use (Evans, 2024; Gordon & Berhow, 2009). UK HEIs have been observed to engage with platforms such as Weibo and WeChat specifically with the aim of recruiting students from China (Zhu, 2019). However, while social media offers potential for dialogue, its use by institutions often remains broadcast in nature (Capriotti & Zeler, 2023). Yang et al. (2024) analysed a large-scale collection of Twitter posts made by 94 UK universities to examine levels of engagement with content, and demonstrated that emotion is important, with affective content garnering higher levels of engagement. Pawar and Islam (2025) conducted research with international students to model the process of engaging with institutional social media on their decision making. Notably, LinkedIn was found to be particularly useful, and differences according to demographic factors were not found

Objectives/Research Questions

While the use of social media by higher education institutions is under-studied in the existing literature, there are two reasons why it is particularly timely to examine this. First, to help understand the impact of the recent changes at X (formerly Twitter), which had previously been a mainstay of institutional communications. Second, the UK higher education sector is currently in a state of financial crisis, and examining the platforms used by universities may help shed light on ways in which institutions are responding to these pressures, particularly given the potential for social media to be used in student recruitment.

Until recently, the use of X/Twitter was almost synonymous with ‘social media’ itself in the context of higher education and its widespread use by academics, to support learning, teaching and research impact (Jordan & Carrigan, 2024). Use of the platform by academics began to decline from 2022 onwards, following the acquisition by Elon Musk and rebranding as X (Jordan & Carrigan, 2024; Swogger, 2023). The move away from X was initially quite slow, reflecting some reluctance to part with networks and communities which had been cultivated over time (Carrigan, 2023; Koutropoulos et al., 2024). A number of alternative platforms emerged, notably Mastodon and Bluesky (Koutropoulos et al, 2024; Shiffman & Wester, 2025), but it is not clear whether the level of engagement and use is sufficient yet to be an effective replacement. However, higher education institutions seem to be notably more reticent toward moving away from X (Tattersall, 2025). Some have reportedly opted for reduced usage of X rather than complete disengagement, while exploring additional alternative platforms (Spencer, 2025).

In light of the recent shifts in the social media landscape, and pressures on the higher education sector at the moment, the overall goal of this study is to understand the landscape of social media tools which UK universities currently use and foreground as

part of their main institutional web presence. This is further broken down into the following specific research questions:

1. Which social media platforms do UK universities feature on their institutional web pages?

- 1a: Are there any patterns in the use of combinations of platforms?

- 1b: To what extent has there been a migration away from X/Twitter?

2. How do follower-type metrics compare across the platforms?

Methods

Data collection took place in May and June 2025. The first step involved populating a list of UK higher education institutions to form the sample. The list of 131 institutions which are included in the Times UK university rankings was used as a starting point (The Times, 2025). Ranking was also recorded. Three prominent distance and part-time education institutions which are not included in the ranking were added. A total of 134 institutions were included at this stage.

Once the list had been populated, the second step involved systematic data collection from social media linked from institutional web pages. The main institutional homepage for each institution was visited, and the platforms which they include links or buttons to on the homepage, were recorded in a spreadsheet. These are typically included as an array in the footer of the page. Note that only platforms and accounts explicitly linked from the homepage were included - it is possible that institutions may have more, but for consistency and focus on those which they opt to foreground, only ones which are included in the homepage were logged. Given that the sampling focused exclusively on web pages and accounts which were institutions' principal accounts, and the main, public-facing representation of each university, there is a reasonable expectation that these sources are intended to be public and as such there are no ethical concerns associated with using this data source (BERA, 2024).

Seven institutions did not have any social media links or buttons on their main homepage, resulting in a final sample of 127 universities. For this sample, further data was collected in terms of the number of followers, or equivalent metric, per platform.

The statistical software package SPSS was used for the analyses. Following import of the data into SPSS, the dataset was then first analysed using descriptive statistics, in order to understand the overall characteristics of the data (Marshall & Jonker, 2010). In order to move beyond considering individual platforms and examine whether there are patterns in use across platforms, k-means cluster analysis was applied (Erbacher, 2020; Everitt et al., 2011). The range of platforms were used as variables for cluster analysis, each recorded as a binary variable of whether the institution used the platform or not. A two-step approach to cluster analysis was used, first undertaking hierarchical

cluster analysis to gain a sense of an appropriate number of clusters (seven was identified as a good fit), followed by k-means cluster analysis to examine the clusters in further detail.

Results

The first research question considered which social media platforms UK universities feature on their institutional web pages. For the final sample of 127 universities, the number of social media platforms featured on the main institutional homepage ranged from three to twelve (Figure 1 - see accompanying PDF). The distribution demonstrates a slight positive skew, with a modal frequency of six.

26 unique platforms or forms of social media were featured on at least one institutional home page. The frequency count per platform is shown in Table 1, and graphically in Figure 2. These demonstrate that while a relatively wide range of platforms were referred to by the sample as a whole, the frequency of use varied to a great extent. Three platforms - Instagram, Facebook, YouTube - were used by all or all but one of the institutions. These are followed by a further three high-use platforms, used by more than two-thirds of institutions: LinkedIn, X and TikTok. Apart from this core of six platforms, the majority of remaining platforms were used by fewer than five institutions.

On sub-question 1a and the move away from X specifically, the frequencies in Table 1 suggest a decrease is underway. While all UK universities were observed to have a Twitter account previously (Jordan, 2017), the percentage of institutions including this on their home page now stands at 79.5%. In terms of alternative platforms with similar functionalities and affordances, Bluesky and Threads are potentially being used to fill this gap, but uptake remains low at 10.2% and 6.3% respectively (no instances of Mastodon were found). Of the 13 institutions now using Bluesky, approximately half (six) had removed X while seven still include it. Similarly, of the eight using Threads, half use X while half do not.

Cluster analysis was applied in order to address sub-question 1b. Hierarchical cluster analysis was first used in order to identify a likely number of clusters (seven was identified as a good fit), before running a k-means cluster analysis using this figure. The number of institutions which were allocated to each cluster, from 1 to 7, are shown in Table 2. The clusters demonstrate different patterns in terms of the platforms used - the platforms which were used by all institutions in a cluster are shown by grey cells in Table 3.

The second question takes the next step beyond whether platforms are used or not, to start to examine the differing extent of engagement at different platforms and how follower-type metrics compare. As a broad indicator of the size of audience at each platform, per institution, the focus here is on follower counts or similar metrics.

Follower counts vary considerably according to platform and institution. An overview is provided by the descriptive statistics shown in Table 4. Note that several platforms do not provide exact figures above certain thresholds, instead rounding to the nearest hundred or thousand for larger follower counts.

Future Work

This presentation reports findings from a study which set out to examine the landscape of platforms being used by UK universities to represent themselves through social media, following broader changes in levels of usage of X and the migration to alternative platforms. Research on higher education social media use at the level of representing HEIs themselves is scant, and this study sheds light upon this by mapping the platforms used to formally represent universities. This further reveals some distinct patterns in platform usage, which may have practical implications for universities considering their social media strategy and representation. The findings suggest that while there is evidence of some movement in the sector away from X, it is not a dramatic shift yet and universities may be reluctant to give up social capital built on the platform. The variation in follower metrics suggest that while the traditionally most popular platforms still command highest audiences, there are a range of platforms in use by HEIs and there is scope for additional audiences to be engaged with through the newer and less prominent platforms.

The study has two main limitations, which could be addressed through future work. The study relies upon the data which is available online through public institutional websites and institutional social media profiles. This provides an overview of whether or not universities use particular platforms, but does not shed light on the ways in which they use them, or the relative effectiveness of engaging with audiences at different platforms. It would be valuable for future studies to carry out content analysis of posts at different platforms, and detailed examination of the extent and nature of engagement with the posts. The second main limitation is that the study focuses solely on institutions in the UK higher education sector. While the study focuses upon the UK higher education sector, the range of platforms are relevant globally. Student recruitment and enhanced research impact are perennial concerns. Further international and comparative work would be useful to establish to what extent the trends shown here are similar in other sectors.

Despite the limitations, this study provides an overview of the landscape of the platforms that UK higher education institutions use to represent themselves online, following a turbulent period for social media. The trends will be discussed in further detail in the presentation. Overall, it is notable that there is a relatively large range of platforms being used, and this opens up new areas for further research on the relationship between higher education and social media, to understand how a wider

range of platforms are being used, why institutions engage with particular platforms, and how engagement varies, particularly with newer social media platforms.

References

- Al-Tamini, A.-K., Salem, M., Alwadan, M., & Rodan, A. (2024a). The Power of the Gram: Leveraging Instagram for Effective Communication in Higher Education. The international Social Media for Learning in Higher Education conference (SocMedHE24), Sheffield Hallam University, UK, 19 December 2024. <https://shura.shu.ac.uk/35619/1/PowerOfGram.pdf>
- Al Tamini, A.-K., Javed, Y., Eliwa, E., & Bani Abdelrahman, R. (2024b). Connecting Beyond Borders: Understanding UK Universities' Social Media Strategies for Global Engagement. The International Social Media for Learning in Higher Education conference (SocMedHE24), Sheffield Hallam University. <https://shura.shu.ac.uk/35620/1/BeyondBorders.pdf>
- BERA (2024). Ethical Guidelines for Educational Research, fifth edition (2024). British Educational Research Association. <https://www.bera.ac.uk/publication/ethical-guidelines-for-educational-research-fifth-edition-2024>
- Capriotti, P., & Zeler, I. (2023). Analysing effective social media communication in higher education institutions. *Humanities and Social Sciences Communications*, 10(1), 656. <https://doi.org/10.1057/s41599-023-02187-8>
- Carrigan, M. (2023) It is time for academics to let go of Twitter. <https://markcarrigan.net/2023/04/21/it-is-time-for-academics-to-let-go-of-twitter/>
- Erbacher, M. (2020) Cluster Analysis. In: Atkinson, P., Delamont, S., Cernat, A., Sakshaug, J.W. & Williams, R.A. (Eds.), *SAGE Research Methods Foundations*. SAGE. <https://doi.org/10.4135/9781526421036846489>
- Evans, K. (2024) Refined vs Raw : Utilizing TikTok for recruitment. The International Social Media for Learning in Higher Education conference (SocMedHE24), Sheffield Hallam University.
- Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis*. Wiley.
- Gordon, J. & Berhow, S. 2009. University websites and dialogic features for building relationships with potential students. *Public Relations Review*, 35(2), 150-152.
- Greenwood, G. (2012) Examining the presence of social media on university web sites. *Journal of College Admission*, 216, 24-28.

- Jordan, K. (2017). Examining the UK higher education sector through the network of institutional accounts on Twitter. *First Monday*.
<https://doi.org/10.5210/fm.v22i5.7133>
- Jordan, K., & Carrigan, M. (2024). Special Collection Editorial – Social Media in Higher Education: What’s Happening? *Journal of Interactive Media in Education*, 2024(1), 6. <https://doi.org/10.5334/jime.911>
- Koutropoulos, A., Stewart, B., Singh, L., Sinfield, S., Burns, T., Abegglen, S., Hamon, K., Honeychurch, S., & Bozkurt, A. (2024). Lines of Flight: The Digital Fragmenting of Educational Networks. *Journal of Interactive Media in Education*, 2024(1).
<https://doi.org/10.5334/jime.850>
- Marshall, G., & Jonker, L. (2010). An introduction to descriptive statistics: A review and practical guide. *Radiography*, 16(4), e1–e7.
<https://doi.org/10.1016/j.radi.2010.01.001>
- Pawar, S. K., & Islam, T. (2025). Hierarchical effects of university social media posts on the enrolment decision-making of international students: Integrating the AIDA and TAM models. *Studies in Higher Education*, 50(12), 2980-2995.
<https://doi.org/10.1080/03075079.2024.2448530>
- Shields, R. (2016). Following the leader? Network models of “world-class” universities on Twitter. *Higher Education*, 71(2), 253–268. <https://doi.org/10.1007/s10734-015-9900-z>
- Shiffman, D.S. & Wester, J. (2025) Scientists no longer find Twitter professionally useful, and have switched to Bluesky, *Integrative and Comparative Biology*, 65(3), 538–545. <https://doi.org/10.1093/icb/icaf127>
- Spencer, W. (2025) Beyond X: How UK universities are adapting to a changing social media landscape. *The EduTimes*.
<https://edutimes.com/news/2025/03/20250321049>
- Swogger, S. 2023. The interactive web — leaving Twitter. *Journal of Electronic Resources in Medical Libraries*. 20(1), 28–32.
- Tattersall, A. (2025). UK Universities are reluctant to break a toxic social media relationship—Impact of Social Sciences. *Impact of Social Sciences - Maximizing the Impact of Academic Research*.
<https://blogs.lse.ac.uk/impactofsocialsciences/2025/02/24/uk-universities-are-reluctant-to-break-a-toxic-social-media-relationship/>
- The Times (2025) UK university rankings 2025. *The Times and The Sunday Times*.
<https://www.thetimes.com/uk-university-rankings>

- Yang, H.-P. (Sophie), Chan, T. K. H., Tran, H.-A., Nguyen, B., & Lin, H. (2024). See it, share it: What makes social media content viral in the higher education context? The power of positive affective content. *Information Technology & People*, 38(4), 1948–1968. <https://doi.org/10.1108/ITP-02-2023-0126>
- Zhu, Y. (2019). Social media engagement and Chinese international student recruitment: understanding how UK HEIs use Weibo and WeChat. *Journal of Marketing for Higher Education*, 29(2), 173–190.

6A.2 Student social media etiquette: Negotiating teacher authority in a Chinese university

Linan Tang (King's College London)

Introduction

Social media platforms such as WeChat and QQ have become infrastructural to everyday teacher-student communication in Chinese higher education, extending interaction beyond classrooms into group spaces and private messaging. This paper examines student social media etiquette as a form of relational work through which teacher authority is sustained and negotiated in routine digital interaction. Educational research on social network sites and out-of-class communication shows that such platforms reshape teacher-student relationships, including expectations of access, responsiveness, rapport and motivation (Camas Garrido et al., 2021; Elhay & Hershkovitz, 2019; Zheng et al., 2022).

Theoretically, the paper draws on the insight that authority is not a stable property of individuals but a relational accomplishment that depends on recognition by those subject to it (Kojève, 2020; Weber, 1979). On social media, this accomplishment is mediated through etiquette, meaning the practical norms students use to manage deference, distance and approachability. Etiquette becomes relational labour that aligns students with teachers' authority claims while managing interactional risk under platform visibility. In the Chinese context, such negotiation is shaped by *guanxi*, understood as relationship cultivation through reciprocal exchange, and by Confucian expectations that students show deference to teachers as moral elders (Bian, 2019). Platforms are not neutral channels. They organise who is visible, who can speak, and how tone, respect, care and discipline are expressed.

A central challenge for students is managing the risks of exposure and peer judgement in class group chats. One participant described this as being watched by silent audiences: 'It feels like all other members are "shijian" (视奸, lurking) on me when I talk with our teachers in the group chat.' This visibility encourages cautious self-presentation and careful boundary management. Platform features such as emojis, memes and red packets introduce affective resources that can soften interaction without dissolving hierarchy, but these cues are ambiguous and easily misread across generational lines, such as the 'death smile' emoji. Red packets operate as 'social currency' that can energise participation while reinforcing vertical relations (Wu & Ma, 2017; Xu, 2021).

This paper argues that teacher authority is interactionally accomplished through students' etiquette practices, including attention to address terms, message

composition, timing and adaptation to teachers' preferences. It contributes to social media studies by showing how hierarchy is reproduced through communicative micro-practices, not only through institutional roles or platform design.

Objectives/Research Questions

This paper asks: How does student social media etiquette shape the negotiation of teacher authority in a Chinese university context?

It addresses these linked questions:

Norms: What counts as 'proper' teacher-student communication on WeChat/QQ (e.g., politeness, timing, boundaries) and why?

Strategies: How do students use etiquette strategies (e.g., visibility management, channel choice, tone work) to organise teacher authority in everyday interaction?

Methods

The study uses a qualitative design at a Chinese university (site anonymised for review). It draws on three linked datasets: (1) focus groups with undergraduate students, (2) semi-structured interviews with teachers and (3) online observation of QQ group chats.

The etiquette analysis is based on recurring patterns across 65 student participants in 16 focus groups, alongside online observations. In line with the project's ethical approach, observational evidence is reported through anonymised and paraphrased fieldnotes rather than screenshots or verbatim chat logs.

Thematic analysis proceeded inductively through iterative coding and comparison across data sources. First, I mapped the communicative landscape (channels, content and roles shaping teacher-student contact). Second, I analysed how platform features (private/group chats; emojis/memes; red packets) structure participation, visibility and tone. Third, I synthesised students' accounts of 'what to pay attention to' when messaging teachers into two overarching etiquette forms: prioritising politeness and adapting to teachers' preferences.

Results

Findings show that student 'social media etiquette' is structured around two forms of relational work that help sustain and negotiate teacher authority in everyday WeChat/QQ interaction.

Prioritising politeness: address, wording, tone, timing, boundaries and closing gestures

Students treat teacher-student messaging as high-stakes, particularly in class group chats where talk is witnessed by peers. Some describe this as being 'shijian' (视奸, lurking) upon. This unease encourages careful self-presentation, including formal

address terms, softened wording, measured tone and cautious timing (e.g., avoiding late-night messages). They manage boundaries by routing sensitive or potentially embarrassing questions into private chat while keeping public spaces for ‘safe’ coordination. Politeness is also performed through closing gestures (thanks, apologies, ritual farewells) and cautious affective signalling. Emojis and memes may soften requests, but students use them selectively to avoid appearing disrespectful. Teachers’ red packets in group chats likewise invite polite collective responses (thanking, quick participation), reinforcing a benevolent tone without dissolving hierarchy.

Adapting to teachers’ preferences: platforms, style and responsiveness

Beyond general politeness, students emphasise adapting to individual teachers’ preferences. They monitor which platform a teacher uses (WeChat vs QQ), whether they prefer short texts or voice notes, how formal their style is, and what counts as ‘prompt’ responding. Students then calibrate message length, frequency and formality accordingly, often mirroring teachers’ style as a ‘safe’ standard. Semiotic uncertainty makes this calibration salient. The jokingly named ‘death smile’ emoji illustrates how a teacher’s seemingly neutral token can be read by students as sarcastic or cold, leading some to avoid emojis unless a teacher uses them first. Students also adapt to teachers’ responsiveness (waiting longer, rephrasing or going via class representatives) rather than pressing publicly.

Overall, these two etiquette practices show how teacher authority is negotiated through interactional choices. Etiquette does not merely reflect hierarchy; it helps reproduce it in everyday social media use while enabling necessary coordination and support.

Future Work

Future work will extend this analysis in three directions.

First, I will develop a more explicit typology of platformed teacher authority by connecting etiquette practices to distinct authority sources (e.g., institutional role, expertise, care and moral expectations). This will clarify how students’ recognition of authority varies depending on whether they perceive a teacher as knowledgeable, caring or holding formal power, and how these sources interact in digital communication.

Second, I will compare how etiquette operates across different group chat types (class, module, check-in groups) to specify when platform visibility strengthens authority versus when it produces avoidance, fatigue or subtle contestation. Preliminary findings suggest that check-in groups, which involve voluntary participation and habit-tracking, may generate different etiquette dynamics than compulsory class groups, with implications for how soft forms of surveillance and motivation are experienced.

Third, I will link these everyday interaction norms to broader debates on platformisation and educational governance, including how incentives (e.g., red packets) blur boundaries between care, obligation and soft control. This connects student etiquette work to questions about how commercial platform affordances become enrolled in institutional authority relations.

References

- Bian, Y. (2019). *Guanxi: How China Works*. Polity Press.
- Camas Garrido, L., Valero Moya, A., & Vendrell Morancho, M. (2021). The Teacher-Student Relationship in the Use of Social Network Sites for Educational Purposes: A Systematic Review. *Journal of New Approaches in Educational Research*, 10(1), 137–156. <https://doi.org/10.7821/naer.2021.1.591>
- Elhay, A. A., & HersHKovitz, A. (2019). Teachers' perceptions of out-of-class communication, teacher-student relationship, and classroom environment. *Education and Information Technologies*, 24(1), 385–406. <https://doi.org/10.1007/s10639-018-9782-7>
- Kojève, A. (2020). *The Notion of Authority*. Verso Books.
- Weber, M. (1979). *Economy and Society: An Outline of Interpretive Sociology* (G. Roth & C. Wittich, Eds). University of California Press.
- Wu, Z., & Ma, X. (2017). Money as a Social Currency to Manage Group Dynamics: Red Packet Gifting in Chinese Online Communities. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2240–2247. <https://doi.org/10.1145/3027063.3053153>
- Xu, Y. (Aveva). (2021). The “Lucky Money” That Started It All—The Reinvention of the Ancient Tradition “Red Packet” in Digital Times. *Social Media + Society*, 7(3), 1–10. <https://doi.org/10.1177/20563051211041643>
- Zheng, W., Yu, F., & Wu, Y. J. (2022). Social media on blended learning: The effect of rapport and motivation. *Behaviour & Information Technology*, 41(9), 1941–1951. <https://doi.org/10.1080/0144929X.2021.1909140>

6A.3 Bluesky as a Teaching and Learning Tool for International Students in Hybrid Communication Courses

Wahiba Chair, Elva Fisher (University Canada West)

Introduction

Numerous studies point to the benefits of using X, formerly known as Twitter, for teaching and learning. A recent review analyzing over 100 X-related studies (including 80+ in higher education) revealed that X can be an effective and innovative teaching and learning tool with positive outcomes for students, educators, and institutions (Malik et al., 2019). X has been shown to enhance student engagement, promote collaborative learning, and support community building and information sharing (Junco et al., 2011; Carpenter & Krutka, 2015).

Since Elon Musk acquired the platform (New York Times, 2022), users, particularly academics, started migrating from X to alternative platforms, most notably Bluesky (Science, 2024). By early 2025, Bluesky had overtaken X in hosting posts related to newly published academic research (Times Higher Education, 2025). As of late 2025, Bluesky has grown to approximately 40 million users (Social Media Today, 2025), with 22% of researchers now using Bluesky as their primary social media channel (Digital Science, 2024).

However, empirical literature on Bluesky remains predominantly focused on scholarly communication and academic networking (Science, 2024)—not teaching and learning pedagogy.

This study may be among the first empirical investigations to examine Bluesky as a teaching and learning tool in formal higher education settings. Beyond platform novelty, the study addresses five additional research gaps on the pedagogical use of microblogging tools such as X.

First, most X pedagogy research has focused on applied sciences, with very few studies in Social Sciences or Communication (Malik et al., 2019).

Second, prior research with language learners has examined EFL contexts in home countries (Taskiran et al., 2018; Arumugam et al., 2022) rather than international students in English-speaking immersion environments, like Canada.

Third, existing studies have predominantly examined traditional in-person or fully online settings (Rohr et al., 2024) rather than hybrid learning environments.

Fourth, although AI writing tools are widely used by students (Grammarly, 2026), no research has examined how students use AI tools during microblogging-based learning activities.

Finally, while research has documented barriers international students face using X, few studies examine pedagogical strategies to mitigate these obstacles.

This work constitutes phase four of a multi-phase mixed methods study on microblogging tools for teaching and learning with international students taking communication courses in hybrid environments. Initiated in Spring 2024, the first three phases focused on X. In winter 2025, the investigators pivoted to Bluesky given the substantial academic migration and emerging interest in the platform. The PI has been using X (and recently Bluesky) for teaching and learning since 2014 and published guides and case studies on the topic. This study was funded by an explore-innovation in teaching research grant at UCW.

Objectives/Research Questions

The objective of this study is to examine international students' perceptions, experiences, adoption barriers, AI tool use, and future use intentions when using Bluesky as a learning tool in hybrid communication courses.

The primary research question is: How do international students perceive Bluesky's impact on their learning, writing skills, sense of community, and overall course experience?

Secondary questions explore:

- How did students feel about the activity overall, and did they enjoy it?
- What barriers, if any, do international students face when using Bluesky for learning?
- Which AI tools do students use in conjunction with Bluesky learning activities, and for what purposes?
- Does students' use of Bluesky in class influence their future use intentions?
- What pedagogical strategies can help increase adoption of Bluesky for learning?

The survey questionnaire is attached as a separate document; however, the branching and display logic are only visible when the survey is taken live.

Methods

Study Design & Context

This mixed methods study examined international students enrolled in two sections of COMM 205: Media Writing at University Canada West during winter 2025 (N = 58). The course followed a hybrid model: two in-person hours followed by two hours of online guided learning (OGL) per week.

Bluesky Activity

During OGL, students participated in a graded activity requiring a minimum of two posts per week using the course hashtag #UCWCOMM205: one original post (response to an instructor prompt) and one reply to a classmate. In the first class, students received a live demonstration of Bluesky and posted their first message with the instructor's guidance. Detailed instructions, community guidelines, and a grading rubric were shared in class and in the course LMS.

Pedagogical Strategies

Throughout the 10-week course, the instructor/PI and co-investigator posted weekly discussion prompts (e.g., reflective questions, resources), moderated discussions, and replied to student posts. Private feedback (e.g. reminder to use the hashtag) was provided as needed. In week 2, students were shown how to use their premium Grammarly university account to write better posts. Class time was dedicated to OGL peer debriefs, addressing student questions, and reminders to apply the course-taught 3Cs framework (clear, correct, concise) to posts

Recruitment & Sample

In Week 10, students were invited to complete a short optional online survey (via Qualtrics) assessing their perceptions of Bluesky's impact on their learning, writing skills, engagement, and sense of community. The survey also examined adoption barriers, motivating factors, AI tool usage, and future intentions.

45 students initiated the survey (78% response rate). After removing incomplete responses, 42 students were included in the analysis.

All 42 respondents were international students; 81% identified as ESL learners, 88% were aged 18–24, and 60% identified as female. 64% had not used X or Bluesky prior to the course (29% and 7% had previously used X and Bluesky, respectively). All survey respondents self-reported participating in the Bluesky activity.

Survey data were analyzed using descriptive statistics for closed-ended questions. Open-ended responses were analyzed qualitatively to identify key themes and patterns. Course hashtag analytics were collected in winter 2026 using bskykeywords.com

Ethics & Transparency

Since this study involves human participants, ethics approval was obtained from the university's research ethics board. Students were informed both during recruitment and in the survey's consent form that their participation was voluntary, did not affect grades, and that results would be aggregated and anonymized. To minimize the PI's influence, the instructor left class during recruitment (conducted by another faculty

member/co-investigator). Survey responses were de-identified and analyzed by a research assistant.

Results

Preliminary Results

Impact on Learning

About 74% of respondents agree or strongly agree that the Bluesky activity improved their learning and course experience. Benefits extend beyond the classroom: over half reported that the activity helped them become more familiar with Canadian news and culture as well as improve their media and information literacy. These skills are essential for international students navigating a new culture in a complex media landscape.

Writing

About 64% of students indicated that the Bluesky activity helped them improve their English writing skills. The activity also gave students an opportunity to consistently apply key class concepts; over 75% agree or strongly agree that the Bluesky activity helped them write better by applying the 3Cs. Together, these patterns suggest that positioning Bluesky as a low-stakes space to write and connect can support writing skills for international students in a hybrid environment.

Connections

Nearly 80% of students agreed that the Bluesky activity helped them feel less isolated with a greater sense of community, an important outcome given international students' higher risk of social isolation in blended settings. Nearly 70% agreed that it helped them connect with their peers and instructors, which is consistent with the requirement to post weekly replies to both. Although the activity was graded, the most common reason for participating was 'getting likes and replies from my peers and instructors,' underscoring the value of these social connections. 83% agreed that the activity was fun and enjoyable. This aligns with very positive sentiment for #UCWCOMM205 posts as reported by bskykeywords.com

Pedagogy

The Bluesky activity was complemented by in-class strategies to improve participation. Over 80% agree or strongly agree that peers' OGL debriefs in class helped them be more engaged and better understand the activity. Instructor support is also important: 83% of students agree or strongly agree that the instructor's feedback on the Bluesky activity, in class and privately, helped them write better posts. No barriers were identified.

AI

While many students use AI tools like Grammarly and ChatGPT for general coursework, fewer reported using AI specifically for the Bluesky activity. Students who did use AI primarily relied on Grammarly, likely influenced by the in-class demo. They mostly used AI to ensure correct grammar and spelling rather than to generate ideas or write full posts. This pattern suggests that the Bluesky activity encouraged intentional, responsible use of AI rather than over-reliance.

Future Use

Most students showed openness to continuing with Bluesky: over 70% are interested in using it in future courses. Over half also plan to use it outside their courses. This is notable given that over 90% of respondents had not used Bluesky before the term, suggesting that the course activity introduced most students to the platform and helped spark ongoing interest.

Future Work

While this study shows promising preliminary results, a few known limitations can be addressed in future research.

- 1) Compare Bluesky for Learning with other microblogging platforms (e.g., X) in similar courses and settings. This includes drawing on prior studies as well as comparing the current phase (4) with earlier phases (1–3), which have larger samples. We are currently analyzing data from phases 1–3 to enable this comparison.
- 2) Although data were collected through bskykeywords.com, further work should integrate these analytics with survey data to develop a more comprehensive understanding of engagement and learning outcomes.
- 3) The study relies on self-reported survey data without direct measures of writing quality or long-term learning. Future studies could combine surveys with performance indicators (e.g., graded writing, discourse analysis) and include longitudinal follow-ups to examine whether Bluesky use predicts sustained community engagement, writing practice, or platform adoption.
- 4) Of the 58 students enrolled, 42 completed the survey. The 16 non-respondents may have faced different or greater barriers to participation. Larger-scale studies, content analysis of posts, and targeted follow-ups with non-participants could clarify how instructor feedback affects contribution quality and uncover under-reported participation barriers.

Finally, because this study involved a small sample of international students from a single institution over one term, and because ethnic diversity was not measured, replicating it across more institutions and in comparable courses with larger, more diverse samples would strengthen generalizability.

References

- Arumugam, N., Shafiqah, N. S., Shanthi, A., Idham, M., & Purwarno. (2022). The use of Twitter in learning English vocabulary: EFL learners. *International Journal of Academic Research in Business and Social Sciences*, 12(6), 1573–1589.
<https://doi.org/10.6007/IJARBSS/v12-i6/14163>
- Carpenter, J. P., & Krutka, D. G. (2015). Engagement through Twitter: A survey of educators' social media practices. *Teacher Development*, 19(3), 443–459.
<https://doi.org/10.1080/19415257.2014.939294>
- Digital Science. (2024, December). Altmetric adds Bluesky social media as a research attention source. Digital Science Blog. <https://www.digital-science.com/blog/2024/12/altmetric-adds-bluesky-social-media-as-a-research-attention-source/>
- Grammarly. (2026). Grammarly for education [Webpage]. Grammarly.
<https://www.grammarly.com/edu>
- Junco, R., Heiberger, G., & Loken, E. (2011). The effect of Twitter on college student engagement and grades. *Journal of Computer Assisted Learning*, 27(2), 119–132.
<https://doi.org/10.1111/j.1365-2729.2010.00387.x>
- Malik, A., Heyman-Schrum, C., & Johri, A. (2019). Use of Twitter across educational settings: A review of the literature. *International Journal of Educational Technology in Higher Education*, 16(1), 36. <https://doi.org/10.1186/s41239-019-0166-x>
- New York Times. (2022, October 27). Elon Musk completes 44 billion dollar deal to own Twitter. The New York Times.
<https://www.nytimes.com/2022/10/27/technology/elon-musk-twitter-deal-complete.html>
- Rohr, L. E., Costello, J., & Squires, L. (2024). Exploring students' Twitter use in the online classroom across 4 years. *E-Learning and Digital Media*, 21(3), 183–201.
<https://doi.org/10.1177/20427530231167644>
- Science. (2024). Old Twitter: Scientific community finds new home on Bluesky. Science. <https://www.science.org/content/article/old-twitter-scientific-community-finds-new-home-bluesky>
- Social Media Today. (2025). Bluesky reaches 40 million users. Social Media Today.
<https://www.socialmediatoday.com/news/bluesky-reaches-40-million-users/804467/>

Taskiran, A., Koral Gümüsoğlu, E., & Aydin, B. (2018). Fostering foreign language learning with Twitter: Reflections from English learners. Turkish Online Journal of Distance Education, 19(1), 100–116. <https://doi.org/10.17718/tojde.382734>

Times Higher Education. (2025). X's dominance over Bluesky becomes new hub for research. Times Higher Education.
<https://www.timeshighereducation.com/news/xs-dominance-over-bluesky-becomes-new-hub-research>

6A.4 AI, Social Media, and Higher Education: A Comparative Policy Analysis across UK, India, and South Africa

Devina Sarwatay, Anubha Sarkar (City St George's, University of London), Swarantra A (Indian Institute of Management Indore), Kealeboga Aiseng (University of the Witwatersrand)

Introduction

AI and GenAI technologies have become central to digital ecosystems in which students and educators operate. Social media platforms now serve as crucial sites where learners first encounter information about these tools, prompting strategies and best practices. Yet, as these practices evolve rapidly and circulate virally, Higher-Education-Institutions (HEIs) and National Bodies are simultaneously publishing guidelines aimed at framing responsible and ethical AI use in formal education. This creates a growing tension between the informal, socially shared norms of AI use shaped by platform cultures and the formal expectations articulated in policy documents.

National and institutional guidance has expanded quickly to promote safe, ethical and effective use, with UK Department for Education highlighting both opportunities (e.g., workload reduction) and risks (e.g., bias, inaccuracy, and privacy), and signposting support for education settings (DfE UK, 2025). Within universities, we observe student-facing expectations for ethical use and integrity and staff-facing guardrails and AI-in-education principles to support teaching, assessment design, and upskilling.

In India, AI policy is framed through #AIforAll, a national strategy to leverage AI for inclusive development (including education) and reinforced through new governance guidance that seeks to balance innovation with accountability and safety (MEITY, Gol, 2025; NITI Aayog, 2018). Sector reports document accelerating adoption and recommend phased, responsible integration across Indian HEIs (EY-Parthenon & FICCI, 2025). Doctoral-level policies in some HEIs tighten authorship/originality norms while allowing limited, transparent uses of GenAI in research processes.

In South Africa, Universities South Africa convenes cross-institutional dialogue on policy and implementation (USAf, 2025), while some universities provide comprehensive guidance spanning ethics, bias, privacy, learning, assessment, and research. Despite these developments, comparative cross-regional analyses of AI-in-education policy, especially in relation to social-media-mediated learning remain limited.

Existing scholarship reveals that digital and AI literacies are not merely technical capacities but socio-cultural, multimodal, and power-laden processes (Buckingham, 2019; Oozeerally et al., 2020; Thomas et al., 2007). Research on algorithmic cultures demonstrates that AI systems redistribute visibility, authority, and labour in uneven ways (Pangrazio et al., 2023; Williamson, 2021).

This paper examines how formal documents conceptualise AI literacy and responsible use and whether they recognise that much AI learning is platform-mediated (e.g., short-form video tutorials and creator prompts) in spaces like TikTok, Instagram, and YouTube. We also examine how formal policy documents frame responsibilities, risks, and expectations of AI-in-education adoption.

Objectives/Research Questions

In this paper we unpack how higher-education policies governing AI use are being formulated and operationalised across different national and institutional contexts, and how these formal documents engage with the platformised environments in which AI learning and practice increasingly occur. Drawing on a comparative perspective across the UK, India, and South Africa, we uncover both what policies explicitly articulate and what they implicitly assume or overlook, particularly in relation to social media-mediated forms of learning, norm formation, and ethical practice.

In this paper, we ask the following questions:

1. How do AI-in-education policies and institutional guidelines in the UK, India, and South Africa conceptualise AI literacy, responsible/ethical use, and academic integrity, and to what extent do they acknowledge social-media-mediated learning?

This question examines how AI competence is defined and bounded in policy, and whether learning through platforms such as YouTube, TikTok, and Instagram is recognised as part of contemporary AI literacy.

2. What assumptions do they make about learner competencies, staff responsibilities, and disclosure/transparency?

Here, we explore how responsibility for ethical AI use is distributed across students, educators, and institutions, including expectations around skills, supervision, and transparency.

3. How do national framings (UK governance and safety; India's #AIforAll plus governance guidelines; South African sector coordination) shape institutional expectations?

This question situates institutional policies within broader national discourses, policy priorities, and regulatory cultures.

4. Where and how are platform-centric learning practices recognised or omitted?

Finally, we identify silences and misalignments between formal governance and the informal, platform-driven practices through which AI knowledge is increasingly produced and shared.

Methods

This study employs a comparative qualitative policy analysis across three contexts: the United Kingdom, India, and South Africa, to examine how AI-in-education is conceptualised, governed, and operationalised at both national and institutional levels. The corpus comprises national strategy documents, sector guidance, and publicly available institutional policies and guidelines produced by higher-education institutions and regulatory or coordinating bodies in each region. Document selection was guided by relevance to higher education and generative AI, recency, and representativeness within each national context.

All documents were systematically coded using a thematic analysis approach. An initial coding framework was developed deductively from the literature on digital and AI literacies, platform governance, and higher-education policy, and iteratively refined through close reading. Core analytic categories included: literacy and skills frameworks; integrity, governance, and accountability; equity and access; research and doctoral norms; teaching and assessment design; and explicit or implicit recognition of social-media-mediated learning practices. Coding was conducted comparatively to identify both convergences across regions and context-specific divergences in policy emphasis, underlying assumptions, and regulatory intent.

Comparative analysis focused not only on stated principles and rules but also on silences and omissions within policy texts. Attention was paid to how, and whether, policies acknowledge that much AI learning and norm formation occurs through platform cultures such as short-form video tutorials, creator-led guides, and peer-to-peer knowledge sharing on platforms including YouTube, TikTok, and Instagram. By examining alignments and tensions between formal institutional governance and informal, platform-mediated practices, the method enables an assessment of the extent to which current AI-in-education policies meaningfully engage with the social media environments in which AI literacy is increasingly developed.

Results

The following are preliminary results from the cross-country policy analysis:

1. Shared core: “responsible, transparent, integrity-first”

Across regions, documents converge on responsible use, transparency/disclosure, and academic integrity. They caution on bias, data protection, and reliability.

2. Distinct national framings

UK: Emphasis on safe, effective adoption to reduce workload and support learning, with practical implementation resources. Institutional guidance operationalises principles for students, doctoral researchers, and staff.

India: Strategic vision for inclusive growth (#AIforAll) plus new governance guidelines foregrounding safety, accountability, and action plans. Institutional rules set stricter authorship/originality norms in doctoral work. Sector analysis indicates widespread, strategic uptake in HEIs.

South Africa: Sector-level coordination via USAf. Comprehensive institutional guidelines covering ethics, bias, privacy, learning, assessment, and research.

3. Social-media-mediated learning: policy blind spot

While guidance is increasingly granular (e.g., assessment permissions, disclosure), documents seldom treat TikTok/Instagram/YouTube as structured AI-learning spaces shaping norms (prompting, heuristics, creator-led practices). This gap appears across UK, India, and South Africa sources reviewed.

4. Doctoral integrity and research transparency

Doctoral-level guidance tightens expectations: UK HEI allows regulated, discussed use with supervisor oversight and potential viva scrutiny. India HEI prohibits thesis text generation while permitting limited, explicitly disclosed uses.

5. Implementation levers

UK institutional materials (principles, guardrails, student guidance) and South Africa guidelines provide actionable levers (assessment design, disclosure templates, data-privacy cautions) that can be adapted cross-regionally.

Taken together, these findings indicate a persistent disconnect between institutional AI governance and the platform cultures through which AI knowledge is socially produced and normalised.

Future Work

Building on the comparative policy analysis presented here, future work will advance in four interconnected directions.

First, the project will develop an AI-literacy framework that explicitly integrates platform-centric learning practices such as short-form video explainers, creator-led tutorials, and prompt-sharing cultures with institutional governance mechanisms, including academic integrity policies, disclosure expectations, and doctoral norms. This framework will be informed by adaptable templates drawn from UK and South African institutional guidance and refined for broader higher-education contexts.

Second, the research will undertake a systematic content analysis of AI-focused creator tutorials and knowledge-sharing practices on platforms such as YouTube, Instagram, and TikTok. This analysis will surface vernacular prompting techniques, informal norms of ethical use, and peer-driven learning heuristics, which will then be compared with formal policy prescriptions. Through this work, social media platforms will be theorised as informal policy actors that actively shape AI practice within higher education.

Third, the project will translate empirical and comparative insights into policy-translation toolkits, including assessment statements, disclosure rubrics, and data-protection checklists. These resources will be aligned with national framings, including UK Department for Education guidance, India's #AIforAll and associated governance guidelines, and South Africa's sector-level coordination efforts, enabling institutions to bridge policy intent and everyday practice.

Finally, these strands converge in Hacking HigherEd (work in progress), a multi-country, participatory, ethical, and inclusive AI-for-Higher-Education project spanning the UK, India, and South Africa. This programme will scale its co-designed AI Toolkit, comparative policy insights, and sustainability-focused, justice-oriented pedagogic frameworks into a longitudinal initiative supporting inclusive AI adoption, capacity-building, and sector-wide transformation.

References

- Buckingham, D. (2019). Teaching media in a 'post-truth' age: Fake news, media bias and the challenge for media/digital literacy education / La enseñanza mediática en la era de la posverdad: fake news, sesgo mediático y el reto para la educación en materia de alfabetización mediática y digital. *Culture and Education*, 31(2), 213–231. <https://doi.org/10.1080/11356405.2019.1603814>
- DfE UK. (2025). Generative artificial intelligence (AI) in education. GOV.UK. <https://www.gov.uk/government/publications/generative-artificial-intelligence-in-education/generative-artificial-intelligence-ai-in-education>
- EY-Parthenon & FICCI. (2025). Future-ready campuses: Unlocking the power of AI in higher education. <https://www.ey.com/content/dam/ey-unified-site/ey-com/en-in/insights/education/documents/ey-harnessing-ai-in-higher-education-opportunities-and-the-road-ahead-v1.pdf>
- MEITY, GoI. (2025). India AI Governance Guidelines: Enabling Safe and Trusted AI Innovation. <https://static.pib.gov.in/WriteReadData/specificdocs/documents/2025/nov/doc2025115685601.pdf>

NITI Aayog. (2018). National Strategy for Artificial Intelligence.

<https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf>

Oozeerally, S., Ramma, Y., & Bhoola, A. (2020). Multiliteracies—New London Group. In B. Akpan & T. J. Kennedy (Eds), *Science Education in Theory and Practice: An Introductory Guide to Learning Theory* (pp. 323–342). Springer International Publishing. https://doi.org/10.1007/978-3-030-43620-9_22

Pangrazio, L., Selwyn, N., & Cumbo, B. (2023). Tracking technology: Exploring student experiences of school datafication. *Cambridge Journal of Education*, 53(6), 847–862. <https://doi.org/10.1080/0305764X.2023.2215194>

Thomas, S., Joseph, C., Laccetti, J., Mason, B., Mills, S., Perril, S., & Pullinger, K. (2007). *Transliteracy: Crossing divides*. First Monday. <https://doi.org/10.5210/fm.v12i12.2060>

USAf. (2025, June 18). Institutional AI Policies and Guidelines in South Africa for Learning and Teaching. Universities South Africa. <https://usaf.ac.za/institutional-ai-policies-and-guidelines-in-south-africa-for-learning-and-teaching/>

Williamson, B. (2021). Digital policy sociology: Software and science in data-intensive precision education. *Critical Studies in Education*, 62(3), 354–370. <https://doi.org/10.1080/17508487.2019.1691030>

Session 6B - Promotional Social Media, Branding, & Marketing

Day 2 · Tues 14 July · 9:30–11:00, Seminar Suite A, ARC

6B.1 Children in the Creator Economy of Family Influencers: Visibility, Engagement, and Regulatory Gaps in Italy

Elisabetta Locatelli (Università Cattolica del Sacro Cuore)

Introduction

The creator economy has rapidly grown into a well-established economic sector governed by its own platform-driven and market-oriented logics. Creators typically earn revenue through partnerships with brands that pay them to publish sponsored content on their profiles (De Veirman et al., 2017). To secure sufficient visibility and engagement to make this model profitable, creators must maintain a continuous editorial plan of organic content, thereby sustaining a recognizable online presence through consistent self-presentation (Arriagada & Ibáñez, 2020). Existing research has highlighted both the benefits and the risks of creator labor, including the commodification of the self and of intimacy and the challenges associated with sustaining constant visibility (Abidin & Thompson, 2012; Duffy & Sawey, 2024).

Recent scholarship has begun to connect research on sharenting with studies on the creator economy, analyzing what happens when parents are family influencers, creators who build an entrepreneurial project around family storytelling (Jorge et al., 2022; Van den Abeele et al., 2024; Van den Abeele et al., 2025b). Influencer value is strongly tied to perceived credibility, authenticity, and intimacy, typically enacted through strategic self-disclosure on social media. For family influencers, these dynamics also involve sharing aspects of their children's everyday lives, including their images and personal information. On one hand, becoming a family influencer may enable parents, particularly women, to achieve a more flexible work-life balance and re-enter the workforce after pregnancy (Archer, 2019a, 2019b). On the other hand, this practice raises concerns about the commodification of family life and the potential exploitation of children's images, thereby heightening privacy risks (such as AI manipulation, digital kidnapping, or forms of sexting) and normalizing minors' presence in both organic and sponsored content (Abidin, 2017; Baxter & Czarnecka, 2025).

There are also relevant regulatory issues. When children appear in sponsored posts, they effectively engage in forms of labor, with both brands and parents acting as their employers. Some countries, such as France, have developed specific rules to protect minors in this context (BBC, 2020). In other countries, such as Italy, no equivalent protections exist. Minors may open a profile, with parental consent, starting at age 14, and their images may be shared by parents if both parents authorize the practice. This regulatory gap becomes even more problematic given the absence of reliable data on the actual prevalence of children's visibility within family influencers' content, and more generally within the broader influencer landscape. It therefore becomes essential to understand to what extent minors are involved in these forms of platform-mediated labor.

Objectives/Research Questions

This research project aims to investigate how children's visibility is mobilized within the influencer economy through the practices of family influencers. The study is part of a collaborative initiative involving an Italian university, an international NGO that has among its main goals to protect childhood, and an Italian regulatory independent institution with the support of an Italian social media consultancy agency [all names omitted for peer review]. This partnership offered a unique opportunity to combine academic rigor, the child-protection advocacy of a major NGO, the regulatory expertise of the advertising industry, and analytical data provided by the agency.

Given the sensitivity and potential risks associated with the visual portrayal of children in family influencers' content, the project first sought to understand to what extent children appear in the sampled profiles (RQ1). Considering that organic content is essential for establishing visibility and that sponsored content is crucial for monetization, the second goal was to assess differences in minors' presence between organic and sponsored posts (RQ2). Thirdly, given the pervasiveness of platform logics, we sought to understand whether children's presence affects engagement metrics in both organic and sponsored content (RQ3).

Methods

The project builds on two earlier studies: one conducted by the German branch of the ONG and another developed jointly by the author of this proposal and the Italian regulatory institution. The first work showed the presence of children's labor among influencers profiles in Germany, while the second involved mapping family influencers' compliance with advertising regulation and documenting the presence of children in their content between March and May 2024.

Based on the findings of these previous initiatives, we conducted an in-depth analysis of 20 Instagram and TikTok family influencer profiles, selected according to their

relevance for prior monitoring activities. The sample included family accounts, parent-managed child profiles, and creators who routinely feature minors.

For each item published by the selected influencers between 26 June and 10 July 2025 (a total of 1,334 posts, reels, stories, and live sessions), we coded variables such as platform (Instagram or TikTok); format (post, story, reel, live); content type (organic vs advertising); engagement metrics (likes, comments, shares, saves, views); presence of minors; and minors' estimated age. Coding was carried out by four analysts. Given the nature of the variables, primarily manifest and not requiring subjective interpretation, formal intercoder reliability testing was deemed unnecessary (Hayes & Krippendorff, 2007). Instead, agreement was ensured through structured training sessions, profile familiarization, a shared codebook, and periodic consensus meetings. To determine minors' ages, we produced a fact sheet compiling information retrieved from social media profiles, influencers' websites, and online articles. Cases of uncertainty or inconsistency were resolved by the project's principal investigator, who reviewed ambiguous content directly (Bauer & Gaskell, 2000).

For a subset of 112 posts and reels (for stories these data are not available) we conducted a deeper engagement analysis using data provided by the consultancy agency, enabling the calculation of Engagement Rate (ER) defined as interactions divided by the creator's follower count and multiplied by 100 (Cazorla Milla et al., 2020). This allowed us to evaluate both representational practices and commercial performance across organic and sponsored content.

Results

The analysis shows that Instagram dominates the content landscape, accounting for 91% of all items. The most frequently used format is stories (82.91%), followed by reels (14.09%), posts (2.85%), and live sessions (0.15%). The vast majority of content (84%) is organic, with sponsored content making up the remaining 16%.

Children's presence is substantial. They appear in 50.62% of organic items and 23.81% of sponsored ones. Many featured minors are aged 0–5 (78.5% of content containing children).

Engagement analysis yields counter-intuitive results. Organic content, on average, performs better than sponsored content, with ER values of 3.02% versus 1.17%, respectively. Within organic content, children's presence does not increase engagement; indeed, posts without minors achieve a higher ER (3.73%) than those with minors (2.71%). Conversely, in sponsored content, children's presence significantly boosts performance, with ER rising above 1.40% when minors are featured, compared to 0.59% in ads without children.

These findings demonstrate, first, the pervasive involvement of influencers' children in everyday storytelling. For family influencers, establishing an online persona often

requires sharing intimate aspects of family life, consistent with previous scholarship (Jorge et al., 2022; Van den Abeele et al., 2024). At the same time, minors appear in a substantial minority of sponsored content, thereby participating, directly or indirectly, in the execution of commercial agreements with brands. In regulatory contexts lacking specific protections for minors involved in influencer labor, such as Italy, this raises significant concerns. Without a coherent legal framework, children may be exposed to unregulated labor conditions and various forms of exploitation. These risks become even more salient when considering the growing presence of minors in the profiles of non-family influencers and the rise of the kidfluencer phenomenon (Clark & Jno-Charles, 2025; Van den Abeele et al., 2025a).

Future Work

Limitations include the relatively small number of analyzed profiles and the exclusive focus on family influencers. Nevertheless, the findings underscore the need for further research to better understand children's visibility in social media content produced by influencers. Future work will expand the analysis to prominent Italian influencers to examine whether and how they portray their children, and will also include kidfluencers, who are often adolescents using social media with explicit visibility goals. Additional research should explore family influencers' subjective perceptions of risks and opportunities through in-depth interviews, building on previous qualitative studies (Van den Abeele et al., 2024). Beyond academic inquiry, the evidence contributes to ongoing debates on the need for legal protections safeguarding minors' privacy and regulating their underage labor on social media, analogous to existing protections in acting and modeling.

References

- Abidin, C. (2017). #Familygoals: Family influencers, calibrated amateurism, and justifying young digital labor. *Social Media and Society*, 3(2), 205630511770719. <https://doi.org/10.1177/2056305117707191>
- Abidin, C., & Thompson, E. C. (2012). Buymylife.com: Cyber-femininities and commercial intimacy in blogshops. *Women's Studies International Forum*, 35(6), 467–477. <https://doi.org/10.1016/j.wsif.2012.10.005>
- Archer, C. (2019a). How influencer 'mumpreneur' bloggers and 'everyday' mums frame presenting their children online. *Media International Australia*, 170(1), 47–56. <https://doi.org/10.1177/1329878X19828365>
- Archer, C. (2019b). Social media influencers, post-feminism and neoliberalism: How mum bloggers' 'playbour' is reshaping public relations. *Public Relations Inquiry*, 8(2), 149–166. <https://doi.org/10.1177/2046147X19846530>

- Arriagada, A., & Ibáñez, F. (2020). “You Need At Least One Picture Daily, if Not, You’re Dead”: Content Creators and Platform Evolution in the Social Media Ecology. *Social Media and Society*, 6(3), 1–12. <https://doi.org/10.1177/2056305120944624>
- Bauer, M. W., & Gaskell, G. (2000). *Qualitative Researching with Text, Image and Sound: A Practical Handbook for Social Research*. SAGE.
- Baxter, K., & Czarnecka, B. (2025). Sharing images of children on social media: British motherhood influencers and the privacy paradox. *PLoS ONE*, 20(1). <https://doi.org/10.1371/journal.pone.0314472>
- BBC. (2020, ottobre 7). France passes new law to protect child influencers. BBC. <https://www.bbc.com/news/world-europe-54447491>
- Cazorla Milla, A., Mataruna-Dos-Santos, L. J., & Muñoz Helú, H. (2020). Content Strategy and Fan Engagement in Social Media The Case of PyeongChang Winter Olympic And Paralympic Games. *Annals of Applied Sport Science*, 8(3), 0–0. <https://doi.org/10.29252/aassjournal.905>
- Clark, D. R., & Jno-Charles, A. B. (2025). The Child Labor in Social Media: Kidfluencers, Ethics of Care, and Exploitation. *Journal of Business Ethics*, 201(1), 35–62. <https://doi.org/10.1007/s10551-025-05953-7>
- De Veirman, M., Cauberghe, V., & Hudders, L. (2017). Marketing through instagram influencers: The impact of number of followers and product divergence on brand attitude. *International Journal of Advertising*, 36(5), 798–828. <https://doi.org/10.1080/02650487.2017.1348035>
- Duffy, E. B., & Sawey, M. (2024). The politics of vulnerability in the influencer economy. *European Journal of Cultural Studies*, 27(3), 352–370. <https://doi.org/10.1177/13675494231212346>
- Hayes, A. F., & Krippendorff, K. (2007). Answering the Call for a Standard Reliability Measure for Coding Data. *Communication Methods and Measures*, 1(1), 77–89. <https://doi.org/10.1080/19312450709336664>
- IAP & ALMED. (2024). Monitoraggio Trasparenza e Influencer Marketing: Beauty, fashion, family e finance. IAP - ALMED. <https://www.iap.it/wp-content/uploads/2025/04/Monitoraggio-Trasparenza-e-Influencer-Marketing-Report-2024-IAP-Almed.pdf>
- Jorge, A., Marôpo, L., Coelho, A. M., & Novello, L. (2022). Mummy influencers and professional sharenting. *European Journal of Cultural Studies*, 25(1), 166–182. <https://doi.org/10.1177/13675494211004593>
- Terre Des Hommes Germany e.V. (2024). “Child labour? In Germany?” – Child labour report 2024.

- Van Den Abeele, E., Hudders, L., & Vanwesenbeeck, I. (2024). The (un)necessity of child portrayal in momfluencer content: Exploring mothers' perspectives on influencer sharenting through in-depth interviews. *Information, Communication & Society*, 1–21. <https://doi.org/10.1080/1369118X.2024.2352625>
- Van den Abeele, E., Hudders, L., & Vanwesenbeeck, I. (2025a). Managing authenticity in a kidfluencers' world: A qualitative study with kidfluencers and their parents. *New Media & Society*, 27(6), 3240–3263. <https://doi.org/10.1177/14614448231222558>
- Van den Abeele, E., Hudders, L., & Vanwesenbeeck, I. (2025b). Tell me why: The impact of mindful sharenting explanations by momfluencers: An experimental study with mothers. *New Media & Society*, 14614448251314869. <https://doi.org/10.1177/14614448251314869>

6B.2 Performing ‘Mountains and Spices’ on Dianping: Cultural Politics and Platform Marketing of “Yunguichuan Bistros” in Urban China

Hao Xie (University of Warwick)

Introduction

In the post-COVID era, China’s urban culinary landscape has been reshaped by the nationwide rise of Yunnan–Guizhou cuisines. By August 2025, more than 42,000 Yunnan–Guizhou-style restaurants were operating across the country, marking the fastest growth among full-service dining sectors (Xinhua, 2025). Within this boom, a new gastronomic genre, the Yunguichuan (Yunnan–Guizhou–Sichuan) bistro, has rapidly proliferated across eastern Chinese cities. These establishments claim to blend sour-spicy flavours, “wild” mountain ingredients, and ethnic motifs from China’s southwest into curated lifestyle experiences that fuse regional novelty with urban sophistication. Promoted as “authentic yet stylish” and “rustic yet refined,” Yunguichuan bistros appeal to a generation for whom dining is increasingly tied to pleasure, identity performance, and social display.

While this trend may appear to signal cultural openness and economic vitality, expanding the visibility of regional cuisines and facilitating interregional exchange, it also carries a more intricate cultural politics. As Yunguichuan bistros repackage southwestern cuisines for metropolitan consumption, they transform difference into a consumable spectacle calibrated to the tastes and aspirations of an expanding urban middle class. Cultural symbols long associated with minority nationalities, peripheral geographies, and imaginaries of socio-economic “backwardness” are recontextualised into a fashionable aesthetic of leisure and refinement (Cheung, 2012; Gros, 2011; Oakes & Zhenting, 2022; Tapp, 2010). Such rebranding both depends upon and obscures the unequal social structures through which the “periphery” becomes desirable to the “centre.”

Inspired by Appadurai’s (1981) notion of gastro-politics, this study examines the platform marketing discourse of Yunguichuan bistros on Dazhong Dianping, a leading consumer review and lifestyle platform in China. It conceptualises Dianping as a contested cultural field where regional identity, class aspiration, and cultural representation intersect within China’s depoliticised multiculturalism. I argue that although the bistro genre celebrates diversity as hybridised lifestyle consumption, it simultaneously reproduces asymmetries between metropolis and periphery, Han-centric imaginaries and minority difference, and urban consumers and rural communities. Extending gastro-politics to a platform context, the paper shows how

Dianping's platformised visibility, such as branding rhetorics and visual norms, helps normalise “authentic” and “wild” as a repeatable lifestyle aesthetic, turning regional/ethnic difference into marketable content while sustaining centre–periphery and class hierarchies in urban China.

Objectives/Research Questions

This study draws on Appadurai's (1981) notion of gastro-politics, which conceptualises the culinary practices as “conflict or competition over specific cultural or economic resources as it emerges in social transactions around food” (p.495). In other words, Appadurai's (1981) conception directs attention to how acts of commensality, exchange, and consumption reproduce, negotiate, or subvert existing hierarchies of class, gender, and status. Food, in this view, operates as a medium through which broader relations of power and inequality are materially and symbolically enacted. Adapting this idea to the platformised food branding in China, this study asks:

- 1) Under the platform infrastructures, how do Yunguichuan bistros construct and circulate their food cultures on Dazhong Dianping through platform-facing promotional discourse and visual branding (e.g., shop-page narratives, images, and recommended-dish scripts)?
- 2) How do platform users' reviews and engagement practices (taste talk, photo norms, recommendation scripts) decode the branded food cultures?
- 3) How do we understand the power structures (centre-periphery, classed distinction, ethnic/regional imaginaries) that are reproduced or reconfigured through this platformised taste-making and promotional culture?

Methods

This study employs a qualitative research design to examine how Yunguichuan bistros are marketed and made meaningful on Dazhong Dianping. Focusing on Shanghai as a major site of platformised consumption, the dataset is constructed in two stages. First, I compile a corpus of Yunguichuan-themed restaurants by searching Dianping with the keyword ‘Yunguichuan bistro’ and recording results on a fixed collection date (i.e., 12 January 2026). Based on the platform visibility at the time of retrieval, the study identifies 34 restaurants explicitly foregrounding the theme of ‘Yunguichuan bistros’ in their names.

In the first stage, I examine how the sampled restaurants present themselves in the digital marketplace. I collected branding and promotional materials on Dianping. Slogans, featured menus, official introductions and relevant posts were gathered and analysed. This aspect of data collection focuses on the discursive construction of taste, experience and identity: how the restaurants narrate “Yunguichuan” culture as both traditional and fashionable, thus transforming regional cuisines into lifestyle commodities for middle-class consumption. In the second stage, I collected all

consumer reviews on the platform posted in the latest week (5 January -11 January 2026) before the data collection date. This includes a total of 754 comments. Their narratives reveal how everyday ‘taste talk’ on the platform decodes the branded food cultures, where authenticity, taste, style, and experience are discussed.

All textual and visual data were analysed according to the principles of Critical Discourse Analysis (CDA) (Fairclough, 2013; Van Dijk, 2005), which conceptualises discourse as a form of social practice. The analysis proceeded in two iterative stages: (1) identifying dominant discursive patterns across the online materials; and (2) interpreting how these discourses reflect, contest, and reproduce power relations involving ethnicity, region and class. Patterns were identified around recurring linguistic framings, visual motifs, and narrative tropes related to the analysed power dynamics.

Ethical handling follows established guidance for research using public online content. Although Dianping reviews are publicly accessible, they may still be traceable to individuals. I therefore do not record or report usernames, profile details, or direct identifiers, and I avoid reproducing screenshots that could enable re-identification. Where review excerpts are used, they are lightly paraphrased or translated and stripped of unique phrasing, while preserving analytic meaning.

Results

By synthesising platform marketing materials and user discourse, this study identifies cultural operations through which Yunguichuan bistros transform southwestern cuisines and identities into consumable urban imaginaries. Three recurring processes structure this transformation: homogenising regional food cultures, exoticising minority ethnic cultures, and romanticising mountainous rural life. These processes expand the visibility of “local” culture while reproducing hierarchies of region, ethnicity, and class in contemporary China.

First, Yunguichuan bistros brand themselves as offering fresh, alternative tastes from an “underexplored” southwest. Yet despite the heterogeneity of Yunnan, Guizhou, and Sichuan culinary traditions (Li & Yang, 2023), they repeatedly promote a singular, legible notion of “Yunguichuan,” converting diverse foodways into a standardised “Southwestern flavour.” Regional and ethnic specificities are flattened into a repertoire calibrated to urban middle-class sensibilities - mildly “exotic” sour-spicy profiles, selectively highlighted techniques, and refined presentations. The result is less an authentic composite than an imagined unity produced for metropolitan Han consumption, echoing “unity in diversity” logics that aestheticise difference while centring the urban mainstream (Tan, 2004; Yan & Vickers, 2023).

Second, exoticisation intensifies this domestication. Menus label dishes as “Dai-style,” “Hani-style,” or “Yunnan-style” to cue an alluring elsewhere, even when items are reworked or invented. Featured interior images deploy stylised ethnic motifs

(embroidered textiles, patterned fabrics, warm lighting) to produce “ethnic authenticity” in a polished grammar. Dianping branding also extends this through images of landscapes, minority attire, and “raw ingredients,” framing the southwest as timeless purity and essence. Such decontextualised fragments convert communities into consumable metaphors (Gorfinkel, 2018): minority figures become emblems of authenticity, stripped of voice, supporting urban consumers’ pursuit of novelty and cultivated taste.

Third, the marketing discourse romanticises mountain and rural life as therapeutic escape. “Natural” interiors (stone, plants, wood) and online rhetoric, e.g., “hidden realms” (mijing), “deep in the mountains” (dashanshenchu), “oil paintings” (youhua), recast uneven development and marginalisation as picturesque restoration.

User discourse shows differentiated decoding (Hall, 2003). Many reviews reproduce “fashionable authenticity” via recommended-dish lists, “must-order” scripts, and photo-centric praise. Negotiated readings hinge on price–value judgements and “stylised but not ‘local enough’” evaluations, while oppositional readings critique hype, inauthentic fusion, or ethnic décor as cliché. Overall, platformed imaginaries are both reproduced and contested, while hierarchy-making through platformised taste remains largely stabilised.

Future Work

Future work may broaden the analysis beyond Shanghai and Dazhong Dianping to test how “Yunguichuan” or similar regional-ethnic food imaginaries travel across platforms, cities and countries. Future work may also incorporate longitudinal sampling to track how platform tags, ranking cues, and visual norms change over time, and how restaurants strategically adapt to these shifts. Finally, targeted interviews with restaurant marketers and consumers will clarify production rationales and interpretive practices that remain opaque in platform text analysis.

References

- Appadurai, A. (1981). gastro-politics in Hindu South Asia. *American Ethnologist*, 8(3), 494-511.
- Cheung, S.-w. (2012). Appropriating otherness and the contention of Miao identity in Southwest China. *The Asia Pacific Journal of Anthropology*, 13(2), 142-169.
- Fairclough, N. (2013). Critical discourse analysis. In *The Routledge Handbook of Discourse Analysis* (pp. 35-46). Routledge.
- Gorfinkel, L. (2018). *Chinese television and national identity construction: the cultural politics and music entertainment programmes*. Routledge.

- Gros, S. (2011). Economic marginalization and social identity among the Drung people of northwest Yunnan. *Moving Mountains: Highland Livelihood and Ethnicity in China, Vietnam, and Laos*, 28-49.
- Hall, S. (2003). Encoding/decoding. *Television: critical concepts in media and cultural studies*, 4, 43-53.
- Li, X., & Yang, H. (2023). Food as a Medium: Ethnic Interactions and Integration Across Yunnan's Dining Tables. 2022 4th International Conference on Literature, Art and Human Development (ICLAHD 2022),
- Oakes, T., & Zhenting, Z. (2022). Remoteness and connectivity: The variegated geographies of the Yunnan-Guizhou Plateau. In *Routledge Handbook of Highland Asia* (pp. 418-429). Routledge.
- Tan, M. G. (2004). Unity in Diversity: Ethnic Chinese and Nation-Building. *Ethnic Relations and Nation-Building in Southeast Asia: The Case of the Ethnic Chinese*, 20.
- Tapp, N. (2010). Yunnan: Ethnicity and economies—markets and mobility. *The Asia Pacific Journal of Anthropology*, 11(2), 97-110.
- Van Dijk, T. A. (2005). Critical discourse analysis. *The handbook of discourse analysis*, 349-371.
- Xinhua. (2025). 《云贵菜发展报告2025》：抖音话题破百亿，五大趋势助力品牌“掘金”增量机会.
Retrieved from:
<http://www.news.cn/food/20250828/6f682743ee8649ca9829e7c319679969/c.html> [Accessed on 10 October 2025]
- Yan, F., & Vickers, E. (2023). Balancing unity and diversity? Shifting state policies and the curricular portrayal of China's minority nationalities. *Comparative Education*, 1-20.

6B.3 Virtual vs. Human Faces: A comparative study of influencer opinions and their effects on user perceptions and pro-environmental behavioral intentions

Son Tung Duong (Deakin University), Asma Husna, Nhung Nguyen (University of South Alabama)

Introduction

A new kind of celebrity is taking over our screens—one that never sleeps, never ages, and never makes a mistake. From Lil Miquela (@lilmiquela)’s polished activism to Rozy (@rozy.gram)’s visuals, activism-themed posts, virtual influencers (VI) are reshaping the boundaries of online persuasion. They look real, speak with emotion, and advocate for causes from fashion to climate change. As this digital phenomenon grows, the question emerges: “Can virtual faces inspire real-world change? Can they motivate people to act just as or even more effectively sustainable than their human counterparts?” Human influencers (HI) have been valued for their authenticity and emotional connection. They evoke trust through lived experience, perceived sincerity, and relatable storytelling. In contrast, VI are meticulously designed and scripted for precision, consistency, and appeal. Yet their lack of real-world identity raises doubts about credibility, truthworthiness, and genuine intent. This tension between authenticity and craftsmanship lies at the heart of modern digital persuasion, especially in promoting pro-environmental behaviors, where moral trust and empathy play central roles. Previous literature content that message processing depends on how individuals evaluate persuasive information through either central or peripheral routes. Emotional appeals often evoke empathy and moral concern, while informational appeals depend on logic and factual evidence. Perceptions such as perceived truthworthiness, perceived attractiveness, inferences of manipulative intent, and environmental consciousnesses may shape how audiences interpret and act upon influencer messages. When messages are seen as truthful and knowledgeable, persuasive impact increases; when motives seem manipulative, intentions weaken. These perceptual factors may moderate how source type and message framing influence pro-environmental behavioral intentions.

Objectives/Research Questions

Drawing from influencer communication frameworks and the Elaboration Likelihood Model, this study examines how emotional and informational messages from HI and VI influence audiences’ pro-environmental behavioral intentions. Theoretically, it advances understanding of persuasion in the context of AI-mediated communication, extending authenticity, credibility, and emotion to VIs. Practically, the findings can

inform sustainability campaigns and brand strategies by identifying effective message types and influencer characteristics that foster trust and motivation toward environmental action.

Green Behavioral Intention

A increasing number of HIs have been found to promote environmentally friendly products and brands that serve the well-being of humanity. They often share posts that support sustainable lifestyles and green products and encourage their followers to protect the environment , thereby shaping the audience's attitudes regarding environmentally friendly products through posts on social media (Freberg et al., 2011; Taylor, 2021). In the context of VIs, Jiang et al. (2024) found that, when the involvement of green messages is low, different types of VIs, including human-like virtual influencers (HVI) and anime-like virtual influencers (AVI) can enhance persuasive power through pride or gratitude appeals. However, it is unclear about the effectiveness of VI's pro-environmental content versus HI's on users behavioral intentions.

RQ1: Does viewing content from a VI lead to stronger stronger pro-environmental behavioral intentions compared to a HI?

Elaboration Likelihood Model (ELM)

Developed by Petty et al. (1983), Elaborate Likelihood Model (ELM) provides a framework to understand the basic psychological processes of recipients in persuasive communications by explaining the roles of central and peripheral routes in persuasion. Studies suggest that the central route has a superior role to the peripheral route in persuasive communication (Allison et al., 2017; Lee & Koo, 2016). In contrast, some literature maintains that the central route is inferior to the peripheral route in persuading the audience (Li et al., 2017; Syrdal et al., 2023). It raises the question of whether either HI's or VI's message (both emotional and informational) more effectively influences users' perceptions and behaviours. Based on the literature showing emotional is more effective than informational messages in shaping users, this study seeks to answer the below question:

RQ2: How do emotional and informational message appeals from HI and VI differentially affect users' green behavioral intentions?

Perceived Influencer Characteristics

Influencers' characteristics—particularly attractiveness and trustworthiness—are key components of the source credibility model (Ohanian, 1990).

RQ3: How do perceived truthworthiness and perceived attractiveness moderate the effects of influencer type (HI vs. VI) and message type on users' pro-environmental behavioral intentions?

Methods

This study consists of a two-by-two factorial experiment to examine the effects of HI vs. VI and emotional vs. informational features (independent variables) on Instagram on the audience's green behavior intentions (dependent variables). The study also explore the effects perceived influencers (i.e., perceived truthworthiness, perceived attractiveness, inferences of manipulative intent, and environmental concerns) on the relationship between independent and independent people. Participants were randomly assigned to view one of two influencer conditions (HI vs. VI) and one of two message conditions (emotional vs. informational). Each stimulus consisted of a short Instagram-style photo and text post created for this study. The VI condition used an AI-generated influencer with realistic features, while the human condition featured a comparable real influencer matched in age, gender, and style. The experiment was hosted on Qualtrics. Participants accessed the study using their own devices; no specialized VR headset was required.

Results

Preliminary Results

Two-way analysis of variance (ANOVA) examined the effects of influencer type (human vs. virtual) and message framing (emotional vs. informational) on participants' attitudes and behavioral intentions toward green behaviors.

Main and Interaction Effects

Results revealed a significant interaction between influencer type and message framing. Participants who viewed emotional messages from human influencers (HIs) reported the most favorable attitudes and strongest intentions to engage in green behaviors, whereas informational messages from virtual influencers (VIs) produced the lowest scores.

Perceived Influencer Characteristics

Human influencers were perceived as significantly more trustworthy and expert than virtual influencers. In contrast, virtual influencers were rated higher in perceived attractiveness, being described as more aesthetically appealing and technologically innovative.

Future Work

Moving forward, we will continue to collect data and apply these preliminary findings to refine the study instruments and proceed with the full-scale data collection.

References

- Allison, T. H., Davis, B. C., Webb, J. W., & Short, J. C. (2017). Persuasion in crowdfunding: An elaboration likelihood model of crowdfunding performance. *Journal of Business Venturing*, 32(6), 707–725.
<https://doi.org/10.1016/j.jbusvent.2017.09.002>
- Freberg, K., Graham, K., McGaughey, K., & Freberg, L. A. (2011). Who are the social media influencers? A study of public perceptions of personality. *Public Relations Review*, 37(1), 90–92. <https://doi.org/10.1016/j.pubrev.2010.11.001>
- Kahle, L. R., & Homer, P. M. (1985). Physical Attractiveness of the Celebrity Endorser: A Social Adaptation Perspective. *Journal of Consumer Research*, 11(4), 954.
<https://doi.org/10.1086/209029>
- Lee, Y., & Koo, J. (2016). Can a Celebrity Serve as an Issue-Relevant Argument in the Elaboration Likelihood Model? *Psychology and Marketing*, 33(3), 195–208.
<https://doi.org/10.1002/mar.20865>
- Li, J. (Jason), Chen, X.-P., Kotha, S., & Fisher, G. (2017). Catching fire and spreading it: A glimpse into displayed entrepreneurial passion in crowdfunding campaigns. *Journal of Applied Psychology*, 102(7), 1075–1090.
<https://doi.org/10.1037/apl0000217>
- Ohanian, R. (1990). Construction and Validation of a Scale to Measure Celebrity Endorsers' Perceived Expertise, Trustworthiness, and Attractiveness. *Journal of Advertising*, 19(3), 39–52. <https://doi.org/10.1080/00913367.1990.10673191>
- Rungruangjit, W. (2022). What drives Taobao live streaming commerce? The role of parasocial relationships, congruence and source credibility in Chinese consumers' purchase intentions. *Heliyon*, 8(6), e09676.
<https://doi.org/10.1016/j.heliyon.2022.e09676>
- Syrdal, H. A., Myers, S., Sen, S., Woodroof, P. J., & McDowell, W. C. (2023). Influencer marketing and the growth of affiliates: The effects of language features on engagement behavior. *Journal of Business Research*, 163, 113875.
<https://doi.org/10.1016/j.jbusres.2023.113875>
- Taylor, J.-A. (2021). Social media and the rise of influencers. *BDJ Team*, 8(7), 12–13.
<https://doi.org/10.1038/s41407-021-0674-0>

6B.4 Enterprising Platform-Dependent Creators and the Rise of Creator-Dependent Labour in India

Srikanth Nayaka (Indian Institute of Management [IIM] Indore)

Introduction

Social media platforms such as Facebook, YouTube, Instagram, and other similar such platforms have disrupted media and entertainment industries. First, these platforms and their various socio-technical affordances have given rise to the platform-dependent commercial content creators who monetize their media content on these platforms (Poell et al., 2021). Second, independent native creators who are part of this parallel platform ecology, dubbed as social media entertainment (SME) (Cunningham & Craig, 2019), acquired micro-celebrity status (Senft, 2008) and identified themselves as social media influencers (Abidin, 2015). While this commercial creator culture (Cunningham & Craig, 2021) has proliferated across the world, including in China (Craig et al., 2021) where social media content creators negotiate precarious platform labour through various strategies (Duffy, 2017). In Chinese context, scholars have examined how platformization and digitalization of economy given rise to “unlikely creative class” (Lin & de Kloet, 2019) and digital “entrepreneurs” (Zhang, 2023). Similarly, in India also, digital and platform economy brought transformative changes where a range of people reinventing themselves “as an entrepreneur” (Athique & Parthasarathi, 2020; Zhang, 2023). While the gig economy and platform work has received critical academic attention (Komarraju et al., 2021; Raval & Pal, 2019), the practices of social media entertainment-based platform entrepreneurs are under examined in India. However, there have been many independent and industry studies which recognized the scale and scope of social media entertainment industry in India (Economics, 2022). Online content creators who have achieved massive visibility transformed into platform entrepreneurs who run their digital enterprises with freelance and other regular full-time employed creative and technical team. Some of the creator-dependent labour includes video editors, thumbnails designers, artists, cameraman, light men, supporting actors, production managers and others. Film scholar Ranjani Mazumdar has brought attention to what she termed “invisible work” in the Indian media industries and asked scholars to consider “stunt artists, body doubles, dancers, junior artists, makeup artists, lower-end technicians, and spot boys” in theorizing media labour (Mazumdar, 2015, p.26). To this end, this paper examines how hyper visible and invisible creators negotiate labour in digital media industries in India.

Objectives/Research Questions

- 1) To identify a select group of social media content creators who have acquired massive platform visibility, digital popularity, and transformed into platform entrepreneurs
- 2) To map various strategies and management practices employed by platform entrepreneurs
- 3) To understand various issues, challenges, and opportunities in platform entrepreneurship
- 4) To highlight the rising supporting creator labour or creator-dependent labour in platform-dependent entrepreneurship

Methods

This paper is part of comparative analysis of rural and urban platform-dependent creators. I draw on fieldwork data conducted in Lambadipally, a small village in the state of Telangana in India between August 2022 and October 2022 as part of my doctoral work. I conducted in-depth interviews with eight creators lasting between 30 and 90 minutes. Except for one woman creator, all my interviewees were predominantly male (7). All these creators have amassed massive online popularity and online visibility through short scripted entertainment YouTube videos. Most importantly, their do-it-yourself (DIY) media creations have been spectacularly formalized and professionalized. They all now approach digital productions as professional media work (Deuze, 2007) and few of them diversified across multiple media platforms. As part of my ethnographic fieldwork, I have engaged in participant observations at their shooting locations and production office. All these creators are associated with My Village Show, a Telugu YouTube channel with over 3 million subscribers, founded by Srikanth Sriram, an engineering graduate from Lambadipally village. Srikanth Sriram transformed My Village Show into 'Village Show Private Limited,' a government-registered private company with fifteen regular employees. He collaborated with few other villagers and established a Human Resource Development Centre (HRDC), a production studio cum office of My Village Show. Village video-makers produce a wide variety of audio-visual content such vlogs, cooking videos, short videos, full-length entertainment videos, short films, interviews, and web series. All creators involve in different roles in their productions such as directors, script-writers, actors, cameraman, production assistant, editors, and publicity managers.

Hailing from different socio-economic and educational backgrounds, rural digital creators as a 'commercializing and professionalizing native social media users' negotiate global media platforms as platform-dependent and platform-independent cultural producers (Cunningham & Craig, 2021: 01). As Poell et al. (2021) point out, 'platform-dependent producers rely on platforms in the creation, distribution, marketing, and monetization of content and services. By contrast, platform-independent producers pursue these activities separately from platforms' (Poell et al.,

2021:13). Rural content creators can be positioned ‘somewhere in between platform-dependence and independence’ (Poell et al., 2021, p.13) as they sustain their income through different activities such as independent business, promotional videos, and pick up any relevant roles in legacy media industries. Following Gina Neff’s work on high-tech workers in New York City’s Silicon Alley (Neff, 2012), I collected data through deep hanging out (Geertz, 1998). I visited their shooting locations, observed their scripting process, sat with them in editing studios, and discussed about their media worlds.

Results

In the context of India, platformization has brought massive changes to creative labor and creative cultures, reshaping how individuals engage with digital media, monetize their content, and navigate new career pathways in the creator economy. This paper situates emerging digital creative work within the broader framework of ‘cultures of work’ in India, examining them through socio-economic, regional, and technological perspectives to understand their impact on labor, livelihood, and digital entrepreneurship (Upadhyay, 2019). From fitness trainers to IT professionals and call center workers, socio-economic mobility in India has traditionally been tied to a limited set of professions, often shaped by educational qualifications, urban opportunities, and industry demands. People from small towns are reinventing themselves as fitness trainers and call centre agents, learning English, and aspiring to middle-class lifestyles (Baas, 2020; Mankekar & Gupta, 2016). Yet, this study highlights a countertrend emerging in India, where digitally networked communities are cultivating aspirational identities by harnessing creative expression and online networks to challenge entrenched norms.

Future Work

In this paper, the analysis is grounded in a rural case study, while interviews with other creator-dependent and platform entrepreneurs in urban settings are currently underway. As creators experience precarious creativity on social media entertainment and the influencer industry, they explore alternative pathways including film and television industries. During interactions with rural creators, I observed that there are many creative labourers who are dependent on the digital folk industry. The rapidly emerging digital folk music is less explored in the Indian regional context and the future studies should examine how the platformization of cultural production transformed the production, distribution, and consumption of folk genres. On a broader level, I argue that creator labour practices should be given much importance like IT labour, call centre work, the platform and gig economy labour practices.

References

- Abidin, C. (2015). Communicative ♥ Intimacies: Influencers and Perceived Interconnectedness. *Ada: A Journal on Gender and New Media and Technology*, 8, 1–15. <https://doi.org/doi:10.7264/N3MW2FFG>
- Athique, A., & Parthasarathi, V. (Eds.). (2020). *Platform Capitalism in India*. Palgrave Macmillan.
- Baas, M. (2020). *Muscular India: Masculinity, Mobility & the New Middle Class*. Context.
- Craig, D., Lin, J., & Cunningham, S. (2021). *Wanghong as Social Media Entertainment in China*. Palgrave Macmillan.
- Cunningham, S., & Craig, D. (2019). *Social Media Entertainment: The New Intersection of Hollywood and Silicon Valley*. New York University Press.
- Cunningham, S., & Craig, D. (Eds.). (2021). *Creator Culture: An Introduction to Global Social Media Entertainment*. New York University Press.
- Deuze, M. (2007). *Media Work*. Polity Press.
- Duffy, B. E. (2017). *(Not) Getting Paid to Do What You Love: Gender, Social Media, and Aspirational Work*. Yale University Press.
- Economics, O. (2022). A Platform for Indian Opportunity: Assessing the Economic, Societal and Cultural Impact of YouTube in India. <https://kstatic.googleusercontent.com/files/45fedbf71b2dd6e7a59b4990a731b6b8067a434accfdc9c6c15df18177e63b4a2f454fe1a165ca47b78412325386c1a101d950355e4811d5e09cb363cb9f3d77>
- Geertz, C. (1998). Deep hanging out. *The New York Review of Books*, 45(16), 69–72.
- Komarraju, S. A., Arora, P., & Raman, U. (2021). Agency and servitude in platform labour: a feminist analysis of blended cultures. *Media, Culture and Society*, 15, 3–6. <https://doi.org/10.1177/01634437211029890>
- Lin, J., & de Kloet, J. (2019). Platformization of the Unlikely Creative Class: Kuaishou and Chinese Digital Cultural Production. *Social Media and Society*, 5(4). <https://doi.org/10.1177/2056305119883430>
- Mankekar, P., & Gupta, A. (2016). Intimate encounters: Affective labor in call centers. *Positions*, 24(1), 17–43. <https://doi.org/10.1215/10679847-3320029>
- Mazumdar, R. (2015). “Invisible Work” in the Indian Media Industries. *Media Industries Journal*, 1(3). <https://doi.org/10.3998/mij.15031809.0001.306>

- Neff, G. (2012). *Venture Labor: Work and the Burden of Risk in Innovative Industries*. MIT Press.
- Poell, T., Nieborg, D. B., & Duffy, B. E. (2022). *Platforms and Cultural Production*. Polity Press.
- Raval, N., & Pal, J. (2019). Making a “Pro”: ‘Professionalism’ after Platforms in Beauty-work. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW). <https://doi.org/10.1145/3359277>
- Senft, T. M. (2008). *Camgirls: Celebrity and Community in the Age of Social Networks*. Peter Lang.
- Upadhy, C. (2019). Cultures of Work in India’s ‘New Economy.’ In S. Srivastava, Y. Arif, & J. Abraham (Eds.), *Critical Themes in Indian Sociology* (pp. 151–163). Sage.
- Zhang, L. (2023). *The Labor of Reinvention: Entrepreneurship in the New Chinese Digital Economy*. Columbia University Press.

Session 6C - Cyberbullying, Trolling and Antisocial Behavior (1)

Day 2 · Tues 14 July · 9:30–11:00, Seminar Suite B, ARC

6C.1 Using YouTube Data to Measure Online Harassment in Canadian Politics

Inessa De Angelis (University of Toronto)

Introduction

“This woman is trained in witchcraft, she's just spewing a word salad” reads a comment left beneath a YouTube video posted by then Canadian Member of Parliament (MP) Chrystia Freeland. Like many comments received by politicians on social media, this remark is not an isolated expression of disagreement, but an example of gendered delegitimization that is normalized and amplified by social media platforms (Sobieraj, 2020). For women politicians in Canada and other democratic nations, navigating online harassment increasingly shapes routine aspects of political life, including communication and engagement with the public (Southern & Harmer, 2021).

Research on Violence Against Women in Politics (VAWIP) shows that women politicians are disproportionately targeted with gendered and personalized harassment on social media (Jane, 2017; Wagner, 2020; Rheault et al., 2019). Women politicians with intersecting identities are particularly impacted by various forms of harassment, including threats, gendered, and sexualized violence (Herrick & Thomas, 2024). Social media can help facilitate this harassment through platform affordances such as anonymity and limited or non-existent content moderation policies (Sobieraj, 2020).

Despite a growing body of gendered online harassment work, existing research largely focuses on Twitter (X). While YouTube plays an important role in political communication, it has received less scholarly attention in Canadian political communication and VAWIP literatures (Arthurs et al., 2018; Esposito & Zollo, 2021; Mattan et al., 2024). YouTube's looser network structure, comment visibility, and engagement-driven recommendation systems can contribute to and magnify harassment, often allowing abusive comments to circulate long after they are posted (Mamié et al., 2021).

To better understand the nature and severity of harassment on YouTube, this paper analyzes one year of public comments left on videos posted by federal Canadian MPs.

Using a supervised multiclass machine learning model to predict the severity of harassment in each comment, this paper focuses on the range of forms and severity of harassing comments, especially more subtle forms like “questioning authority.” Following Rossini (2022), this paper recognizes that shades of civil and uncivil speech simultaneously exist online and must be understood along a spectrum. By moving beyond binary categorizations of harassment, this paper contributes to a more nuanced understanding of the forms and severity of harassment and its implications for democratic representation.

Objectives/Research Questions

The goal of this paper is to use supervised machine learning to understand how the harassment faced by federal Canadian MPs on YouTube differs based on their gender, racialized or Indigenous background, and political party affiliation. Accordingly, my research questions are:

1. Are women MPs subjected to more severe forms of harassment than men MPs on YouTube?
2. How do party and background shape the severity of harassment experienced by MPs?
3. What is the distribution of severity of harassment in comments on YouTube videos posted by MPs?

I hypothesize that the harassment will be systematically gendered, with women MPs receiving higher proportions of harassing comments focused on their gender, appearance, and qualifications categorized as questioning authority and name-calling/gender insults (Jane, 2017). I also expect that racialized MPs and MPs from centre and centre-left leaning political parties will receive more severe harassment overall (Herrick & Thomas, 2024; Pedersen et al., 2025). Furthermore, I hypothesize that more subtle forms of harassment, such as questioning authority and name-calling/gender insults will be the most prevalent, while some more severe forms like credible threats and hate speech will also be detected (De Angelis, 2024).

Methods

To understand how the harassing YouTube comments uniquely impact MPs on the basis of their demographic characteristics, I collected one year worth of public comments left on YouTube videos posted by federal Canadian MPs. Data were collected by writing custom R and Python scripts that queried the YouTube API. I collected data for both women and men MPs to account for “all possible victims” (Bardall et al., 2020, p. 928). After data collection and cleaning processes, I had 672,758 comments left by the public on YouTube videos posted by 116 active MPs between January 1, 2024, and December 31, 2024.

I trained a supervised multiclass machine learning classification model in Python to predict the severity of harassment in each comment, following the seven-point scale developed by De Angelis (2024). The seven categories are: positive (1), neutral (2), questioning authority (3), namecalling/ gender insults (4), vicious language (5), credible threats (6), and hate speech (7). A specific contribution of this seven-point scale is the “questioning authority” category, which emphasizes how much of the harassment faced by women politicians is more nuanced and subtle, focusing on their physical appearance and assumed qualifications, but still undermines their equal participation in politics (Krook, 2020). The categorization of each comment is mutually exclusive and must belong to exactly one of these seven categories. The model was trained on a manually coded subset of comments from both Twitter and YouTube, with cross-validation used to optimize performance.

Following classification, I added MP-level demographic data to my dataset, focusing on gender, background, and political party affiliation. To predict whether a certain combination of MP-level characteristics lead to more severe forms of harassment (such as credible threats or hate speech), I estimate a multinomial logistic regression model in R (R Core Team, 2025). I used multinomial logistic regression because the severity of harassment scores are categorical (1 to 7), allowing me to model differences across multiple levels of severity. Multinomial logistic regression is suitable for this analysis because the dependent variable (severity of harassment) does not need to be collapsed into a binary variable, which otherwise would impact our ability to understand the relationship between MPs’ personal characteristics and the harassment they receive.

Finally, I qualitatively analyze select comments to contextualize the quantitative findings and to illustrate the intersectional nature of harassing language.

Results

Preliminary results from the supervised machine learning classifications reveal that harassment on YouTube is widespread and often appears in more subtle forms. 48 percent of comments were classified as questioning authority, suggesting that much of the harassment directed toward MPs is through subtle microaggressions and undermining comments. The second most identified category was name-calling/gender insults (22 percent), followed by vicious language (14 percent). The more severe categories of credible threats and hate speech were also detected, accounting for just under 6 percent of all comments. Likewise, smaller numbers of positive and neutral categories were detected, totaling around 10 percent of the comments. 90 percent of comments were classified as containing some form of harassment, ranging from questioning authority to hate speech.

Results from my model suggest systematic differences in the predicted severity of harassment based on MPs' gender, background, and party. When holding other characteristics constant, women MPs received higher predicted levels of harassment than men MPs. Differences by background were more pronounced: racialized MPs and especially Indigenous MPs were predicted to receive more severe harassment than white MPs, emphasizing the intersectional nature of harassment (Sobieraj, 2020; Herrick & Thomas, 2024).

Political affiliation also shapes the severity of harassment. Preliminary estimates suggest that Conservative MPs receive lower predicted levels of harassment than Liberal MPs. There were no meaningful differences predicted between Liberal and NDP MPs. These findings align with Pedersen et al. (2025) who found that centre and centre-left leaning politicians receive more online harassment and perceive the harassment they receive as being more severe.

Future Work

This study examines public comments left on videos posted by Canadian MPs to assess the forms and severity of online harassment. Preliminary results indicate that harassment is widespread, with 90 percent of comments containing some harassment. Subtle forms of harassment, such as questioning authority and name-calling/gender insults were the most prevalent, with some instances of more severe forms like credible threats and hate speech still being detected. Results from my model further suggest that harassment on YouTube is shaped by MPs' gender, background, and party, highlighting the intersectional nature of online harassment.

This study contributes to the VAWIP and political communication literature by demonstrating how YouTube's platform affordances and recommendation systems shape the nature and severity of harassment faced by Canadian MPs. Future research should include a wider range of demographic information about MPs, such as their age, social class, and occupation prior to election to better understand the dynamics of online harassment. Moreover, comparative work across platforms such as YouTube, Twitter (X), and Instagram would help clarify how platform-specific affordances and logics shape the forms, severity, and impact of harassment. Understanding how online harassment operates across platforms is critical, as online harassment should not be a barrier for equal political participation.

References

Arthurs, J., Drakopoulou, S., & Gandini, A. (2018). Researching YouTube. *Convergence: The International Journal of Research into New Media Technologies*, 24(1), 3–15.
<https://doi.org/10.1177/1354856517737222>

- Bardall, G., Bjarnegård, E., & Piscopo, J. M. (2020). How is political violence gendered? Disentangling motives, forms, and impacts. *Political Studies*, 68(4), 916–935. <https://doi.org/10.1177/0032321719881812>
- De Angelis, I. (2024). Torrential Twitter? Measuring the Severity of Harassment When Canadian Female Politicians Tweet About Climate Change. *Social Media + Society*, 10(4), 1–17. <https://doi.org/10.1177/20563051241304493>
- Esposito, E., & Zollo, S. A. (2021). “How dare you call her a pig, I know several pigs who would be upset if they knew”: A multimodal critical discursive approach to online misogyny against UK MPs on YouTube. *Journal of Language Aggression and Conflict*, 9(1), 47–75. <https://doi.org/10.1075/jlac.00053.esp>
- Herrick, R., & Thomas, S. (2024). An Intersectional Exploration Of Psychological Violence, Threats, And Physical Violence Of Mayors In 2021. *American Politics Research*, 52(3), 264–278. <https://doi.org/10.1177/1532673X231220648>
- Jane, E. A. (2017). *Misogyny online: A short (and brutish) history*. SAGE Publications.
- Krook, M. L. (2020). *Violence Against Women in Politics*. Oxford University Press.
- Mamié, R., Horta Ribeiro, M., & West, R. (2021). Are anti-feminist communities gateways to the far right? Evidence from Reddit and YouTube. *Proceedings of the 13th ACM Web Science Conference 2021*, 139–147.
- Mattan, A. J., Harris, A., & Small, T. A. (2024). Mapping Canadian Leaders’ Use of YouTube. *Canadian Parliamentary Review*, 47(3).
- Pedersen, R. T., Petersen, N. B. G., & Thau, M. (2025). Online Abuse of Politicians: 40 Experimental Evidence on Politicians’ Own Perceptions. *Political Behavior*, 47(1), 119–139. <https://doi.org/10.1007/s11109-024-09944-8>
- R Core Team. (2025). *R: A Language And Environment For Statistical Computing*. R Foundation For Statistical Computing. <https://www.R-project.org/>
- Rheault, L., Rayment, E., & Musulan, A. (2019). Politicians in the line of fire: Incivility and the treatment of women on social media. *Research & Politics*, 6(1), 1–7. <https://doi.org/10.1177/2053168018816228>
- Rossini, P. (2022). Beyond Incivility: Understanding Patterns of Uncivil and Intolerant Discourse in Online Political Talk. *Communication Research*, 49(3), 399–425.
- Sobieraj, S. (2020). *Credible Threat: Attacks Against Women Online and the Future of Democracy*. Oxford University Press.
- Southern, R., & Harmer, E. (2021). Twitter, Incivility and “Everyday” Gendered Othering: An Analysis of Tweets Sent to UK Members of Parliament. *Social Science Computer Review*, 39(2), 259–275. <https://doi.org/10.1177/0894439319865519>

Wagner, A. (2020). Tolerating the trolls? Gendered perceptions of online harassment of politicians in Canada. *Feminist Media Studies*, 22(1), 32–47.
<https://doi.org/10.1080/14680777.2020.1749691>

6C.2 Differentiating Contact and Fantasy Online Grooming Offenders

Jake Street, Isibor Ihianle, Ahmad Lotfi (Nottingham Trent University)

Introduction

Online Grooming attacks are deceptive attacks involving adults targeting children online for predominately sexual gratification. These attacks have a significant impact on their victims with numerous recent news stories reporting on the death of a child following one of these attacks (ITV News, 2024; Australian Associated Press, 2024).

The two main legal criteria that are broken in Online Grooming attacks are 'Sexual Communication with a Child' (Sexual Offences Act, 2003) and 'Meeting a Child Following Sexual Grooming' (Sexual Offences Act, 2003). With the former having a maximum sentence of two years in prison and the latter having a maximum sentence of ten years in prison, highlighting the severity that contact offences have.

Throughout literature there have been numerous solutions to the binary classification of Online Grooming with near-perfect classification performance metrics ($F1 = 0.98$) (Nguyen, 2023; Hamm, 2025). However these solutions have not been implemented in social media platforms due to the concerns surrounding privacy that these solutions would have and the generalisability of these approaches due to the datasets commonly used not representing modern day interactions due to the dates this data was collected, with the most utilised dataset in literature (PAN12) having interactions up to 2012 (Inches, 2012) and the most recent open dataset on this including interactions up to 2016 (Street, 2025).

However there has been limited attempts to identify other insights that can help law enforcement prioritise different potential grooming offences. From a legal perspective the most significant insight to obtain, once making the binary classification, is identifying if a contact offence has occurred. Due to the potential for physical harm to the child as well as the sentencing difference between a non-contact grooming offence, 2 years (Sexual Offences Act, 2003), and a contact grooming offence, 10 years (Sexual Offences Act, 2003), the importance of making this distinction is clear.

Being able to make this distinction is also important from a psychological perspective, as Child Contact Sex Offenders (CCSOs) (Chiu, 2018) are more likely to be repeat offenders and are more likely to have cognitive distortions which lead to an internal denial of impact that the offence they committed caused (McCarthy, 2010).

The other offender type in grooming offenses is a Fantasy Child Sex Offender (FCSO) (Chiu, 2018), with these offenders having no interest in having a physical meet up with the child and instead is focused on sexual gratification through the creation of fantasy, perhaps discussing hypothetical sex acts.

It should be noted that CCSOs can also partake in this fantasy creation, so therefore the presence of fantasy is not a reliable indicator of offender type.

Objectives/Research Questions

This investigation aims to answer the following questions 1) How effectively can a model differentiate between CCSO and FCSO language? and 2) How effectively can a model provide confidence to the different degrees of contact indicative language?

The first research question shall utilise a F1 score to determine this, performing binary classification on if a test message block is categorised as CCSO or FCSO. Whereas the second shall observe the precision of the prediction by using Mean Squared Error (MSE) to be able to highlight the high confidence CCSO language (DF).

Methods

As part of this investigation 1136 message blocks from the extended PJ dataset (Street, 2025) were labelled as either being Fantasy Only (FO) (61.3%), Ambiguous Language (AL) (2.9%), Indicative Language (IL) (11.2%), Direct Facilitation (DF) (21.0%), or No Category (3.5%). FO is language that relates to fantasy without any evidence to suggest the fantasy relates to a planned contact event, AL is language that could reasonably be classed as DF or IL however not enough evidence is present to make this choice, and not enough to disprove as FO, IL is language that suggests that a genuine meet up has been arranged, but does not contain content that allows the meet up to occur or a request of this, and DF is requesting to meet up, discussion of sentiments to meeting up, or discussion of details that will allow the meet up to occur at present or in future.

The labelling of these message blocks was conducted by one expert reviewer. Following the labelling of these messages a 30:70 Test-Train split was utilised to fine tune BERT (Devlin, 2019). Predictions were made over 3 k-folds with the following independent variables; including/excluding honeypot child messages, differing weightings for IL (0.7, 0.8, 0.9), and including/excluding AL.

Results

This investigation gives promise that CCSO/FCSO language can be differentiated from one another with F1 scores of different configurations ranging from 0.84-0.92. Additionally MSE values obtained ranged from 0.065-0.025, showing that configuration of these variables strongly affects the error value.

It was found that the most significant change came from the removal of “child honeypot” message lines from the dataset which caused a systematic increase in the MSE value as well as a decrease in the F1 score. Therefore it can be concluded that the additional responses from a child help to build this evidence in determining if the offender is a CCSOs. It should be noted however that as these messages came from a honeypot child and not a real child this may not be the case when implemented in a real world setting, and this may be the product of the honeypot investigator repeating similar phrases when faced with fantasy and contact language.

It was also found that by increasing the IL value used, from 0.7 to 0.9, the F1 score obtained increased (+0.026) however this same increase in IL value caused the MSE value to marginally increase as well (+0.006).

However with the removal of ambiguous message there was little impact on the MSE as well as the F1 score, which can be explained by these messages only constituting 2.9% of the dataset. Therefore a reliable conclusion on this cannot be given and a larger sample is required.

Therefore this investigation has found that to determine contact language from a binary perspective the best configuration, based on the F1 score, is Removing Ambiguous messages, including offender and honeypot child messages, and labelling the IL as 0.9.

Whereas in the use case of highlighting contact language with a high degree of confidence the best configuration, based on the MSE value, is Removing Ambiguous messages, including offender and honeypot child messages, and labelling the IL as 0.8.

Future Work

As part of the future work of this investigation, testing this approach on more Online Grooming negative transcripts will allow for an assessment of the generalisability of the approach to a real world setting. Additionally the linguistics used in the dataset are unlikely to be an accurate representation of modern day direct messages (Beck, 2025) therefore future work should be focused on observing the robustness of this approach on modern day linguistics.

It is suggested that this approach alone does not give enough information to determine an Online Grooming attack, as all transcripts are classified as either fantasy or contact, and this should be implemented alongside traditional binary classification to provide additional insight into the transcript.

Additionally as mentioned only one expert reviewer was used to label these message blocks, future work should incorporate more reviewers to improve the validity of the dataset.

References

Australian Associated Press. (2024, July). Victorian teenager's suicide after sextortion scam prompts call for improved education. The Guardian. Retrieved from <https://www.theguardian.com/australia-news/article/2024/jul/22/victorian-teenagers-suicide-after-sextortion-scam-prompts-call-for-improved-education>

Chiu, M. M., Seigfried-Spellar, K. C., & Ringenberg, T. R. (07 2018). Exploring detection of contact vs. fantasy online sexual offenders in chats with minors: Statistical discourse analysis of self-disclosure and emotion words. *Child Abuse & Neglect*, 81, 128–138. doi:10.1016/j.chiabu.2018.04.004

Hamm, L. (2025). Advancing Grooming Detection in Chat Logs: Comparing Traditional Machine Learning and Large Language Models with a Focus on Predator Tone.

Her Majesty's Government. Sexual Offences Act 2003. , (2003).

Inches, G., & Crestani, F. (2012). Overview of the International Sexual Predator Identification Competition at PAN-2012. webis.de.

Leader, S. (2024, April). 'They Killed Our Son': Online Scammers Tearing Families Apart as Sextortion Cases Soar. ITV News. Retrieved from <https://www.itv.com/news/2024-04-17/they-killed-our-son-the-devastating-impact-of-sextortion-scams-as-cases-soar>

McCarthy, J. A. (2010). Internet sexual activity: A comparison between contact and non-contact child pornography offenders. *Journal of Sexual Aggression*, 16(2), 181–195.

Nguyen, T. T., Wilson, C., & Dalins, J. (2023). Fine-Tuning Llama 2 Large Language Models for Detecting Online Sexual Predatory Chats and Abusive Texts. arXiv [Cs.CL]. Retrieved from <http://arxiv.org/abs/2308.14683>

Street, J., Ihianle, I. K., & Lotfi, A. (2025). Extended Online Grooming Transcript Dataset [Data set]. doi:10.5281/zenodo.15554093

6C.3 Online Racism, Digital Mental Health Tools, and Online Mental Health Communication: A Cross Sectional Study Among Black Young Adults With and Without Depression or Anxiety

Melissa Holland, Kyle Walker, Oreoluwa Badejoh, Vanessa Volpe (North Carolina State University), Brian TaeKyuk Keum (University of California, Berkeley), Jimi Huh, Hans Oh (University of Southern California)

Introduction

Given their convenience and accessibility, digital mental health tools and communication about mental health online are popular among young adults (ages 18–29), who are almost constantly online to connect with others and seek advice [1]. Digital mental health tools are technologies (e.g., apps) that provide support for mental health symptoms [2,3] and are used by those with [4] and without [5] formal mental health diagnoses. Young adults may also engage in online mental health communication, such as by joining a mental health-related online group and sharing personal mental health experiences [6].

Black young adults may particularly benefit from the use of digital mental health tools and online mental health communication due to barriers they face to receiving traditional mental health care (e.g., stigma) [7]. However, they may also face negative race-related experiences as they use digital mental health tools and engage in online mental health communication. One such experience is exposure to anti-Black racism (hereafter "racism"), a sociopolitical system that manifests in differential and negative treatment of and cultural messages about Black individuals [9].

Online racism can entail being the direct target of racist attacks as well as vicariously witnessing racist interactions online, [9,10] such as content circulating on social media of people being assaulted or killed [11]. Exposure to online racism is stressful [12] and threatens one's emotional well-being [13,14]. Such exposure may be associated with the degree to which Black young adults use digital mental health tools and engage in online mental health communication.

The theory of online spaces as both a stressor and a coping tool [15] suggests that digital mental health tools and online mental health communication could be engaged as ways of coping with the stressor of online racism. Black young adults frequently go online to seek support and find racially affirming social connections [16,17].

Conversely, using digital mental health tools or engaging in online mental health communication may itself be associated with greater exposure to online racism, as online racism is structurally embedded in digital technologies and platforms [10]. Experiencing online racism may lead some Black young adults to cope by disengaging from online spaces [18].

Associations between exposure to online racism, digital mental health tool use, and online mental health communication may also differ for those with versus without depression or anxiety. Although individuals with depression or anxiety use digital mental health tools at higher rates than those without [19], these tools are also used by the general public to manage stress and negative emotions [20]. Whether exposure to online racism is associated with use of digital mental health tools and engagement in online mental health communication differently for Black young adults with versus without depression or anxiety remains an empirical question.

Objectives/Research Questions

This study examined whether Black young adults' use of digital mental health tools and frequency of engagement in online mental health communication are associated with exposure to online racism. We also examined whether these associations differed for Black young adults with versus without depression or anxiety. Given the potential bidirectional nature of these relationships, we conducted an initial exploratory correlational investigation. As digital mental health resources are growing in popularity, such initial descriptive research can elucidate considerations for seeking digital mental health tools among Black young adults using digital and online technologies.

Methods

Data came from a larger cross-sectional online survey of people in the United States conducted from July to September 2024. Participants were recruited through Qualtrics Panels using quota sampling for gender, education, and geographic region, and multiple procedures were used to ensure data quality. The analytic sample included 1,005 monoracial Black young adults ages 18–29 ($M_{age} = 24.07$, $SD = 3.04$). Approximately half identified as women (50.6%).

Measures

Use of digital mental health tools was assessed by asking participants whether they had used any digital tools for their mental health or emotions in the past year, including wellness apps, mental health apps, online or telephone-based therapy services, and online support groups. Responses were coded dichotomously (any use vs. no use).

Participants indicated how often they had communicated about mental health online in the past year using four items developed for the survey. The items included: (1) join forums or closed social media groups on specific mental health topics, (2) do hashtag

searches for mental health topics on social media, (3) follow people online who share about mental health conditions, and (4) share about your own poor mental well-being online. Participants indicated the frequency of each item using a Likert-type response scale from 0 (never) to 4 (once a week or multiple times a week). Items were averaged, with higher scores indicating more frequent engagement (sample $\alpha = .87$).

Exposure to online racism was measured using the Perceived Online Racism Scale–Very Brief, which assesses the frequency of six race-related online experiences in the past six months, using a Likert-type scale from 1 (never) to 5 (always). Higher mean scores indicated more frequent exposure (sample $\alpha = .90$).

Depression was measured using the Patient Health Questionnaire-9 (sample $\alpha = .90$), and anxiety was measured using the General Anxiety Disorder 7-item scale (sample $\alpha = .92$). Scores were dichotomized using established clinical cutoffs to indicate the presence or absence of clinically significant depressive or anxiety symptoms. Age, gender, and daily internet use were included as analytic covariates.

Analysis Plan

Descriptive statistics were first computed. Two separate path analysis regression models were then conducted in Mplus using maximum likelihood estimation given the presence of missing data. Model 1 included predictors of online racism and depression, and the interaction between the two, on outcomes of use of digital mental health tools (modeled with logistic regression) and engagement in online racism (modeled with linear regression). Model 2 followed the same analytic approach, substituting anxiety for depression. As depression and anxiety were relatively strongly correlated ($r(1003) = .61, p < .001$), we included these variables in separate models and used a Bonferroni alpha adjustment ($\alpha = .05$ divided by two models for a corrected $\alpha = .025$).

5. Results

Many participants resided in the southern United States. Additionally, 88.66% reported exposure to online racism in the past six months, with exposure occurring on average between “sometimes” and “half the time.” Approximately 51.5% reported using one or more digital mental health tools in the past year, and participants engaged in online mental health communication less than once per month on average. Nearly half of the sample met criteria for depression, and 42.4% met criteria for anxiety.

More frequent exposure to online racism was associated with a significantly greater likelihood of using digital mental health tools (ORs ranged from 1.72 to 1.84, $p < .001$) and more frequent engagement in online mental health communication (β s ranged from .31 to .36, $p < .001$). Controlling for online racism, Black young adults with depression or anxiety were also more likely to use digital mental health tools and engaged more frequently in online mental health communication than those without

these conditions. Neither depression nor anxiety significantly moderated the association between exposure to online racism and either outcome.

Future Work

As digital and online mental health solutions continue to be developed as promising and accessible new treatments [3], the current study considers the role of online racism in how Black young adult populations seek digital mental health resources and provides design and research implications for such solutions for this population.

Because this study was cross-sectional, future research should examine these variables using longitudinal and daily diary designs to measure use and engagement specifically in response to both general stress and specific stressors like online racism. Another future direction is to investigate the types of online spaces where Black young adults use digital mental health tools and engage in online mental health communication. Research suggests that Black young adults may feel more comfortable discussing these experiences in anonymous communities than on public, social media platforms like Facebook [22]. Future work could explore how platform type, privacy, and identity shape help-seeking. As digital mental health resources continue to expand, ensuring these tools and opportunities for communication are safe, free of online racism, and culturally relevant remains an ongoing parallel priority for researchers, policymakers, and developers.

References

1. Pempek, T. A., Yermolayeva, Y. A., & Calvert, S. L. (2009). College students' social networking experiences on Facebook. *Journal of Applied Developmental Psychology*, 30(3), 227-238.
2. Hamdoun, S., Monteleone, R., Bookman, T., & Michael, K. (2023). AI-based and digital mental health apps: Balancing need and risk. *IEEE Technology and Society Magazine*, 42(1), 25-36.
3. Meyerhoff, J., Kornfield, R., Mohr, D. C., & Reddy, M. (2022). Meeting young adults' social support needs across the health behavior change journey: Implications for digital mental health tools. *Proceedings of the ACM on Human-computer Interaction*, 6(CSCW2), 1-33.
4. Nicholas, J., Boydell, K., & Christensen, H. (2017). Beyond symptom monitoring: Consumer needs for bipolar disorder self-management using smartphones. *European Psychiatry*, 44, 210-216.
5. Pretorius, C., Chambers, D., Cowan, B., & Coyle, D. (2019). Young people seeking help online for mental health: cross-sectional survey study. *JMIR Mental Health*, 6(8), e13524.

6. Neal, D. M., Campbell, A. J., Williams, L. Y., Liu, Y., & Nussbaumer, D. (2011). "I did not realize so many options are available": Cognitive authority, emerging adults, and e-mental health. *Library & Information Science Research*, 33(1), 25-33.
7. Willis, H. A., & Neblett, E. W. (2023). Developing culturally-adapted mobile mental health interventions: a mixed methods approach. *Mhealth*, 9, 1.
8. Williams, D. R., Neighbors, H. W., & Jackson, J. S. (2003). Racial/ethnic discrimination and health: Findings from community studies. *American Journal of Public Health*, 93(2), 200-208.
9. Banks, J., & Ward, L. M. (2025). From microaggressions to viral trends: A qualitative examination of Black emerging adults' experiences with online racial discrimination. *Emerging Adulthood*, 13(2), 340-354.
10. Volpe, V. V., Hoggard, L. S., Willis, H. A., & Tynes, B. M. (2021). Anti-Black structural racism goes online: A conceptual model for racial health disparities research. *Ethnicity & Disease*, 31(Suppl 1), 311.
11. Motley Jr, R. O., Chen, Y. C., Johnson, C., & Joe, S. (2020). Exposure to community-based violence on social media among Black male emerging adults involved with the criminal justice system. *Social Work Research*, 44(2), 87-97.
12. Carter, R. T. (2007). Racism and psychological and emotional injury: Recognizing and assessing race-based traumatic stress. *The Counseling Psychologist*, 35(1), 13-105.
13. Grapin, S. L., Warner, C. M., Bixter, M. T., Cunningham, D. J., Bonumwezi, J., Mahmud, F., ... & Nisenson, D. (2024). Online racial discrimination and mental health among Black undergraduates: The moderating role of gender. *Journal of American College Health*, 72(1), 310-318.
14. Park, J., Hallman, J., Liu, X. S., & Hancock, J. (2024). Black representation in social media well-being research: A scoping review of social media experience and psychological well-being among Black users in the United States. *New Media & Society*, 26(3), 1670-1702.
15. Wolfers, L. N., & Utz, S. (2022). Social media use, stress, and coping. *Current Opinion in Psychology*, 45, 101305.
16. Volpe, V. V., Buhrman, G. W., Boaheng, P., Holliday, D., Nick, E. A., & Criss, S. (2023). Speak [ing] My Mind. *Journal of Media Psychology*.
17. George Mwangi, C. A., Bettencourt, G. M., & Malaney, V. K. (2018). Collegians creating (counter) space online: A critical discourse analysis of the I, Too, Am social media movement. *Journal of Diversity in Higher Education*, 11(2), 146.

18. Musgrave, T., Cummings, A., & Schoenebeck, S. (2022, April). Experiences of harm, healing, and joy among Black women and femmes on social media. In Proceedings of the 2022 CHI conference on human factors in computing systems (pp. 1-17).
19. Sorkin, D. H., Janio, E. A., Eikey, E. V., Schneider, M., Davis, K., Schueller, S. M., ... & Mukamel, D. B. (2021). Rise in use of digital mental health tools and technologies in the United States during the COVID-19 pandemic: Survey study. *Journal of Medical Internet Research*, 23(4), e26994.
20. Montagni, I., Tzourio, C., Cousin, T., Sagara, J. A., Bada-Alonzi, J., & Horgan, A. (2020). Mental health-related digital use by university students: A systematic review. *Telemedicine and e-Health*, 26(2), 131-146.
21. To, A., Sweeney, W., Hammer, J., & Kaufman, G. (2020). " They Just Don't Get It": Towards Social Technologies for Coping with Interpersonal Racism. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1), 1-29.

6C.4 Social Media Use, Online Racism, and Mental Health among Black Emerging Adults

Oreoluwa Badejoh (North Carolina State University), Shedrick Garrett (University of North Carolina at Chapel Hill), Melissa Holland, Vanessa Volpe (North Carolina State University), Brian TaeHyuk Keum (University of California, Berkeley), Jimi Huh, Hans Oh (University of Southern California)

Introduction

Social media use has been flagged as a key risk factor for mental health symptomatology among emerging adults. Emerging adulthood (ages 18 -29), a distinct developmental period characterized by increased independence and ushering of adult responsibilities has been linked to strains on mental health (Arnett, 2000, 2006). For instance, 12-month prevalence of psychiatric disorders is highest among individuals aged 18 to 29 years. (Kessler et al., 2005). This is especially relevant for Black emerging adults who are some of the highest users of social media platforms, with over half being active on social media platforms several times a day (Auxier & Anderson, 2021).

For this population, race-related stressors, such as discrimination, coupled with stressors tied to emerging adulthood can pose an increased risk for depression and anxiety (Hope et al., 2015). In the online environment, exposure to online racial discrimination is a common occurrence– with most Black emerging adults reporting experiencing online racism at least several times per year (Volpe et al., 2023; Onuoha et al., 2024). Therefore, social media use may be linked to online racism– one part of the association between frequent social media use and more depression and anxiety.

According to racial trauma theory (Carter, 2007), exposure to racism – whether offline or online – exerts a psychological toll on the individual who is directly or indirectly the target. However, no studies to our knowledge have assessed online racism as a pathway through which social media use can cause negative mental health outcomes. Therefore, we aimed to fill this gap in two ways: (1) by examining online racism as a mediator (see Figure 1), and (2) by examining platform-specific social media use, rather than grouping all social media platforms into a general measure of social media use. Discerning between social media platforms may provide specific implications for social media policy and intervention efforts.

Objectives/Research Questions

Given the frequency of online racism exposure and social media use among Black emerging adults, the goal of the current study was to examine the frequency of exposure to online racism as a mediator in the association between the frequency of social media use and depression and anxiety symptoms, and deciphering how specific platforms (i.e., YouTube, Instagram, Facebook, Twitter, Reddit, TikTok) are linked to mental health symptomatology. Drawing from the racial trauma framework (Carter, 2007), We hypothesized that there would be a significant indirect effect of online racism exposure, such that more frequent social media use would be associated with more frequent exposure to online racism and that more frequent exposure to online racism would be associated with greater odds of depression and anxiety. We also considered whether there were significant associations between social media use and mental health variables for Black emerging adults. Although we did not have distinct hypotheses for each social media platform, we considered social media use separately for each social media platform (i.e., YouTube, Instagram, Facebook, Twitter, Reddit, TikTok), to better inform clinical and policy efforts.

Methods

Data come from a larger investigation of emerging adults (ages 18-29), residing in the United States, which was conducted online from July to September 2024. Participants were recruited via Qualtrics Panels, which generated a non-probability sample of 1005 monoracial Black emerging adults (Mage = 24.07, 50.6% women). Participants completed an online cross-sectional survey, in which they reported their social media use, exposure to online racism, and mental health symptoms.

Measures

Depression and anxiety were each measured using two self-report scales to index clinically significant mental health symptomatology: The Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001) was used to measure the degree to which participants were bothered by nine common depressive symptoms (e.g., “little interest or pleasure in doing things,”) over the past two weeks on a Likert-type scale from 1 (Not at all) to 4 (Nearly every day). The General Anxiety Disorder 7-item scale (GAD-7; Spitzer et al., 2006) was used to measure the degree to which participants were bothered by seven common anxiety symptoms (e.g., “feeling nervous or anxious, etc.) over the past two weeks on a Likert-type scale from 1 (Not at all) to 4 (Nearly every day). In consideration of clinically significant depression (sample $\alpha = .90$) and anxiety (sample $\alpha = .92$), each continuous sum scale score was dichotomized to correspond to established clinical cutoffs for major depressive disorder (PHQ-9 score ≥ 10) and general anxiety disorder (GAD-7 score ≥ 10), respectively. Depression or anxiety scores exceeding the cutoffs were coded as 1, and scores below the cutoffs were coded as 0.

The Perceived Online Racism Scale - Very Brief (Keum, 2021) was used to index exposure to online racism. This six-item scale asks participants to report their frequency of experiencing online racism on a Likert-type scale from 1 (Never) to 5 (Always). A mean score was calculated (sample $\alpha = .90$), with a higher score representing more frequent exposure to online racism.

The frequency of using specific social media platforms was assessed using six items developed for the survey. Participants were asked to rate the frequency of their use of specific social media platforms: YouTube, Instagram, Facebook, Twitter (X), Reddit, and TikTok on a typical day, rating each platform using a Likert-type response scale from 0 (Not at all) to 4 (Almost constantly). These platforms were considered because they were among the most popular platforms used by Black emerging adults when data were collected (Gottfried, 2024). Each item was included as a separate predictor variable in our model to examine platform-specific associations, controlling for the use of other platforms. Higher scores indicate more frequent daily use of that social media platform.

Results

On average, participants reported using YouTube many times a day ($M = 3.12$), Instagram ($M = 2.67$) and TikTok ($M = 2.48$) between a few and many times a day, Facebook ($M = 2.13$) and Twitter/X ($M = 1.92$) a few times a day, and Reddit less than a few times a day ($M = 1.28$). They reported exposure to online racism between “sometimes” and “half the time” in the past six months ($M = 2.58$). About half of the sample had PHQ-9 scores indicating depression (49.5%), and about forty percent had GAD-7 scores indicating anxiety (42.4%).

The indirect effect of social media use on depression and anxiety through exposure to online racism was supported for use of Twitter (depression indirect effect = .04, $p = .027$, 95% CI [.01, .09]; anxiety indirect effect = .04, $p = .027$, 95% CI [.01, .08]), Reddit (depression indirect effect = .10, $p < .001$, 95% CI [.06, .15]; anxiety indirect effect = .10, $p < .001$, 95% CI [.06, .14]), and TikTok (depression indirect effect = .06, $p = .005$, 95% CI [.02, .10]; anxiety indirect effect = .05, $p = .005$, 95% CI [.02, .09]), but not YouTube, Instagram, or Facebook.

More frequent daily use of Twitter ($\beta = .09$, $p = .020$, 95% CI [.01, .16]), Reddit ($\beta = .19$, $p < .001$, 95% CI [.12, .26]), and TikTok ($\beta = .11$, $p = .002$, 95% CI [.04, .18]) were each associated with more frequent exposure to online racism (i.e., a paths of the mediation model). In turn, more frequent exposure to online racism was associated with a 98% increased odds of depression (OR = 1.98, $p < .001$, 95% CI [1.74, 2.31]) and a 90% increased odds of anxiety (OR = 1.90, $p < .001$, 95% CI [1.46, 2.48]) (i.e., b paths of the mediation model).

Finally, although online racism did not mediate the association between daily Instagram use and depression or anxiety, more frequent daily Instagram use itself was also associated with 19% decreased odds of depression (OR = .81, $p < .001$, 95% CI [.72, .92]).

Future Work

Social media is an important landscape in the lives of Black emerging adults' online experiences—providing opportunities for collective identity and belonging in the midst of life transitions (Florini, 2014), but also exposure to negative race-related experiences, such as online racism. In summary, the current study considered the role of exposure to online racism as one underlying factor that could explain the relationship between social media use and mental health for Black emerging adults, a population that is frequently online, while highlighting these differences per platform.

Findings underscore the importance of understanding the features or interactions unique to different platforms, namely Reddit, Twitter/X, and TikTok, that expose Black emerging adults to online racism and increase their risk of mental health challenges. Future research could explore whether platform features like content warnings, hate speech moderation, and bias response teams could also help reduce exposure. Additionally, future studies should further consider the role of intersectional factors (e.g., gender) in influencing both the online experiences of Black emerging adults and also how such varied experiences may affect their mental health.

References

- Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, 55(5), 469–480.
<https://doi.org/10.1037/0003-066X.55.5.469>
- Arnett, J. J. (2006). Emerging Adulthood: Understanding the New Way of Coming of Age. In J.J. Arnett & J. L. Tanner (Eds.), *Emerging adults in America: Coming of age in the 21st century*. (pp. 3–19). American Psychological Association.
<https://doi.org/10.1037/11381-001>
- Auxier, B., & Anderson, M. (2021). Social Media Use in 2021. Pew Research Center.
https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2021/04/PI_2021.04.07_Social-Media-Use_FINAL.pdf
- Carter, R. T. (2007). Racism and Psychological and Emotional Injury: Recognizing and Assessing Race-Based Traumatic Stress. *The Counseling Psychologist*, 35(1), 13–105. <https://doi.org/10.1177/0011000006292033>
- Florini, S. (2014). Tweets, Tweeps, and Signifyin': Communication and Cultural Performance on "Black Twitter." *Television & New Media*, 15(3), 223–237.
<https://doi.org/10.1177/1527476413480247>

- Gottfried, J. (2024). Americans' Social Media Use. Pew Research Center.
https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2024/01/PI_2024.01.31_Social-Media-use_report.pdf
- Hope, E. C., Hoggard, L. S., & Thomas, A. (2015). Emerging into adulthood in the face of racial discrimination: Physiological, psychological, and sociopolitical consequences for african american youth. *Translational Issues in Psychological Science*, 1(4), 342–351. <https://doi.org/10.1037/tps0000041>
- Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, Severity, and Comorbidity of 12-Month DSM-IV Disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6), 617.
<https://doi.org/10.1001/archpsyc.62.6.617>
- Keum, B. T. (2021). Development and validation of the Perceived Online Racism Scale short form (15 items) and very brief (six items). *Computers in Human Behavior Reports*, 3, 100082. <https://doi.org/10.1016/j.chbr.2021.100082>
- Kroenke, K., Spitzer, R. L., & W. Williams, J. B. (2001). The PHQ-9. *Journal of General Internal Medicine*, 16(9), 606-613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Onuoha, A. C., Matsuzaka, S., Stanton, A. G., Volpe, V. V., & Avery, L. R. (2024). Young Black American Women's Social Media Use and Online Victimization. *Social Media + Society*, 10(3), 20563051241277610.
<https://doi.org/10.1177/20563051241277607>
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine*, 166(10), 1092-1097. doi:10.1001/archinte.166.10.1092
- Volpe, V. V., Buhrman, G. W., Boaheng, P., Holliday, D., Nick, E. A., & Criss, S. (2023). "Speak[ing] My Mind": Reasons for Using Twitter and the Online Experiences, Critical Media Literacy, and Racial Identity of Black American Emerging Adults. *Journal of Media Psychology*, 35(5), 268–279. <https://doi.org/10.1027/1864-1105/a000387>

Session 6D - Social Media Cultures and Everyday Life (5)

Day 2 · Tues 14 July · 9:30–11:00, Seminar Suite C, ARC

6D.1 From Forums to Feeds: The Platformized Transformation of Chinese K-pop Fandom

Shangru Jiang (University of Bristol)

Introduction

As one of the largest overseas markets for Korean popular culture, China has played a central role in the transnational circulation of K-pop. However, Chinese fans have long engaged with K-pop under infrastructural and political constraints. In the early 2000s, Korean idols were introduced through traditional media such as television and print magazines. As digital access expanded, fans relied on subtitling teams to access content and later built vibrant online communities through forums and Baidu Tieba. However, China's Great Firewall restricted access to Korean platforms like Daum cafés, Twitter, and Instagram, creating a domestically bounded K-pop fandom. Today, Chinese fans operate primarily within highly platformized environments such as Weibo, Bilibili, and Xiaohongshu. These constraints deepened after the 2016 THAAD dispute, which led to a ban on Korean idols in China, and later during the COVID-19 pandemic, which disrupted offline fan gatherings. Despite these obstacles, China continues to host the world's largest number of Hallyu fan clubs—nearly 68% of which are K-pop related (The Korea Foundation, 2023).

This study focuses on fans of Super Junior as a case study to explore how platform transformations have influenced Chinese K-pop fandom and identity construction over the past two decades. Debuting in 2005, Super Junior was among the earliest K-pop groups to achieve sustained popularity in China, aided by the inclusion of Han Geng, the first Chinese member in a major Korean boy band. The group's visibility expanded through Korean variety shows such as X-Man and Love Letter, which were broadcast on Chinese television, and through the establishment of Super Junior-M in 2008, which specifically targeted the Chinese market with Mandarin-language releases. As Super Junior celebrates its twentieth anniversary in 2025, the group's longevity provides a unique lens for examining how Chinese fans have navigated evolving social media infrastructures from early online forums to algorithmically driven platforms and how

these shifts have reconfigured the social, emotional, and participatory conditions of fandom.

Objectives/Research Questions

The primary objective of this research is to examine how the evolving ecosystem of Chinese social media platforms has reshaped the identity construction and participatory practices of K-pop fans. By centering on the fandom of Super Junior, a group that bridges the “Hallyu 2.0” era and the contemporary algorithmic age, the study seeks to understand how platformization and generational experience co-produce distinct fan subjectivities. It explores how veteran fans who have experienced twenty years of technological transformation negotiate fandom differently from newer participants, and how platform logics interact with personal trajectories to create diverse modes of engagement within the same fandom.

Specifically, the research investigates how the spatial and discursive boundaries of fandom have evolved alongside shifts in China’s social media landscape, how fans at different stages of their “fan life” negotiate agency within algorithmic systems and institutionalized digital labor, and how platform affordances intersect with personal aging and life-course transitions to reconstruct fan identity. Through this investigation, the study illuminates how fans’ evolving relationships with platforms reflect broader transformations in digital citizenship, affective labor, and mediated selfhood in China’s platformized cultural environment.

Methods

This research adopts a qualitative methodology combining netnography with semi-structured interviews to capture fans’ lived experiences and their evolving engagement practices. Data were collected from ten Chinese Super Junior fans, seven of whom have followed the group for over 15 years and continue to express active interest, including plans to attend the group’s twentieth-anniversary concert. The researcher’s insider positionality as a former fan and fan-site moderator provided valuable ethnographic access and reflexive depth. Interviews were conducted via WeChat and Tencent Meeting and fully transcribed

Thematic analysis followed Braun and Clarke’s (2022) five-step framework, supported by NVivo for data organization. Coding proceeded through three iterative stages: open, focused, and axial (Saldaña, 2021), to identify recurring themes across different platform contexts. Cross-platform comparison was emphasized to trace how fan practices, discourse, and identity construction evolved in tandem with China’s shifting digital infrastructures.

Results

This case study explores how long-term Chinese fans of Super Junior negotiate platform transitions and how such shifts reshape fan identity over time. By tracing participants' experiences from early fan forums to contemporary platforms such as Weibo and Xiaohongshu, this analysis demonstrates that platform changes not only introduce new technical affordances but also fundamentally reconfigure the social, emotional, and relational conditions of fandom participation.

The findings reveal three interrelated transformations. First, platform transitions restructure the spatial organization of fandom from enclosed, community-oriented spaces to individualized participation, and more recently, to fluid, algorithmically mediated modes of engagement. Second, these spatial transformations coincide with a reversal in relationship formation: from offline-first, socially embedded networks to online-first encounters characterized by weak ties and transient interactions. Third, these structural changes intersect with fans' life-course trajectories, contributing to a reorientation of fan identity from collective responsibility and affective labor toward individualized enjoyment and emotional self-regulation.

Crucially, the study shows that shifts in fan identity cannot be attributed solely to technological innovation. Platform affordances interact with broader social factors such as age, employment, and time availability to shape how long-term fans recalibrate their participation. Rather than disengaging from fandom, these fans adopt more flexible, selective, and cross-platform practices that allow fandom to coexist with everyday obligations.

While this case focuses specifically on Super Junior fans, its broader implications lie in revealing the dynamics of platform-mediated fandom in the Chinese context. The identified patterns: spatial restructuring, relational reorientation, and individualized identity, offer a conceptual framework for understanding how social media platforms reshape fan identities beyond a single fandom.

Moreover, the findings suggest that the impact of platformization is significantly moderated by the idol's career trajectory. As Super Junior has transitioned into a "mature" stage, the demand for aggressive algorithmic visibility and large-scale data labor has comparatively diminished. This evolution fosters a distinctive fandom atmosphere in which both veteran fans and newcomers prioritize emotional sustainability and personal enjoyment over collective performative labor. Consequently, the emergence of "invisible fandom" reflects a co-produced outcome of platform affordances and the natural maturation of the idol-fan relationship, allowing fandom to be integrated into everyday life as a form of emotional self-regulation rather than an external obligation.

Future Work

Future research will extend this case study by incorporating a wider range of Chinese K-pop fans across different age groups, fandom stages, and platform contexts. Through additional interviews and survey data, the next phase will compare how fans construct identity and negotiate participation across diverse digital ecosystems from algorithmic feeds on Xiaohongshu to video-based engagement on Douyin and community-based interactions on Douban. By juxtaposing the longitudinal experiences of veteran fans with the emerging practices of newer participants, this study aims to develop a comparative framework for understanding how platform governance, algorithmic visibility, and social infrastructure jointly shape fandom in China's evolving digital environment.

References

Braun, V., & Clarke, V. (2022). *Thematic Analysis: A Practical Guide*. Sage.

The Korean Foundation. (2024, April 5). 2023 Analysis of Global Hallyu Status by The Korea Foundation - Issuu.
https://issuu.com/the_korea_foundation/docs/0402_global_hallyu_status_eng

Saldaña, J. (2021). *The Coding Manual for Qualitative Researchers*. 4th ed. Los Angeles: Sage.

6D.2 Chords of Small Media Through Time

Sahar Sagha (Durham University)

Introduction

This paper explores how ordinary images circulating on social media become politically meaningful in contexts of protest. Focusing on the Woman, Life, Freedom movement in Iran, it examines how everyday photographs—such as family images, portraits, and personal snapshots—shift from intimate forms of self-expression to collective symbols of solidarity and resistance.

Building on the concept of small media (Sreberny and Mohammadi 1990), the paper argues that images function as small media in two distinct but connected ways. Before moments of protest, personal images circulate within social networks where they quietly contradict state-imposed norms and generate shared sensibilities through repetition and recognition. During protest, this logic transforms: solidarity emerges not through individual self-representation, but through the collective propagation of the same images of protesters, which gradually acquire iconic status.

The paper conceptualises this transformation through the notions of remediation and visual dialogue. Rather than treating images as static representations, it foregrounds how meaning is produced through circulation, repetition, and variation. Each act of sharing, remaking, or recontextualising an image is understood as a dialogic response that strengthens affective connections among dispersed participants. Through this process, images become anchors of collective memory and emotional alignment.

Objectives/Research Questions

This paper aims to investigate how images circulating on social media operate as forms of small media in different phases of political mobilisation. Rather than treating protest imagery as something that emerges only during moments of uprising, the research examines how everyday visual practices gradually prepare the conditions for collective action.

The paper pursues three main objectives. First, it seeks to analyse how personal images shared on social media before protest moments function as relational practices that quietly challenge state norms and produce shared sensibilities within social networks. Second, it aims to explore how, during protest, the circulation of images shifts from personal self-representation to the collective propagation of the same visual references, through which solidarity and iconic meaning are produced. Third, it aims to conceptualise remediation and visual dialogue as key mechanisms through which images acquire political significance, emotional resonance, and durability over time.

Guided by these objectives, the paper addresses the following research questions:
How do ordinary images circulate as small media in everyday digital life prior to moments of protest?

How does the function of images change during protest, from personal expression to collective symbolic action?

How do processes of remediation and visual dialogue shape the production of meaning, affective connection, and iconicity in protest imagery?

By addressing these questions, the paper seeks to contribute to a deeper understanding of how intimacy, repetition, and visual circulation transform everyday images into collective political resources within networked publics.

Methods

This paper adopts a qualitative and interpretive approach grounded in visual culture studies and critical media analysis. Rather than treating protest images as fixed objects or data points, it approaches them as relational and affective forms whose meanings emerge through circulation, repetition, and transformation across social media environments.

The analysis focuses on selected images associated with the Woman, Life, Freedom movement, including personal photographs shared by families of victims, widely circulated portraits of protesters, and their remediations across Instagram.

Through this approach, the paper seeks to understand protest imagery not simply as representation, but as a process through which affective connection, collective memory, and political meaning are continuously produced within networked publics.

Results

The analysis reveals that images circulating on social media operate as small media in two distinct but interconnected modes, corresponding to different phases of political mobilisation.

First, before moments of protest, personal images function as relational practices within social networks. These images—such as everyday photographs, profile pictures, and intimate visual moments—quietly contradict state-imposed norms and gradually normalise alternative ways of appearing and living. Through repetition and recognition within circles of friends and followers, such images generate shared sensibilities and affective alignment without explicit political intent. In this phase, images remain personal, but they cultivate habits of visibility and connection that prepare the ground for collective action.

Second, during protest, the function of images shifts from personal circulation to collective propagation. Solidarity is produced not through individual self-

representation, but through the widespread repetition of the same images of protesters and victims. Through processes of remediation—such as redrawing, stylising, and recontextualising—these images become sites of visual dialogue. Meaning emerges not from a single authoritative interpretation, but from the accumulation of responses that reinforce emotional connection and collective identity.

The findings also show that certain images undergo rapid iconic condensation, transforming from ordinary photographs into shared symbols of the movement. This process does not erase intimacy; rather, it reconfigures it, allowing personal loss to be articulated as collective memory. Iconic images thus function as affective anchors that stabilise meaning within the fluid dynamics of social media.

Overall, the results demonstrate that the political power of protest imagery lies not merely in visibility or documentation, but in the capacity of images to mediate relationships, sustain affective ties, and transform everyday visual practices into collective political resources over time.

Future Work

Future work will further develop the theoretical framework of visual dialogue and small media by engaging more systematically with platform-specific dynamics. While this paper focuses on the affective and relational dimensions of image circulation, subsequent research will examine how platform affordances, algorithms, and moderation practices influence the visibility, longevity, and transformation of protest imagery.

References

- Bolter, J. D., & Grusin, R. (1999). *Remediation: Understanding new media*. MIT Press.
- Casas, A., & Williams, N. W. (2019). Images that matter: Online protests and the mobilizing role of pictures. *Political Research Quarterly*, 72(2), 360–375.
<https://doi.org/10.1177/1065912918786805>
- Hariman, R., & Lucaites, J. L. (2007). *No caption needed: Iconic photographs, public culture, and liberal democracy*. University of Chicago Press.
- Hariman, R., & Lucaites, J. L. (2015). Icons, iconicity, and cultural critique. *Sociologica*, 9(1). <https://doi.org/10.2383/80393>
- Papacharissi, Z. (2016). Affective publics and structures of storytelling: Sentiment, events and mediality. *Information, Communication & Society*, 19(3), 307–324.
<https://doi.org/10.1080/1369118X.2015.1109697>

Sekula, A. (1982). On the invention of photographic meaning. In V. Burgin (Ed.), *Thinking photography* (pp. 84–109). Macmillan.

Sreberny, A., & Mohammadi, A. (1990). Small media for a big revolution: Iran. *International Journal of Politics, Culture, and Society*, 3(3), 341–371.
<https://doi.org/10.1007/BF01384966>

Tilly, C. (1989). *Big structures, large processes, huge comparisons*. Russell Sage Foundation.

6D.3 Masculinity and Social Change: A dramaturgical explanation of online misogyny

Shona Rooney (Queen's University Belfast)

Introduction

In recent years, online misogynistic platforms and forums, often grouped under the umbrella term the manosphere, have received increasing public, political, and academic attention. These loosely connected online spaces, including Men's Rights Activists (MRAs), Pick-Up Artists (PUAs), Men Going Their Own Way (MGTOW), and Incels, have been linked to the normalisation of misogynistic narratives, harassment campaigns, and, in extreme cases, offline violence. While this is a relatively new area of study, previous literature has explored the misogynistic online content posted by social actors, highlighting the anger such members appear to hold against women (Rafail and Freitas, 2019; Ging, 2017; Sugiura, 2021; Sparks, et al., 2022). This research frames the manosphere as emerging during a period where there is significant change regarding gender norms. Across these forums, participants repeatedly articulate a belief that traditional masculine norms are being eroded by shifting economic conditions, changing intimate norms, and feminist political gains. These losses are not only expressed through explicit misogyny, but through complex emotional performances of grievance, resentment, anger and shame. Rather than viewing these expressions as purely irrational, this study approaches them as socially meaningful responses to normative conflict, where masculinity is felt to be under threat. This research thus explores the emotion culture underpinning misogynistic online forums which is crucial for addressing the persistence of online misogyny, informing effective interventions, and supporting individuals impacted by these communities.

Objectives/Research Questions

Previous literature has explored the misogynistic online content posted by social actors, focusing in particular on the anger and resentment that such individuals direct towards women. As well as the real-world implications of such environments, including offline violence (Laskovtsov, 2020:10; Hoffman, et al., 2020:4; Ging, 2017:647). This research aims to address the academic gaps within this relatively new research topic by taking a dramaturgical approach to explore the emotional culture that exists within these spaces

Research aims and objectives

This research aims to

- 1) Understand how the normative qualities of gender dramas are conveyed through emotional expression
- 2) Explore how modern masculinity can be an arena for normative conflict.
- 3) Uncover shared emotional narratives among members of misogynistic online networks

To achieve these aims, this study plans to:

- 1) Explore how modern conceptions of masculine norms are debated, contested, and redefined by collecting research material from online misogynistic networks
- 2) Embrace a dramaturgical framework to effectively identify the emotional underpinning of conflicts, discussions, and areas of tension among forum members.

Research Questions

This study aims to address the following research questions:

- 1) Within online misogynistic networks, what is the normative content of forum threads concerning the changing status of men within society in relation to sexual relationships, economic opportunities, and political and social values?
- 2) What does the emotional character of these posts occurring within the 'manosphere' reveal about normative conflict and change concerning gender?
- 3) Are there any other polarised and intense social and political claims that these groups are making, outside of promoting misogynistic norms?

Methods

A primary framework underpinning this study is dramaturgy with a particular emphasis on emotion culture to critically examine how gender norms are performed and conveyed through emotional expression in online spaces. Research material was selected using purposive sampling from a range of publicly accessible online forums representing the most notable manosphere communities, MRA, MGTOW, PUA, and Incels. These forums were found on mainstream social media sites such as 'Reddit' but also on fringe social media platforms such as 'Communities.win', which are known for their lack of policies and regulations. These spaces also included a variety of interactions, including commentary about politics and news stories, memes, personal anecdotes, and long comment threads that many users engaged in. The research gathered material from archived threads posted between January 2021 and January 2024, initially observing 5,760 posts before selecting 920 relevant posts regarding masculinity and social change. Qualitative data analysis provides this research with a nuanced lens for investigating the intricate ways in which gender, norms, and emotions interact within online misogynistic networks (Smilde and Hanson, 2018:333). This study also uses an abductive approach to analysis, which should allow the researcher to

develop an in-depth understanding of those involved within these spaces. An abductive approach acts as a form of reasoning whereby the researcher attempts to find a situational fit between observed facts and norms (Timmermans and Tavory, 2012:171). Unlike other approaches, the goal of abduction is not to arrive at a specific or generalised conclusion, but rather to make sense of and better understand a social phenomenon (Brinkmann, 2014:722).

This study also recognised some key ethical concerns, including obtaining informed consent. Some researchers argue that it is unnecessary when studying publicly available information, as this is no different to traditional media research (Sugiura et al., 2016; Langer and Beckman, 2005). Moreover, making participants aware of a researcher's presence could have had negative repercussions, especially for female researchers in misogynistic online environments. Confidentiality and data anonymisation were therefore critical to this study. The researcher avoided collecting personal data and used pseudonyms, anonymised quotes, and stored materials securely. Aggregated quotations were also used with slight wording changes, while maintaining the initial meaning, to ensure anonymity by preventing quotes from being located by back-tracing them through search engines.

Results

Data analysis is nearing completion, with preliminary findings presented across four sections.

Legitimate Victims

This section explores the MRA forum. Here, misogynistic comments are often cloaked in the language of rationality, logic, and social reform. Within these forums, specifically MRA, users appear to perform an identity of a good man wronged, a person who claims to have fulfilled societal expectations only to be rejected, mischaracterised, and punished. Their emotional expressions of frustration and injustice are articulated and framed as rational and reasonable. This also acts to distinguish MRA forums from what they perceive as the more extreme factions within the manosphere, such as Incels.

Intensification and Escalation

This section focuses on findings from fringe platforms hosting MGTOW and incel communities. Here, exclusion or self-exclusion from mainstream society seemingly produces heightened emotional norms. Forum members appear to construct social worlds where sexual and romantic access is governed by Darwinian dynamics. 'Aphas' or 'Chads' are depicted as the ultimate beneficiaries of this system. They are presented as a projection of everything that forum users are not; Chads/Aphas are attractive, dominant men who effortlessly secure attention from women. While MRA focuses less

on this social hierarchy, they still express frustration towards men who, in their view, uphold a gynocentric order for personal gain.

Managing Intimacy

This section explores how PUA and MGTOW communities develop emotional strategies to manage intimacy, rejection, and identity amid perceived male decline. PUAs frame intimacy as a learnable skillset, promising control and sexual success while reframing rejection as a challenge to be mastered. MGTOW, by contrast, frames heterosexual relationships as inherently dangerous. Women, feminists, and legal institutions are positioned as threats, generating persistent anxiety around interaction. Avoidance of women, relationships, and marriage is promoted as a means of regaining control and preventing inevitable harm.

Failed masculinity

Incels appear to construct their entire identities around the internalisation of failing to obtain this ideal masculine status. Their world is structured by the belief that they are genetically inferior, have poor bone structure, are short, and or are not white. This is further emphasised through their anecdotes of social ostracism and being mocked, ridiculed, and sometimes physically assaulted by outsiders. Shame is apparent within the forums and seemingly is key in constructing the Incel masculine identity.

Future Work

There are several potential avenues for future work in this area. Firstly, while misogyny was a predominant aspect of these forums, it was not the only intense form of prejudice. Alongside misogyny, there were also numerous instances of blatant racism and anti-Semitism, which warrant further investigation. Secondly, future research could explore how these forum users become involved with these platforms. Although distinct differences were noted between misogynistic forums on mainstream and fringe social media platforms, it would be valuable to explore how and why men seek out these increasingly extreme forums. Finally, focusing on the technical features of online platforms, future research could examine how platform affordances, moderation practices, and visibility shape emotional expression and radicalisation. Overall, this work aims to contribute to the growing academic discourse on gender norms in online environments, particularly regarding the rise of extreme misogyny in digital spaces.

References

- Brinkmann, S. (2014) 'Doing without data', *Qualitative Inquiry*, 20(6), pp. 720–725. doi:10.1177/1077800414530254.
- Ging, D. (2017) 'Alphas, Betas, and Incels: Theorizing the Masculinities of the Manosphere'. *Men and Masculinities*, 22(4), p.638-657.

- Hoffman, B., Ware, J. and Shapiro, E., 2020. Assessing the Threat of Incel Violence. *Studies in Conflict & Terrorism*, 43(7), pp.565-587.
- Langer, R. and Beckman, S.C. (2005). Sensitive research topics: netnography revisited. *Qualitative Market Research: An International Journal*, 8(2), pp.189–203
- Laskovtsov, A. (2020) 'Navigating the Manosphere: An Examination of the Incel Movements Attitudes of Sexual Aggression and Violence Against Women': Online Thesis and Dissertations.
- Rafail, P. and Freitas, I. (2019) 'Grievance articulation and community reactions in the men's rights movement online', *social media + Society*, 5(2), p. 1-11
doi:10.1177/2056305119841387.
- Smilde, D. and Hanson, R. (2018) 'Studying gender and sexualities with qualitative methods', *Qualitative Sociology*, 41(3), pp. 333–335. doi:10.1007/s11133-018-9395-x.
- Sparks, B., Zidenberg, A.M. and Olver, M.E. (2022) 'Involuntary celibacy: A review of incel ideology and experiences with dating, rejection, and associated mental health and emotional sequelae', *Current Psychiatry Reports*, 24(12), pp. 731–740.
doi:10.1007/s11920-022-01382-9.
- Sugiura, L. (2021). *INCEL REBELLION: the rise of the manosphere and the virtual war against women*. S.L.: Emerald Group Publ.
- Timmermans, S. and Tavory, I. (2012) 'Theory construction in qualitative research', *Sociological Theory*, 30(3), pp. 167–186. doi:10.1177/0735275112457914.

Session 6E - Affordances

Day 2 · Tues 14 July · 9:30–11:00, Studio 2, ARC

6E.1 “What Should I Wear?” Facebook Groups as Sociotechnical Infrastructures for Embodied Rhetorical Agency

Najia Nazir (Texas Tech University)

Introduction

Saudi Arabia’s Vision 2030 positions tourism and global talent mobility as strategic pillars, aiming to attract on the order of 150 million visitors annually by 2030 while expanding a diversified, international workforce. Recent government reports also show women’s labor-force participation surpassing the original 30% Vision benchmark, reflecting rapid changes in public-facing norms and workplaces. For newcomers, this policy environment turns everyday dress into a live question of situated social competence. Private expat Facebook groups function as sociotechnical infrastructures where members test “modest-enough” choices across micro-locations (street, mall, compound, office, mosque) and rehearse embodied etiquette before entry. These threads do more than exchange tips; they function as rehearsal spaces where respect, safety, and identity are coordinated in advance of offline encounters, with the group acting as a sociotechnical infrastructure for learning local norms. We analyze how platform affordances, such as privacy, threading, persistence, and moderation, shape that learning, and how inquiry tone and guidance shift around the recent relaxation of abaya-only expectations for visitors and, increasingly, in everyday practice.”). This study treats these inquiry threads as embodied intercultural communication and shows how rhetorical agency is shaped by platform affordances and circulating emotions.

Objectives/Research Questions

As Vision 2030 aims to boost tourism and grow the international workforce, and with noted improvements in women's economic involvement, expat Facebook groups have become useful tools for understanding local dress codes, which we explore through the features of the platform and changes in policies. Hence, this study asks:

RQ1: How do inquiry posts spatialize dress rules across micro-locations, and which platform features enable that mapping?

RQ2: Which facework markers (hedging, humility, respect) co-occur with reassurance responses and practical guidance?

RQ3: How do participants articulate identity calibration (what counts as “modest-enough”) across settings, and how is this linked to affordances?

RQ4: (policy segment): Did the recent relaxation of abaya-only expectations coincide with measurable shifts in inquiry volume, tone, and content?

Methods

This study employs descriptive, exploratory netnography to investigate private, women-focused expatriate Facebook groups as sociotechnical systems conducive to learning about dress norms (Kozinets, 2020). The analysis centers on inquiry posts and their comment threads, utilizing a dataset of 162 posts from five active, well-moderated groups, gathered up until January 10, 2026. Posts were selected based on their explicit requests for context-specific dress advice, with a focus on various settings like workplaces and mosques, and were identified through commonly used terms related to modest apparel.

The research examines platform affordances as analytical variables, documenting elements such as thread organization, references to prior discussions, community privacy rules, and moderation practices. This approach connects rhetorical practices to specific platform features, allowing for the assessment of variables like response times and the inclusion of policy-linked resources in replies.

A coding scheme aligns with the theoretical framework, incorporating elements such as situational modulation, affective deference, and identity calibration. This entails careful tracking of various factors including clothing references and the nature of replies from participants, supported by reliability measures ensuring the integrity of data analysis.

A significant change in national dress policies during the study period provided a natural experimental backdrop. The corpus was divided into segments based on pre- and post-policy change, assessing various dimensions including inquiry frequency, tone shifts, and response helpfulness through both descriptive averages and lightweight modeling techniques.

Ethical considerations were paramount, treating Facebook groups as semi-private spaces with participant expectations for audience and context. The study adhered to data minimization principles and incorporated a protocol grounded in Contextual Integrity (Nissenbaum, 2004), avoiding the storage of identifiable information and minimizing risks of reverse lookup. Documentation of governance practices for each group and respect for community guidelines governed the sharing of research findings (Markham, 2012). Public materials are designed to be anonymized, while a sealed appendix allows for independent verification without compromising participant

identities (Lincoln & Guba, 1985). Throughout the analysis, emphasis was placed on the discourse of the platform and situational learning, rather than on individual member behaviors (Zimmer, 2010).

Results

Situational modulation: the thread as a map.

Across groups, guidance is consistently place-bound: the same outfit is framed as appropriate in one setting and not another (street vs. mall vs. compound vs. workplace vs. religious sites). Threading keeps these micro-location truths from collapsing into a single rule, effectively producing a geosemiotic “map” newcomers can read in advance (Scollon & Scollon, 2003). This pattern aligns with Discourses in Place and Hall’s “silent language,” now made legible by comment nesting in a semi-private setting (Hall, 1976).

Affective deference: the currency of inquiry.

Inquiry posts are routinely couched in humility and respect (“not wanting to offend,” “happy to respect local ways”), which reliably draw reassurance and detailed, context-savvy responses. Users manage this anxiety through what Ting-Toomey (2019) describes as facework. This observed circulation of anxiety, care, and reassurance is consistent with Ahmed (2015)’s affective economies and with Face-Negotiation Theory on facework in intercultural exchanges. The platform’s asynchronous text affordances enable careful hedging that performs deference without loss of agency.

Identity calibration: finding a “modest-enough” self.

Participants iteratively test sleeves, hemline, layering, head coverings, and fit to balance respect with self-presentation, using the group as a backstage before stepping into specific spaces. This rehearsal logic is consistent with Goffman’s frontstage/backstage and Butler’s performance, and with my framing of clothing as a communicative mode within a multimodal ensemble (Butlet, 2006; Goffman, 1959).

Affordance effects: structure to practice to outcome.

Where threads exhibit visible nesting and users reference prior discussions, coexisting micro-location truths persist without conflict; in more private groups, deference markers appear stronger and helpful replies arrive faster; moderator presence and pinned guidance correlate with replies that include links to rules. Together these observations underscore that affordances do not merely host interaction, and while they condition it, this is an interpretation coherent with this platform-aware reading and with netnographic approaches to sociotechnical context (Kozinets, 2020; Markham, 2012).

Policy-shift signal: from single rule to contextual choice.

In the post-policy segment, threads feature fewer “abaya-only” answers and more option-framing (long cardigans, shawls, tailored pantsuits), with advice increasingly keyed to venue and audience (e.g., government offices vs. malls). This emerging distribution of “acceptable” solutions is visible in your captured exemplars and is theoretically consonant with a geosemiotic account of norms that vary by place.

References

- Ahmed, S. (2015). *The cultural politics of emotion* (2nd ed.). Routledge.
- Butler, J. (2006). *Gender trouble*. Routledge.
- Goffman, E. (1959). *The presentation of self in everyday life*. Doubleday.
- Hall, E. T. (1976). *Beyond culture*. Anchor.
- Kozinets, R. V. (2020). *Netnography* (3rd ed.). SAGE.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.
- Markham, A. (2012). Fabrication as ethical practice. *Information, Communication & Society*, 15(3), 334–353.
- Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review*, 79(1), 119–158.
- Scollon, R., & Scollon, S. (2003). *Discourses in place*. Routledge.

6E.2 The logistics of visibility: Reconfiguring illicit drug trade on Telegram

Piotr Siuda (Kazimierz Wielki University in Bydgoszcz)

Introduction

Research on online drug trade has largely been structured around two dominant analytical models. The first focuses on darknet cryptomarkets, where exchange is governed through formalized infrastructures of trust such as escrow systems (third-party payment mechanisms in which funds are held until the buyer confirms successful delivery), feedback mechanisms, and vendor reputations (e.g., Aldridge & Décary-Hétu, 2016; Martin, 2019; Siuda et al., 2024). The second model examines drug dealing on mainstream social media (e.g., Facebook, Instagram) emphasizing relational interaction and semiotic concealment as strategies for managing platform moderation and legal risk (e.g., Bakken & Demant, 2019; Bakken & Harder, 2023; Moyle et al., 2019). Despite their differences, both strands share an implicit assumption: that online drug trade is fundamentally relational, negotiated, and embedded in interpersonal trust work.

This paper challenges that assumption by examining Telegram as a platform where drug-related exchange operates through a markedly different socio-technical configuration. Telegram combines public broadcast channels, asymmetric communication, weak moderation, and rapid mechanisms of replication and migration (see Bal, 2023; Buehling & Heft, 2023; Rolando et al., 2023 for analyses of Telegram affordances). Unlike cryptomarkets, it lacks escrow, stable vendor identities, integrated, platform-native reputation systems, or durable feedback infrastructures; unlike mainstream social media, it does not primarily rely on interpersonal interaction or semiotic ambiguity. Yet drug-related communication on Telegram is often highly explicit, standardized, and persistent.

Rather than interpreting this explicitness as a lack of concealment or a deviation from established models, the article argues that Telegram reorganizes illicit exchange around what is conceptualized as a logistics of visibility. In this configuration, concealment is displaced from the level of content to the level of infrastructure. On Telegram, what persists is not the vendor as a recognizable actor, but the communicative flow itself: actors can disappear, be replaced, or migrate, while the stream of offers remains visible through rhythm, repetition, and infrastructural redundancy. Telegram thus functions not as a market or a community, but as an infrastructure for maintaining the visibility of illicit exchange without relational or reputational anchoring.

Objectives/Research Questions

The primary objective of this paper is to problematize relational and market-based theories of online drug trade by examining a platform where exchange is organized without sustained interaction, narrative reputation-building, or semiotic concealment of offers. Instead of asking how the drug trade operates on Telegram, the paper asks what Telegram reveals about the limits of existing theoretical frameworks for studying illicit online exchange. The study addresses the following research questions:

RQ1: How does Telegram-based drug communication challenge relational models of online drug trade centered on negotiation, interaction, and interpersonal trust?

RQ2: How is visibility transformed from a risk to be managed into an infrastructural resource that sustains illicit exchange?

RQ3: What replaces reputation and trust mechanisms in platform environments lacking formal market governance?

RQ4: How do rhythmic posting, standardization of communicative formats, and infrastructural redundancy function as alternative modes of coordination and stability?

By shifting analytical attention from actors and transactions to communication streams and platform affordances, the article reconceptualizes online drug trade as a form of platform-mediated visibility management. Conceptually, it introduces the notion of temporal trust: a mode of credibility produced through continuity, regularity, and infrastructural persistence rather than through feedback, reputation, or relational exchange.

Methods

The study draws on a corpus of publicly accessible, Polish-language Telegram posts related to drug distribution. The unit of analysis is the individual post, treated as a communicative artifact rather than as evidence of a completed transaction. Data collection followed a multi-stage procedure. An initial set of 70 Telegram groups advertising drugs in the Polish context was identified through manual search. Using the Telethon library (Telethon's Documentation — Telethon 1.42.0 Documentation, n.d.), messages and metadata were scraped from these groups, while simultaneously collecting links to additional drug-related channels referenced within them. This snowball sampling strategy expanded the dataset to approximately 1,500 Telegram channels, including both open and closed groups.

From each channel, up to the most recent 1,000 messages were collected (or fewer where historical data were limited). Data collection was conducted once, during the first week of January 2025, yielding an initial corpus of approximately 500,000 messages, with the earliest retained posts dating back to 2021. Subsequent preprocessing involved removing duplicates, empty or image-only messages, and masking sensitive identifiers (cf. Harviainen et al., 2021). This resulted in a cleaned

dataset suitable for analyzing large-scale patterns of communicative visibility on Telegram.

Analytically, the study employs a mixed-methods computational approach to examine how communicative visibility is structured and sustained within Telegram-based drug channels. First, posts were classified into recurring communicative formats—such as sales offers, logistics information, contact redirections, so-called “legit checks” (see Results), and migration or backup announcements—using rule-based text classification supported by manual validation. Second, temporal metrics were calculated at the channel level, including posting frequency, inter-post intervals, and measures of regularity. Third, inequality measures, including Pareto distributions and Gini coefficients, were used to examine the distribution of communicative visibility across channels. Rather than indicating economic inequality, these measures captured the concentration of communication, revealing whether visibility is broadly shared or dominated by a small number of highly active channels. Finally, substance-related content was identified using a curated lexicon of drug names and slang terms, enabling an analysis of diversification and specialization across channels. Together, these steps provide an integrated account of how Telegram organizes illicit exchange through standardized formats, temporal rhythms, and infrastructural concentration.

Results

The analysis demonstrates that drug-related communication on Telegram is dominated by standardized, template-like formats characterized by brevity, functional markers (e.g., substance, price, contact, delivery), and minimal narrative elaboration (RQ1). Unlike darknet marketplaces, where explicitness is embedded within reputational and market governance structures, Telegram communication decouples explicit offers from durable identities, interaction, and feedback mechanisms, thereby challenging relational models of online drug trade. While externalized forms of validation do exist—most notably so-called “legit check” channels, where buyers briefly confirm successful transactions—these practices remain weakly institutionalized, highly fragmented, and functionally instrumental (RQ1, RQ3).

Rather than fostering relational continuity or community-based trust, legit checks operate as peripheral verification devices, separated from the channels where transactions occur. As such, they support transactional reliability without producing durable social ties, shared norms, or reputational capital comparable to darknet marketplaces. Visibility, credibility, and verification are distributed across loosely connected Telegram channels rather than consolidated within a single market architecture. Trust is not accumulated but momentarily signaled, reinforcing Telegram’s role not as a market or community, but as a flexible logistical infrastructure for coordinating visibility, contact, and delivery (RQ1, RQ3).

In addition, communicative visibility is unevenly distributed: a small subset of channels accounts for a disproportionate share of all posts, indicating a strongly centralized communicative ecology (RQ2). Crucially, channel prominence correlates more strongly with temporal regularity than with diversity of substance. Channels that maintain consistent posting rhythms occupy dominant positions within the visibility hierarchy, suggesting that continuity and predictability function as substitutes for reputation in platform environments lacking formal market governance (RQ3, RQ4).

Finally, migration-related communication—such as announcements of backup channels or platform movement—emerges as a persistent infrastructural layer rather than an exceptional response to disruption (RQ2, RQ4). These practices enable anonymity without reducing visibility, allowing actors to remain opaque and replaceable while communication streams persist. Taken together, the findings show that Telegram-based drug trade operates through a logistics of visibility in which coordination and stability are achieved via rhythm, redundancy, and format standardization rather than interaction or interpersonal trust narratives (RQ1–RQ4).

Future Work

Future research should further interrogate the implications of infrastructural visibility for theories of illicit exchange. The present study is limited by its focus on observable communication practices and platform-level metadata, which do not allow direct access to participants' intentions, interpretations, or off-platform interactions. In particular, the analysis cannot fully capture how users subjectively understand trust, risk, and legitimacy, nor how these meanings travel across interconnected Telegram channels and external verification spaces.

Combining computational mapping with qualitative digital ethnography and interview-based approaches could illuminate how participants interpret, negotiate, and strategically navigate these infrastructures, including peripheral mechanisms such as external feedback or verification channels. More broadly, the concept of logistics of visibility invites reconsideration of how illicit economies operate in platform environments where markets and communities give way to infrastructural coordination—raising new questions about governance, responsibility, and power in digitally mediated illicit exchange

Moreover, future studies could address the limitations of this study by conducting comparative research across platforms, examining how different affordances redistribute the balance among visibility, anonymity, and governance, and how infrastructural trust is stabilized—or destabilized—under varying technical and regulatory conditions. Longitudinal analyses may reveal how rhythmic, redundant, and template-based communication adapts to moderation pressure, platform shutdowns,

or legal intervention, capturing infrastructural change over time rather than static snapshots.

References

- Aldridge, J., & Décary-Hétu, D. (2016). Cryptomarkets and the future of illicit drug markets. In *The Internet and drug markets* (pp. 23–32).
<http://daviddhetu.openum.ca/files/sites/39/2018/12/Future.pdf>
- Bakken, S. A., & Demant, J. J. (2019). Sellers' risk perceptions in public and private social media drug markets. *International Journal of Drug Policy*, 73, 255–262.
<https://doi.org/10.1016/j.drugpo.2019.03.009>
- Bakken, S. A., & Harder, S. K. (2023). From dealing to influencing: Online marketing of cannabis on Instagram. *Crime, Media, Culture: An International Journal*, 19(1), 135–157. <https://doi.org/10.1177/17416590221081166>
- Bal, M. (2023). Audio-visual piracy on Telegram: A perspective on monetization models, pirate strategies and industrial pathways. *Contemporary South Asia*, 31(2), 311–325. <https://doi.org/10.1080/09584935.2023.2204220>
- Buehling, K., & Heft, A. (2023). Pandemic Protesters on Telegram: How Platform Affordances and Information Ecosystems Shape Digital Counterpublics. *Social Media + Society*. <https://doi.org/10.1177/20563051231199430>
- Harviainen, T. J., Haasio, A., Ruokolainen, T., Hassan, L., Siuda, P., & Hamari, J. (2021). Information Protection in Dark Web Drug Markets Research. In *Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS) / ed. Tung X. Bui* (pp. 4673–4680). University of Hawaii at Manoa.
<http://hdl.handle.net/10125/71184>
- Martin, J. (2019). *Cryptomarkets: A Research Companion*. Emerald Publishing.
- Moyle, L., Childs, A., Coomber, R., & Barratt, M. J. (2019). #Drugsforsale: An exploration of the use of social media and encrypted messaging apps to supply and access drugs. *International Journal of Drug Policy*, 63, 101–110.
<https://doi.org/10.1016/j.drugpo.2018.08.005>
- Rolando, S., Arrighetti, G., Fornero, E., Farucci, O., & Beccaria, F. (2023). Telegram as a Space for Peer-Led Harm Reduction Communities and Netreach Interventions. *Contemporary Drug Problems*, 50(2), 190–201.
<https://doi.org/10.1177/00914509221145196>
- Siuda, P., Świeca, L., Cheba, P., & Shi, H. (2024). “Similar to Legal Companies”: Vendor Brands and Other Aspects of Marketing Communication in the Darknet Drug Trade. *Deviant Behavior*, 0(0), 1–16.
<https://doi.org/10.1080/01639625.2024.2428711>

Telethon's Documentation—Telethon 1.42.0 documentation. (n.d.). Retrieved January 25, 2026, from <https://docs.telethon.dev/en/stable/>

6E.3 Palestinian women's social media practices in London: platformed strategies of resistance, solidarity and political expression

Sevilay Cesurer (UAL)

Introduction

Drawing on the experiences of Palestinian women living in London, this paper examines how video-based social media platforms such as Instagram and TikTok are strategically used alongside offline practices to communicate migration trajectories, negotiate cultural identities, engage in political activism, and maintain forms of mediated co-presence (Baldassar et al., 2016) with social networks and unfolding events in Palestine. By foregrounding platformed strategies of resistance, the paper aims to contribute to debates in digital media, platform studies, migration, and feminist studies by highlighting both the possibilities and constraints that shape diasporic political engagement.

Against the backdrop of ongoing genocide and settler-colonial violence in Palestine—namely in Gaza, described as the “first livestreamed genocide in history” (Shaw, 2025)—social media platforms have (re)emerged as crucial spaces for witnessing and connecting (Baym, 2015). Through their distinctive affordances, these platforms enable the negotiation of identity and belonging, the performance of long-distance nationalism (Skrbiš, 2017), and alternative modes of artistic and political expression as gestures of solidarity.

Therefore, as the homeland continues to be shaped and re-imagined through discursive and visual regimes, the plethora of existing literature demystifies how these practices intensify during times of crisis, and among communities whose voices are systematically marginalised or excluded from mainstream media and public discourse (Aouragh, 2011; Godin and Dona, 2016; Tawil-Souri and Matar, 2024). However, focusing on the gendered, affective, and mundane forms of resistance enacted through everyday social media practices in diaspora is essential to understanding the dynamic relationships and tensions between platform affordances and ongoing negotiations of identity, belonging, and political engagement.

This paper seeks to unpack these platformed practices, highlighting how everyday acts of connection and solidarity can constitute forms of mediated resistance, even within and against algorithmic constraints.

Objectives/Research Questions

My doctoral research examines how Palestinian women living in London engage with social media platforms and offline spaces to mobilise politically, express solidarity, and sustain connections to their homeland. While the broader project and sub-research questions explore how practices of resistance are shaped, mediated, and transformed through gendered and diasporic experiences, as well as the interweaving of cultural, artistic, and affective practices across online and offline contexts, this paper focuses specifically on Sub-research Question 1, which asks:

How do platform-specific affordances assist or constrain Palestinian women's modes of mobilisation, solidarity, and political expression related to the homeland?

Accordingly, the paper foregrounds the role of platform-specific affordances in shaping these practices. The objective is not only to identify the possibilities these platforms offer but also to critically examine their constraints, including issues related to visibility, circulation, the digital labour of becoming micro-celebrities, undertaking both micro and macro activism, and other challenges as experienced by Palestinian women in diaspora.

Methods

This paper draws on a doctoral research project currently in progress, with fieldwork commenced officially in December 2025. The study adopts a qualitative approach, and data collection comprises two interrelated components. First, semi-structured interviews are being conducted with Palestinian women to explore their experiences of political mobilisation, solidarity practices, and engagements with Palestine across both online and offline spaces. Second, digital ethnography (Pink et al., 2015) is being employed to observe and analyse social media practices, focusing on how they post, share, consume, and interact with Palestine-related content across two different platforms. The paper explores how practices and expressions of grief, care, and the politics of solidarity circulate through platform-specific affordances as these women continue to develop their own situated strategies for posting or consuming content within digitally mediated environments.

Results

Preliminary analysis of interview data identifies several interrelated themes illustrating how platform-specific affordances shape Palestinian women's political expression and solidarity practices in everyday life, with some of them being:

Cross-border connectivity and situated visibility: Participants emphasise that social media enables them to reach both local and transnational audiences, allowing them to maintain connections with diaspora communities as well as with the homeland. At the same time, visibility is understood as shaped by platform algorithms and moderation practices.

Curating the self, curating the message: Participants describe careful strategies of self- and content-curation through features such as story highlights, private groups, co-posting, and selective posting. These practices are often used to manage visibility, protect themselves from harassment, and mitigate risks of shadowbanning or reduced reach. Importantly, curation here is not only aesthetic but political, enabling polysemous performances (Papacharissi, 2011).

Everyday algorithmic negotiation: A key insight concerns how participants intentionally intersperse political content with mundane or intimate posts, such as selfies, pets, daily routines, or light-hearted content, to “break the algorithm”. This alternation between Palestine-related posts and everyday life content reflects an embodied, tactical engagement with platform affordances.

Emotional and political labour: Using social media for solidarity requires navigating emotional exposure, political pressure, and potential backlash. Platform affordances both enable and constrain these experiences, shaping how women negotiate identity and activism online.

These early findings reveal the complex interplay between technology, political engagement, and affective labour, suggesting, in a way, that understanding platform affordances is crucial for appreciating digital activism among marginalised communities, countering the taken-for-granted techno-utopian narratives often associated with these technologies and platforms as giving voice by default (Tazzioli, 2023).

Future Work

Building on these preliminary insights, the next phase of research will extend data collection through digital ethnography on Instagram and TikTok. Data will be used along with interviews to provide a nuanced understanding of how platform affordances shape engagement, content creation, consumption, and visibility over time, framing the online and offline not as separate nexuses but as a continuous, mutually constituting continuum.

References

- Aouragh, M. (2011) *Palestine Online: Transnationalism, the Internet and Construction of Identity*. I. B. Tauris.
- Baldassar, L., Nedelcu, M., Merla, L. and Wilding, R. (2016) ‘ICT-Based Co-presence in Transnational Families and Communities: Challenging the Premise of Face-to-Face Proximity in Sustaining Relationships’, *Global Networks*, 16(2), pp. 133-144.
- Baym, N. K. (2015) ‘Connect With Your Audience! The Relational Labor of Connection’, *The Communication Review*, 18(1), pp. 14-22.

- Godin, M. and Dona, G. (2016) ““Refugee Voices,” New Social Media and Politics of Representation: Young Congolese in the Diaspora and Beyond’, *Refuge: Canada’s Journal on Refugees*, 32(1), pp. 60-71.
- Papacharissi, Z. (2011) *A Networked Self: Identity, Community, and Culture on Social Network Sites*. New York, Abingdon, Routledge.
- Pink, S., Horst, H., Postill, J., Hjorth, L., Lewis, T. and Tacchi, J. (2015) *Digital Ethnography: Principles and Practice*. SAGE.
- Shaw, M. (2025) ‘The Genocide that Changed the World’, *Journal of Genocide Research*, pp. 1-15.
- Skrbis, Z. (2017) *Long-Distance Nationalism Diasporas, Homelands and Identities*. Routledge.
- Tawil-Souri, H. and Matar, D. (2024) *Producing Palestine: The Creative Production of Palestine Through Contemporary Media*. I. B. Tauris.
- Tazzioli, M. (2023) ‘Counter-mapping the techno-hype in migration research’, *Mobilities*, 18(6), pp. 920-935.

6E.4 Undoing Dravidian Feminism: How Tamil Women Influencers Are Re-Traditionalizing Gender Through Digital Culture

Sinthu Bhairavi Govindasamy Vinoba (Oktopost Technologies) Neelambari Govindasamy Vinoba (Community College, Central University of Tamil Nadu)

Introduction

Tamil Nadu, a South Indian state of 78 million, represents a political anomaly in contemporary India. For nearly a century, Dravidian parties rooted in E.V. Ramasamy "Periyar's" anti-Brahmin Self-Respect Movement have governed the state, promoting rationalism, women's autonomy, and resistance to Hindu nationalism. This legacy produced India's highest women's workforce participation and strong electoral resistance to the BJP. Dravidian ideology explicitly rejected Brahminical Hindu festivals like Varalakshmi Vratam (women fasting for husbands' welfare) as patriarchal, instead promoting egalitarian Tamil agrarian festivals like Pongal and Aadi Perukku.

Yet on Instagram and YouTube, young Tamil women influencers are becoming agents of re-traditionalization. A growing cohort performs "trad wife" aesthetics through elaborate homemaking content and devotional practices, amplifying the very Brahminical festivals (Varalakshmi Vratam, Navarathri, Gokulashtami) that Dravidian rationalism rejected. Meanwhile, Dravidian festivals receive minimal elaborate treatment. These influencers redefine "Tamil culture" such that it closely aligns with Hindu Brahminical practice, while distancing it from the Dravidian discourse that positioned Tamil identity as resistant to Hindu nationalism.

This paper argues that the Tamil trad wife phenomenon represents a gendered digital counter-reformation using social media affordances (visibility, aestheticisation, virality) and postfeminist discourse (choice, empowerment) to re-traditionalize Tamil women's identity in alignment with Hindutva's cultural project. If re-traditionalization succeeds in Tamil Nadu with its century-long feminist legacy, what does this reveal about digital culture's capacity to erode progressive movements globally?

Objectives/Research Questions

Drawing on postfeminist media studies (Gill, 2007; McRobbie, 2009), scholarship on digital Hindutva (Udupa, 2019; Chopra, 2020), and Sanskritization theory (Srinivas, 1956), and Gibson's (1979) theory of affordances, this study examines how the affordances of digital platforms encourage Tamil women influencers to repackage Brahminical rituals as an empowered cultural choice.

Research Questions:

RQ1: How do Tamil women influencers use social media affordances to aestheticise and "modernise" Brahminical religious practices within a historically Dravidian/rationalist political landscape?

RQ2: How does platform choice (Instagram vs. YouTube) mirror class/caste stratifications, and how do platform-specific affordances shape performances of "tradition"?

RQ3: To what extent does "apolitical" lifestyle content discursively align with soft-Hindutva propaganda and Hindu cultural nationalism's broader project?

We examine multiple overlapping mechanisms: (1) soft Hindutva advancing Hindu nationalism through "apolitical" content; (2) digital Sanskritization spreading upper-caste Brahmin practices across class/caste lines through aspirational influencer culture; (3) postfeminist sensibility framing patriarchal rituals as empowering "choice"; (4) algorithmic amplification rewarding visually spectacular religious content; (5) platform capitalism brands preferentially sponsoring Brahminical festivals; and (6) cultural amnesia younger generations encountering Brahminical practices as "authentic Tamil culture" without knowledge of Dravidian critique.

Methods

Approximately 80-100 Tamil women lifestyle influencers were identified through systematic hashtag search (#TamilCulture, #Pongal, #Navarathri), location tags, and snowball sampling. Accounts were screened for Tamil self-identification, lifestyle/cultural focus, minimum 10K followers (or 1K+ with $\geq 3\%$ engagement), monetization evidence, active status, and ≥ 12 -month operation. A preliminary 3-month content audit calculated Brahminical content percentage and identified temporal patterns, classifying accounts into two performance types: Type A (festive/occasionalist) characterized by $< 30\%$ baseline Brahminical content with festival spikes $\geq 3\times$ baseline, and Type B (ideological/consistency) maintaining $\geq 40\%$ year-round Brahminical content with low variance. The final sample ($n=27$) was stratified by typology (Type A $n=15$, Type B $n=9$, Hybrid $n=3$), platform (Instagram $n=16$, YouTube $n=8$, dual-platform $n=3$), follower tier, and geography (urban Tamil Nadu, tier-2/3 cities, diaspora). Extended validation examined 6-12 months of content to ensure pattern stability across multiple festival cycles.

Public posts spanning 12 months were archived using platform-compliant tools, with video transcripts manually verified for Tamil/Sanskrit accuracy. Accounts under 50K followers were pseudonymized following AoIR ethical guidelines.

Caste/class positioning was inferred through convergent self-presentation markers: linguistic patterns (Brahmin Tamil dialect features, Sanskrit density), food content (vegetarian advocacy, Brahmin-specific festival foods), ritual practices, dress markers (9-yard saree, specific jewelry), and surnames when disclosed. Two independent

coders established inter-rater reliability ($\kappa=0.82$, $n=15$ accounts). We analyzed performed identity rather than genealogical status, with class markers including production quality, domestic settings, consumption displays, and brand partnerships.

Ongoing analysis addresses our research questions through: (RQ1) A composite elaboration index (1-4 scale) scoring posts on length, production quality, posting frequency, emotional intensity, ritual completeness, and aesthetic labor, comparing Brahminical festivals (Diwali, Navarathri, Varalakshmi Vratam) with Dravidian festivals (Pongal, Tamil New Year) using Mann-Whitney U tests. (RQ2) Systematic platform comparison of Instagram versus YouTube content across production values, language ratios, caste markers, and brand partnerships to analyze how platform-specific affordances shape caste/class performance. (RQ3) Code-switching analysis across Tamil/English/Sanskrit with stratified comment coding ($n=200$) for affirmation, learning, nostalgia, critique, community-building, and aspirational language; comprehensive commercial analysis identifying all sponsored content and analyzing frequency, brand categories, timing, and festival alignment patterns.

Results

Preliminary observations suggest Brahminical Hindu festivals exert disproportionate influence on Tamil women influencers' content through superior aesthetic affordances, algorithmic rewards, and commercial opportunities, facilitating digital Sanskritization while rendering Dravidian cultural practices invisible.

Brahminical festivals (Varalakshmi Vratam, Navarathri, Gokulashtami) offer high visual spectacle, elaborate kolams, multi-tiered golu displays, silk sarees, temple jewellery, and coordinated presentations. These provide "aesthetic density": multiple photographable elements translating effectively to Instagram's visual economy and YouTube's tutorial format. Dravidian festivals (Pongal, Aadi Perukku) centre on simpler agrarian practices with lower visual complexity and fewer extended content opportunities.

This aesthetic differential creates algorithmic advantages. Visually elaborate Brahminical content generates higher engagement through shareability and tutorial potential. Preliminary observations suggest these posts receive more comments, shares, and saves - metrics platforms reward with increased visibility. This creates feedback loops: Brahminical content performs better algorithmically, incentivising greater production effort, further amplifying reach and attracting more creators.

Early commercial analysis indicates Brahminical festivals attract more brand partnerships. Elaborate aesthetic requirements, silk sarees, jewellery, puja supplies, premium cookware, and decor create natural sponsorship opportunities. Brands recognize these as monetizable cultural moments, aligning with aspirational

consumption. Dravidian festivals, requiring fewer material inputs and emphasizing egalitarian simplicity, offer limited commercial integration.

These combined factors: aesthetic potential, algorithmic reward, and commercial opportunity appear to drive gradual cultural adaptation. Influencers, including non-Brahmins, adopt Brahminical practices because they are platform-optimised. What begins as strategic content creation (choosing high-performing festivals) evolves into actual cultural practice (observing festivals, teaching children, and family integration). Platform structural biases thus facilitate Sanskritization: upper-caste practices spread as aspirational and authentically "Tamil" while Dravidian alternatives become invisible.

For younger audiences encountering Tamil culture through social media, Brahminical aesthetics become synonymous with "Tamil tradition" while Dravidian alternatives vanish.

Platform architecture itself, not just user preferences, drives re-traditionalization. Digital platforms structurally favor religious spectacle, aesthetic complexity, and consumption-intensive practices, making Brahminical Hinduism more viable than Dravidian rationalism in the attention economy. This represents platform-mediated Sanskritization.

Future Work

Analytical completion will finalize quantitative statistical testing of elaboration differentials with effect size reporting, complete systematic linguistic coding achieving full inter-rater reliability across the corpus, and comprehensive commercial categorization of brand types and sponsorship strategies. Definitive platform stratification analysis will test class/caste reproduction hypotheses.

This research makes several theoretical contributions. First, it extends postfeminist media studies beyond Western contexts, demonstrating how "empowerment feminism" can facilitate rather than resist religious conservatism in regions with strong feminist legacies. Second, it advances digital nationalism scholarship by revealing how Hindu cultural nationalism operates through seemingly apolitical lifestyle content, showing crucial regional variation within "Indian social media." Third, the comparative festival analysis methodology offers a replicable technique for revealing selective cultural production and ideological erasure in digital contexts.

Key preliminary findings demonstrate that platform capitalism structurally incentivizes Brahminical content through differential sponsorship patterns, revealing commercial interests' alignment with religious nationalist projects. Platform stratification patterns suggest Instagram and YouTube reproduce offline caste/class hierarchies digitally. Triglossic linguistic performance reveals sophisticated ideological work erasing Dravidian politics through strategic language deployment.

The Tamil Nadu case holds particular significance: if digital re-traditionalization succeeds in India's most rationalist state, with century-long resistance to Hindu nationalism and active Dravidian governance, it demonstrates platform culture's formidable capacity to erode progressive movements globally. The commercial dimension reveals this is not organic cultural revival but a monetized project where platform capitalism profits from women's unpaid ritual labor while advancing conservative gender politics.

Future extensions beyond this study could examine: (1) audience demographics and reception patterns through focus groups to understand consumption and political consciousness formation, (2) longitudinal analysis tracking correlations with electoral shifts or declining Dravidian party support, (3) comparative analysis with other Indian states where progressive legacies are weaker, (4) creator motivations through in-depth interviews exploring political consciousness and economic pressures, and (5) counter-publics analysis identifying whether Dravidian-rationalist content exists but suffers algorithmic marginalization.

Limitations: Our sample focuses on visible, successful influencers. Caste identification through performance markers has inherent limitations. Causality between platform economics and content choices requires future creator interviews to establish definitively.

References

- Banet-Weiser, S. (2018). *Empowered: Popular feminism and popular misogyny*. Duke University Press.
- Chopra, R. (2020). Hindutva and technology: Digital networking in the making of a Hindu cultural nationalist public. *Social Identities*, 26(2), 187–203.
- D, M., Bagilesh, K., & Shanmugam, K. (2025). Social media as a catalyst: How Instagram is enabling Indian housewives to become digital entrepreneurs. *Journal of Information Technology Teaching Cases*, <https://doi.org/10.1177/20438869251363995>
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.
- Gill, R. (2007). Postfeminist media culture: Elements of a sensibility. *European Journal of Cultural Studies*, 10(2), 147–166.
- McRobbie, A. (2009). *The aftermath of feminism: Gender, culture and social change*. Sage.
- Srinivas, M. N. (1956). A note on Sanskritization and Westernization. *The Far Eastern Quarterly*, 15(4), 481–496.

Udupa, S. (2019). Nationalism in the digital age: Fun as a metapractice of extreme speech. *International Journal of Communication*, 13, 3143–3163.

Session 6F - Power, Platforms, and Politics

Day 2 · Tues 14 July · 9:30–11:00, 225, ARC

6F.1 MAGA on TikTok: Issue Mapping a Political Network

Gil Sharon (University of Leeds)

Introduction

The constantly evolving digital landscape is significantly reshaping our understanding of political movements, activism, and community formation. In my PhD project, I study Donald Trump's MAGA (Make America Great Again) movement on TikTok, specifically its use of the platform around the 2024 US presidential election.

Analysing the MAGA movement on TikTok presents several technical challenges. One key challenge is how to effectively collect data that accurately reflects the movement's actions on the platform. Another challenge is mapping the movement's online interactions and associations. Unlike Facebook or YouTube, TikTok does not provide the same level of community-building features, making it more difficult to trace a network or community on the platform. Previous studies have modelled connections within political networks using concepts such as followers, friends, hashtags and appearances (Lewis, 2018). These models have been successfully applied to their respective platforms. However, political activism on TikTok, particularly within the MAGA movement, does not follow the same logic.

In this paper, I present a new approach to fill this gap by modelling the MAGA network on TikTok through their shared issues and topics. This method is based on Rogers et al.'s (2015) Issue Mapping framework, grounded in Latour's Actor-Network-Theory, to identify a collective of actors based on the issues they discuss. The issue-mapping approach helps reveal the prominent actors pushing the MAGA agenda on TikTok and the issues they collectively discuss.

Objectives/Research Questions

Identifying the network actors

The issue-mapping approach effectively highlights the connections between actors through shared discussions. However, this framework does not specify TikTok's platform scenario. As such, the key question remains: how can we identify relevant users and determine the issues discussed in the videos? For example, Burgess and Matamoros-Fernandez (2016) propose a model for mapping a network of hashtag

associations to identify controversies on social media. However, hashtag sampling has many limitations, making it insufficient for the MAGA TikTok scenario, as many prominent MAGA users do not use hashtags in their videos.

To explore connections among TikTok users, I propose analysing their associations based on the issues they discuss in their videos. Unlike hashtag linkage, which only requires scraping metadata, identifying the issues discussed in each video requires manual coding. To collect this data, I employed Schellweld's (2021) digital ethnographic approach, which he argues is well-suited for researching TikTok due to the platform's "ephemerality," as videos and trends are often short-lived.

Ephemerality of TikTok videos

Tracing back past MAGA content can be problematic due to the ephemerality of far-right content, as platforms might remove or ban access to hate-speech videos (Muis et al., 2021). Moreover, the ephemerality of the platform poses a more profound concern in defining an online group. Venturini (2010) emphasises the difficulty of observing controversies due to the complexity of the active social scenario. He compares this ongoing configuration of social ties to magma, which is simultaneously liquid and solid. Hence, describing a social controversy forces an artificial freezing of an active actor-network formation. Rogers et al. (2015) elevate this complexity to the idea of understanding groups, as group-like formations are "in continual development" (p. 16). Groups are continually evolving, with members frequently changing, topics of discussion shifting, and even the group's identity transforming over time. This dynamic nature reflects the complexities of relationships and interactions that define a group's essence. Therefore, it is essential to select a specific timeframe and code the videos in real time.

Methods

I conducted the coding process from October 1 to November 30, 2024. The timeframe spans about one month before and one month after the November 5 presidential election, which was the peak of engagement in this community. I have created a spreadsheet to manually code the videos as I watch them on the platform. I watched TikTok content on the "following" and "for you" pages. Videos coded on the day of their release had their metadata (likes, comments, saves, and shares) re-recorded after a week for better accuracy.

The coding process has concluded, with 3,492 videos totalling 484 million likes across 239 accounts. The coding process was designed to identify the issues raised by each video. The videos were categorised into 104 issues.

Once coded, I have used a social network analysis (SNA) to visualise the actor-network. The network was based on the issue associations between actors.

Thus, in the network, each node represents an actor, and each edge (connection) represents the number of issues shared between two given actors. The number of issues shared by two actors determined the weight of the linkage. Consequently, the more issues discussed together by associated actors, the stronger the link between them. Similarly, a second SNA was developed to connect the issues based on the actors.

As a result, the issue-mapping approach facilitated the creation of two visual representations of the MAGA network on TikTok. These visual maps illustrate the positions of actors and the issues frequently discussed within the MAGA movement on TikTok.

Results

Discussion

This mapping and data-collection model enables insights into the MAGA movement's actions on TikTok around the presidential election. First, the intra-dynamics of MAGA's network on TikTok link grassroots influencers, official political campaigners, and media outlets through a hybrid connective action that can be described as non-hierarchical yet top-down influence. This formation contributes to the discussion of the MAGA coalition structure as outlined by Bennett and Livingston (2025). Second, the issues illuminate the narratives that the MAGA network socially constructed around the election. These narratives share a common sense of victimhood based on being Trump supporters.

Conclusion and future work

To summarise, the issue-mapping approach helps reveal the MAGA actor-network on TikTok. The issues discussed collectively by MAGA actors on TikTok demonstrate the formation of a cohesive community of shared ideas through connective actions. Visualising the MAGA network on TikTok helps reveal the issues and actors in the centre of the movement, those in the periphery and clusters. Additionally, this model can enhance our understanding of political communities on TikTok and potentially other short-form video platforms.

Future Work

The next step in my PhD project is to conduct an in-depth analysis to contextualise the data and illuminate the movement's use of TikTok affordances to advance their messages.

References

- Bennett WL, Livingston S. (2025). Platforms, Politics, and the Crisis of Democracy: Connective Action and the Rise of Illiberalism. *Perspectives on Politics*. Published online 2025:1-20. doi:10.1017/S1537592724002123
- Burgess, J., & Matamoros-Fernández, A. (2016). Mapping sociocultural controversies across digital media platforms: one week of #gamergate on Twitter, YouTube, and Tumblr. *Communication Research and Practice*, 2(1), 79-96.
<https://doi.org/10.1080/22041451.2016.1155338>
- Lewis, R. (2018). *Alternative Influence: Broadcasting the Reactionary Right on YouTube*.
<https://datasociety.net/library/alternative-influence>
- Muis, J., Klein, O., & Dijkstra, G. (2021). Challenges and opportunities of social media research: Using Twitter and Facebook to investigate far right discourses. In J. B. Stephen D. Ashe, Graham Macklin and Aaron Winter (Ed.), *Researching The Far Right: Theory, Method and Practice* (Vol. 147-163). Routledge.
- Rogers, R., Sánchez-Querubín, N., & Kil, A. (2015). *Issue Mapping for an Ageing Europe*. Amsterdam University Press.
- Schellewald, A. (2021). Communicative Forms on TikTok: Perspectives From Digital Ethnography (Vol. 15) [TikTok, short-video, forms of communication, digital ethnography, digital culture]. <https://ijoc.org/index.php/ijoc/article/view/16414>
- Venturini, T. (2010). Diving in magma: how to explore controversies with actor-network theory. *Public Understanding of Science*, 19(3), 258-273.
<https://doi.org/10.1177/0963662509102694>

6F.2 From Encyclopedic Authority to Platformed Knowledge: Wikipedia, Social Media, and the Reconfiguration of Epistemic Power

Giota Alevizou (King's College London)

Introduction

Encyclopedias occupy a paradoxical position in contemporary digital culture. Long understood as stabilising reference points within media systems, they have increasingly become contested infrastructures embedded within platform ecosystems shaped by algorithmic governance, political polarisation, and automation. Nowhere is this tension more visible than in the case of Wikipedia, which functions simultaneously as a public knowledge commons, a site of ideological attack, and a foundational resource for commercial AI extraction.

This paper draws on work of encyclopedic media genealogies (Iliadis, 2023; Alevizou, 2026) and critical Wikimedia research (Jankowski et al, 2025; Vetter et al., 2025) to situate Wikipedia within broader transformations of social media, platform capitalism, and digital epistemology. It argues that encyclopedic media should not be understood as neutral repositories or tertiary information sources, but as symbolic and infrastructural sites where authority, power, and legitimacy are continuously negotiated. In the platform age, encyclopedic knowledge is no longer bounded by genre or medium; it circulates across search & answer engines, social media feeds, knowledge panels, and generative AI systems, often detached from the procedural practices that originally rendered it accountable.

Recent public controversies — from political attacks framing Wikipedia as ideologically biased to anxieties about its appropriation by large language models — expose a deeper epistemic struggle. At stake is not merely the credibility of a single platform, but the future of shared informational baselines in a fragmented digital public sphere. As encyclopedic content is increasingly abstracted, summarised, and recombined by opaque algorithmic systems, the deliberative, procedural, and collective dimensions of knowledge production risk being flattened into seamless outputs.

Against deterministic narratives of technological displacement, this paper positions Wikipedia as a living epistemic experiment: imperfect, contested, and increasingly political, yet vital for understanding how authority is reconfigured under conditions of platformisation and automation.

Objectives/Research Questions

The paper advances three interrelated objectives:

1. To theorise encyclopedias as platformed epistemic infrastructures, embedded within social media ecologies and computational knowledge systems.
2. To examine how epistemic authority is redistributed across symbolic, algorithmic, and participatory forms of power in the digital age.
3. To assess the implications of generative AI for procedural knowledge practices, epistemic justice, and public accountability.

The central research questions are:

- How has encyclopedic authority shifted from institutional expertise to collective, procedural, and now increasingly algorithmic forms?
- In what ways do social media dynamics and AI systems recondition the visibility, legitimacy, and reuse of encyclopedic knowledge?
- What is at stake for epistemic justice when commons-based knowledge infrastructures are absorbed into commercial platform ecosystems?

Methods

The paper adopts a qualitative, interpretive, and historically grounded approach, synthesising insights from media genealogy, platform studies, cultural political economy, and epistemology. Rather than presenting new empirical data, it draws on a comparative analysis of encyclopedic, spanning print encyclopedias (e.g. Britannica), early digital models (e.g. Encarta), Wikipedia, and contemporary semantic and AI-driven knowledge infrastructures.

Primary materials within the larger project have included platform governance documents and interviews, media discourse, and scholarly debates on algorithmic power, collective intelligence, and epistemic justice. These are analysed to trace how authority is operationalised through editorial norms, technical mediation, and infrastructural integration across platforms.

Methodologically, the paper treats encyclopedias as socio-technical systems rather than discrete media objects, situating them relationally within broader information ecologies that include search engines, social media platforms, and generative AI models. This allows for an analysis of how knowledge is not only represented but also routed, abstracted, and instrumentalised within contemporary platform capitalism.

Results

The analysis advances four core arguments.

First, encyclopedic media have transitioned from bounded repositories of codified knowledge into dynamic semantic infrastructures that underpin platformed knowledge economies. Wikipedia no longer merely represents knowledge for human readers; it functions as an infrastructural node whose content is continuously reused, recombined, and monetised across platforms.

Second, the shift from institutional to collective authority introduced by Wikipedia did not eliminate epistemic gatekeeping but reconfigured it procedurally. Authority now emerges through policies, citation norms, bots, and community governance, redistributing power while simultaneously stabilising particular epistemic values.

Third, the rise of generative AI introduces a new epistemic tension between collective intelligence and synthetic automation. While AI systems depend heavily on Wikipedia as training data, they obscure the platform's procedural epistemology — deliberation, contestation, provenance — replacing visible friction with fluent abstraction.

Finally, these developments raise urgent questions of epistemic justice and sovereignty. As encyclopedic labour becomes increasingly invisible within automated systems, the capacity to contest bias, omission, or misrepresentation is weakened, particularly when knowledge is mediated through black-boxed infrastructures. These findings position encyclopedias as critical sites for analysing how power, authority, and knowledge are being reshaped in the age of social media and AI.

Future Work

Future research will extend this framework by examining emerging models of knowledge governance, including data trusts, federated infrastructures, and human-in-the-loop AI systems inspired by commons-based epistemologies. It will also explore how encyclopedic platforms can inform critical AI literacy by foregrounding procedural transparency, pluralism, and contestation.

More broadly, the paper argues that the future of public knowledge depends not on technological inevitability but on political, cultural, and design choices about how authority, accountability, and collective intelligence are institutionalised within platformed environments.

References

- Alevizou, G. (2026). *The Web of knowledge: Encyclopedias and authority in the digital age*. Cambridge: POLITY PRESS.
- Iliadis, A. (2023) *Semantic media: mapping meaning on the internet*. Cambridge: Polity Press
- Jankowski, S., Ford, H., Iliadis, A., and Sidoti, F. (2025). Uniting and reigniting critical Wikimedia research. *Big Data & Society*, 12(3).
- Vetter, M. A., Jiang, J., & McDowell, Z. J. (2025). An endangered species: How LLMs threaten Wikipedia's sustainability. *AI & SOCIETY*, 40(6), 4309–4322.
<https://doi.org/10.1007/s00146-025-02199-9>

6F.3 Platform Nonmarket Strategies as Power Projection

Gordon Dimitrieff (Antwerp Management School)

Introduction

When regulatory or legal intervention threatens platform business arrangements, platforms rarely concede discretion over audience access and visibility directly. Instead, they adapt—reframing obligations, shifting costs, mobilizing intermediaries, and engaging multiple institutional arenas. Existing scholarship examines platform governance (Gillespie, 2018; Van Dijck et al., 2018), infrastructure (Plantin et al., 2016), cultural production (Nieborg & Poell, 2018), and political economy (Srnicsek, 2016), demonstrating how platform power operates across technical, market, and social domains. This paper extends that work by examining how platforms coordinate strategic responses across institutional arenas when that power is contested—how legislative engagement, regulatory participation, and judicial positioning operate as interdependent components of nonmarket strategy rather than as reactive, venue-specific adjustments. This paper traces platform nonmarket engagement during Canada's Online Streaming Act, examining coordination patterns across Parliament, the CRTC, and Federal Court that reveal mechanisms through which platforms defend algorithmic discretion.

Canada's Online Streaming Act (Bill C-11) provides a well-documented case for examining this coordination. The Act brings streaming platforms under Canadian broadcasting regulation, requiring them to support domestic content production and enhance discoverability of Canadian programming. This directly implicates both algorithmic recommendation systems and the revenue-sharing arrangements through which platforms allocate audience attention and economic value within their ecosystems. The policy process has unfolded across overlapping institutional timelines since 2021, with parliamentary debate, regulatory consultations, and prospective judicial review proceeding in parallel rather than in sequence. This concurrency makes platform coordination observable as it occurs: arguments, organizational positioning, and institutional engagement are documented across multiple venues as the regulatory process continues.

Objectives/Research Questions

The paper pursues two related objectives.

First, it seeks to examine platform nonmarket strategy as an analytically integrated domain of platform business organization, rather than as a set of episodic or venue-specific political activities. In doing so, the paper focuses on how engagement with

legislatures, regulators, and courts can operate as coordinated strategic practice when regulation targets core coordination infrastructures.

Second, it uses Canada's Online Streaming Act (Bill C-11) as an empirical case to examine how platforms organize nonmarket engagement across multiple institutional arenas, and what patterns become visible when these engagements are analyzed together rather than in isolation.

The paper is guided by the following research questions:

RQ1: How do platforms coordinate nonmarket engagement across legislative, regulatory, judicial, and public arenas when regulation targets algorithmic coordination infrastructures such as ranking and discoverability?

RQ2: What recurring patterns of argumentation, timing, and cross-venue alignment emerge when platform nonmarket strategy is examined across institutional arenas rather than within them?

Rather than testing hypotheses, the paper uses these questions to examine how nonmarket strategy operates as a business-relevant domain that becomes most visible during periods of regulatory contestation over platform coordination mechanisms.

Methods

The study uses qualitative analysis to examine how platforms engage legislative, regulatory, and judicial arenas during the development and ongoing implementation of Canada's Online Streaming Act (Bill C-11). The analysis focuses on patterned similarities across venues rather than on attributing intent or strategy to individual actors.

The empirical analysis draws on publicly available materials produced during the legislative development of Bill C-11 and its continuing regulatory and judicial consideration. Sources include parliamentary testimony and written submissions; CRTC consultation filings and hearing transcripts; judicial materials associated with Federal Court challenges, where available; platform public communications; and statements by organizations presenting themselves as representing creators or users.

To analyze these materials, the paper introduces and applies the Rent Networks Model, an original analytical framework that conceptualizes platform ecosystems as structured across three interdependent circles: a core of actors governing coordination infrastructures, a middle circle of market participants whose value creation depends on those infrastructures, and an outer circle of nonmarket institutions such as legislatures, regulators, and courts. The model's analytical utility derives from its structural-position logic, which links actors' positions within the network to the strategic pathways available to them across domains.

Within this framework, nonmarket strategy is examined by distinguishing vertical strategies that project influence from the infrastructural core into outer-circle institutions from horizontal strategies that coordinate actors within the same circle. This distinction enables assessment of whether platform engagement across legislative, regulatory, and judicial venues reflects coordinated defense of core infrastructural arrangements or isolated responses to venue-specific constraints.

The analysis compares how platforms frame their positions across parliamentary, regulatory, and judicial settings. Although arguments vary by institutional context—emphasizing innovation and creator opportunity in Parliament, technical feasibility in CRTC proceedings, and constitutional limits in judicial contexts—the same underlying object is defended across venues: platform discretion over algorithmic coordination, including attention allocation and revenue relations with complementors. By identifying this consistency beneath venue-specific variation, the analysis tracks functional invariance and treats patterned alignment across arenas as evidence of coordination without requiring proof of centralized intent.

Results

Preliminary findings point to three interrelated mechanisms through which platform nonmarket strategy operates when coordination infrastructures are politically problematized.

First, platforms engaged in parallel institutional coordination. Rather than addressing legislative, regulatory, and judicial arenas sequentially, platforms advanced arguments across multiple arenas concurrently, anticipating future stages of the policy process while earlier stages remained unresolved. This suggests that nonmarket engagement is organized around the interdependence of institutional outcomes rather than reactive movement between venues.

Second, platforms employed venue-adapted argumentation with functional invariance. Arguments varied systematically across arenas—emphasizing innovation and expression in Parliament, feasibility and proportionality in CRTC proceedings, and constitutional limits in judicial contexts—yet consistently opposed regulatory specification or oversight of algorithmic ranking and discoverability systems. What changed was justificatory language; what remained constant was the defended business arrangement: platform discretion over visibility infrastructure.

Third, the analysis identifies the use of intermediary organizations as direct projection mechanisms rather than as agents of collective action. In parliamentary hearings, at least one organization presenting itself as representing creators lacked an independent membership base, received direct funding from major platforms, and included platform employees in its governance. Under these conditions, the organization did not mediate between platform and nonmarket institutions but functioned as an organizational

extension of the core, transmitting platform positions into outer-circle arenas under the appearance of independent constituency advocacy.

Examined across nonmarket venues, platform responses to Bill C-11 exhibit a patterned consistency not apparent within any single one: although platforms advance substantively different arguments in parliamentary, regulatory, and judicial settings, these engagements consistently defend discretionary control over how attention is allocated within the multisided markets they operate. Taken together, these findings indicate that platform nonmarket strategy is organized less around resisting particular rules governing market competition than around maintaining control over the mechanisms that coordinate participation, visibility, and value capture across platform ecosystems.

Future Work

This study documents how platforms coordinate nonmarket engagement across legislative, regulatory, and judicial arenas during Bill C-11, revealing patterns of parallel institutional engagement, venue-adapted argumentation, and organizational intermediation. These findings raise a further analytical question: under what conditions does such coordination remain effective, and when does it begin to break down?

Addressing this question requires closer examination of how non-platform actors (such as creators, trade associations, and advocacy organizations) operate within platform-organized ecosystems. While this study observes organizational participation with platform-aligned positions, it does not assess how organizational structure, resource access, or structural position within the rent network shape the range of strategic options available to different actors. Comparative analysis of cases where non-platform actors influence regulatory outcomes would clarify how coordination capacity varies within platform ecosystems.

More broadly, the findings in this paper suggest that nonmarket strategy should be analyzed as an integral component of platform business models rather than as a reactive response to regulation. Understanding how firms manage institutional engagement across arenas is essential for explaining how infrastructural control over market coordination persists or changes under evolving governance conditions.

References

- Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. Yale University Press.
<https://doi.org/10.12987/9780300235029>

Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society*, 20(11), 4275-4292. <https://doi.org/10.1177/1461444818769694>

Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2016). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, 20(1), 293-310. <https://doi.org/10.1177/1461444816661553>

Srnicek, N. (2016). *Platform Capitalism*. Polity Press.

Van Dijck, J., Poell, T., & De Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford university press.

6F.4 The multimodal use of Instagram in Italian politics: How do candidates use verbal and non-verbal strategies during regional elections?

Umberto Famulari (Colorado State University)

Introduction

Instagram has become a social media platform that politicians, political parties and government officials employ strategically as a tool to reach potential voters in both election and non-election years, allowing candidates to tailor their messages using a combination of verbal and non-verbal strategies (Bast, 2021). Instagram enables users to post either still and moving images accompanied by text, giving politicians and their consultants the opportunity to develop and craft effective multimodal messages (Filimonov et al., 2016) to promote and enhance the brand of candidates and to increase the followers' engagement and participation with political content during campaigns (Peng, 2021).

While research on political communication via social media has extensively examined images and text separately (Boulianne & Larsson, 2023), fewer studies have addressed the significance of multimodality (Metz et al., 2020). Regarding Instagram, the bulk of the research has looked at the diverse use of visuals by politicians as a vehicle for self-representation (Grusell & Nord, 2023) and to appeal to voters (Navarro et al., 2023). A few publications have examined Instagram posts as multimodal channels that incorporate text and visuals (Bossetta & Schmøkel, 2023; Haßler et al., 2023; Geise et al., 2024) during elections in the United States and Germany.

Studies treat Instagram posts as cohesive multimodal messages without analyzing how verbal and non-verbal strategies interact or diverge. This study instead conceptualizes Instagram campaign posts as multimodal devices in which verbal and non-verbal strategies may or may not align. In other words, while Instagram posts function as single multimodal messages, their verbal and non-verbal components can convey different meanings, underscoring the need for empirical analysis of both modalities to identify their overlaps and discrepancies.

As most of the previous research has concentrated on national election campaigns, more scholarship is needed to investigate how Instagram is utilized during political campaigns at the local level. Responding to the recent attention to regional elections and social media (Pineda et al, 2025), a quantitative content analysis was employed to study how the main candidates for president during the 2024 Italian regional elections used verbal and non-verbal strategies on Instagram.

Objectives/Research Questions

First, the study identified the main verbal and non-verbal strategies used by political candidates in seven regional elections. Second, it quantitatively analyzed the frequency of these strategies, comparing the candidates leading the two main coalitions (center-left and center-right) in each election. Third, while existing research on social media, politics, and gender (McGregor et al., 2017; Geise et al., 2024) is valuable, more studies are needed to understand how gender influences the use of verbal and non-verbal strategies, particularly in contexts outside the United States. Fourth, responding to a growing call for studies on the emotional expressions of candidates on social media (Bossetta & Schmøkel, 2023), this study analyzed differences in the emotional facial expressions of candidates. Gender can affect how emotional reactions are portrayed, and this study contributes to the literature by exploring how local politicians' emotional expressions on Instagram differ by gender.

Lastly, incumbency can be a variable that influences the campaign of politicians and their tactics on social media (Auter & Fine, 2018), but it is not clear how incumbents and challengers adapt their strategies that are conveyed verbally and non-verbally and in a comparative regional political landscape. This study aims to be an original addition to the scholarship on political incumbency and social media.

Given the increasingly multimodal digital media environment of political campaigns, this study aims to advance the literature on social media and politics by examining the interplay between visual and textual content, offering practical insights for political campaigns, consultants, and candidates running for office.

RQ1. Which (a) verbal strategies and (b) non-verbal strategies do the candidates in the 2024 Italian regional elections use in their Instagram posts?

RQ2. How did the use of (a) verbal strategies and (b) non-verbal strategies in Instagram posts during the 2024 Italian regional elections differ by political leaning of the candidates?

RQ3. How did the use of (a) verbal strategies and (b) non-verbal strategies in Instagram posts during the 2024 Italian regional elections differ by gender of the candidates?

H1. In posts on Instagram during the 2024 Italian regional elections candidates were portrayed significantly more often with a positive facial expression

H2. In posts on Instagram during the 2024 Italian regional elections female candidates were portrayed significantly more often with a positive facial expression than male candidates

RQ4. How did the use of (a) verbal strategies and (b) non-verbal strategies in Instagram posts during the 2024 Italian regional elections differ by incumbency?

Methods

A quantitative content analysis was conducted to study how the main candidates running for president in seven Italian regions used verbal and non-verbal strategies on Instagram.

Italian political leaders and parties are also active on X, Facebook, and TikTok, but Instagram was selected for this study due to its growing popularity in Italy (with about 31 million active accounts in 2024) and the limited research on Italian politics and multimodality on Instagram. The unit of analysis is the entire post, including both text and visuals (Geise et al., 2024). Following Farkas and Bene (2021), only still images were analyzed, excluding video posts due to the different coding methods required for video content. While coding videos could provide additional insights, it is time-consuming and would require multiple coders and additional resources not available for this study. Each post was coded for presence (Yes) or absence (No) of verbal and non-verbal strategies. Except for the variable facial expression of the candidate, all the other variables were considered as non-mutually exclusive (Farkas & Bene, 2021). A verbal strategy includes all textual information within the post. A non-verbal strategy includes the non-verbal elements contained in each still image (for instance the portrayal of the candidate on stage, a candidate surrounding by a cheering crowd, the candidate with family members).

All the posts that the candidates for regional elections posted in the six weeks preceding the elections were retrieved and screenshots were taken and stored. A total of N = 1194 Instagram posts were included in the analysis. The Italian regional elections do not have an official start of the election campaign, but the four- or six-weeks' timeframe is commonly used in the scholarship to collect a substantial sample for quantitative studies that focus on election campaigns (Steffan, 2020).

Variables (variables were operationalized for text and visuals within each Instagram post. A full description of the variables is available but removed from this form due to words limit).

The scholarship on politics and social media guided the adaptation of a set of variables for this study. The author and coder also conducted a qualitative overview of the sample to identify verbal and non-verbal strategies not previously found in the literature.

Professional

Private

Intimate

Mobilization

Interaction

Cultural/heritage advocate

*Attack**Policy**Facial expression***Results***Selected results*

RQ1 asked which (a) verbal strategies and (b) non-verbal strategies do the candidates in the 2024 Italian regional elections use in their Instagram posts. Descriptive statistics were employed to count the frequency of each verbal and non-verbal strategy.

Verbal strategies: Professional was the most common strategy employed (74.7%), followed by policy (40.8%), mobilization (31.1%), attack (15.8%), cultural/heritage advocate (10.4%), interaction (7.7%), private (3.8%), intimate (2.7%).

Non-verbal strategies: Professional was the most common strategy (54.4%), followed by interaction (13.4%), policy (9.8%), cultural/heritage advocate (8.8%), mobilization (7.1%), private (3.8%), intimate (1.8%), attack (.7%).

RQ2 asked how did the use of (a) verbal strategies and (b) non-verbal strategies in Instagram posts during the 2024 Italian regional elections differ by political leaning. Chi-square tests of goodness-of-fit with Chi-square test with Fisher's exact tests and adjusted standardized residuals were used to detect differences between variables.

Verbal strategies: A statistical difference was found for mobilization $X^2(1, N = 1194) = 92.02, p < .001$, attack, $X^2(1, N = 1194) = 39.55, p < .001$, policy $X^2(1, N = 1194) = 4.63, p = .033$ and cultural/heritage advocate $X^2(1, N = 1194) = 5.73, p = .018$. The verbal strategy mobilization was significantly more used by candidates of center-left coalitions (44.4%) than center-right coalitions (18.7%). Attack was significantly more common in Instagram posts of center-left candidates (22.6%) than center-right candidates (9.4%). Center-left (43.2%) candidates also used more frequently the verbal strategy policy compared to center-right candidates (37.1%). The strategy cultural/heritage advocate was used more often by center right candidates (12.4%) than center-left candidates (8.2%). No significant differences were found for professional, interaction, private and intimate.

Non-verbal strategies: A significant difference was found for interaction, $X^2(1, N = 1194) = 4.94, p = .027$, cultural/heritage advocate, $X^2(1, N = 1194) = 4.59, p = .040$, and policy, $X^2(1, N = 1194) = 39.55, p < .001$. Center-left candidates (15.7%) used the non-verbal strategy interaction significantly more often than center-right candidates (11.3%). Center-right candidates used more frequently then cultural/heritage advocate non-verbal strategy (10.5%) than center-left candidates (7%). Center-right candidates were significantly more likely to use the non-verbal strategy policy (14.75%) than

center-left candidates (4.5%). No significant differences were found for professional, private, intimate, attack and mobilization.

RQ3 asked how did the use of (a) verbal strategies and (b) non-verbal strategies in Instagram posts during the 2024 Italian regional elections differ by gender

Verbal strategies: A significant difference was found for mobilization, $X^2(1, N = 1194) = 4.91, p = .028$, attack, $X^2(1, N = 1194) = 21.46, p < .001$, a

Future Work

This study identified verbal and non-verbal strategies used in seven Italian regional elections, which can be adapted for future comparative studies at local, national, and international levels, offering insights for political consultants, candidates, and journalists. Candidates primarily used Instagram to convey professionalism (verbally and non-verbally) and to some extent to describe policies but notable differences by political leaning, gender, and incumbency were observed.

Politicians from center-left coalitions, female candidates, and non-incumbents were more likely to use Instagram to mobilize voters, attack opponents, and describe policies. Overall, candidates used a broader range of verbal strategies than non-verbal ones. Non-verbal strategies mainly focused on projecting professionalism and depicting formal campaign events. The data also revealed that female candidates employed a wider variety of strategies, particularly verbally, compared to male candidates. This may be linked to the demographics of Instagram users in Italy.

Nearly 55% of Instagram users in Italy are women, with the largest group being women aged 25 to 34. It is plausible that female candidates in Italian regional elections used a variety of strategies to target female voters on Instagram. Future research should explore how female politicians in Italy use different social media platforms to confirm or challenge these findings. Comparative and transnational studies are also needed to better understand why gender influences strategic self-representation in text and visuals.

As in other countries, this research confirms that Italian politicians use Instagram to post images with facial expressions conveying optimism and positivity. Surprisingly, no differences between female and male politicians were found for any type of facial expression, contradicting studies that suggest gender influences visual emotionalization on social media. Interviews with political consultants and candidates could help explain the strategic choice of specific facial expressions on social media. Comparing facial expressions in news coverage and politicians' self-representation on social media would also be a valuable addition to the literature.

This study has several limitations. It focuses on one social media platform in a single country. Expanding the research to include multiple platforms and/or countries would

improve the generalizability of the findings. While this quantitative content analysis provides empirical insights through a relatively large dataset, interviews with political consultants and candidates would offer valuable context on why multimodality matters in the current political landscape and how verbal and non-verbal strategies are chosen for specific elections and platforms. Finally, experimental studies could be conducted to test the impact of these strategies on voting behavior, social media engagement, and voter mobilization.

References

- Auter, Z. J., & Fine, J. A. (2018). Social media campaigning: Mobilization and fundraising on Facebook. *Social Science Quarterly*, 99(1), 185-200.
- Bast, J. B. J. (2021). Politicians, parties, and government representatives on Instagram: A review of research approaches, usage patterns, and effects. *Review of Communication Research*, 9.
- Bast, J. (2024). Managing the image. The visual communication strategy of European right-wing populist politicians on Instagram. *Journal of Political Marketing*, 23(1), 1-25.
- Bossetta, M. (2018). The digital architectures of social media: Comparing political campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 US election. *Journalism & mass communication quarterly*, 95(2), 471-496.
- Boulianne, S., & Larsson, A. O. (2023). Engagement with candidate posts on Twitter, Instagram, and Facebook during the 2019 election. *New Media & Society*, 25(1), 119-140.
- Boulianne, S., & Larsson, A. O. (2024). Facebook campaigning in the 2019 and 2021 Canadian federal elections. *Politics and Governance*, 12.
- Farkas, X., & Bene, M. (2021). Images, politicians, and social media: Patterns and effects of politicians' image-based political communication strategies on social media. *The international journal of press/politics*, 26(1), 119-142.
- Filimonov, K., Russmann, U., & Svensson, J. (2016). Picturing the party: Instagram and party campaigning in the 2014 Swedish elections. *Social media+ society*, 2(3), 2056305116662179.
- Geise, S., Maubach, K., & Boettcher Eli, A. (2024). Picture me in person: Personalization and emotionalization as political campaign strategies on social media in the German federal election period 2021. *New Media & Society*, 14614448231224031.
- Geise, S., & Xu, Y. (2024). Effects of visual framing in multimodal media environments: A systematic review of studies between 1979 and 2023. *Journalism & Mass Communication Quarterly*, 10776990241257586.

- Grusell, M., & Nord, L. (2023). Not so intimate Instagram: Images of Swedish political party leaders in the 2018 national election campaign. *Journal of Political Marketing*, 22(2), 92-107.
- Haßler, J., Wurst, A. K., & Pohl, K. (2024). Politicians over issues? Visual personalization in three Instagram election campaigns. *Information, Communication & Society*, 27(5), 815-835.
- Jungblut, M., & Haim, M. (2023). Visual gender stereotyping in campaign communication: Evidence on female and male candidate imagery in 28 countries. *Communication Research*, 50(5), 561-583.
- Haim, M., & Jungblut, M. (2021). Politicians' self-depiction and their news portrayal: Evidence from 28 countries using visual computational analysis. *Political Communication*, 38(1-2), 55-74.
- Larsson, A. O. (2019). Skiing all the way to the polls: Exploring the popularity of personalized posts on political Instagram accounts. *Convergence*, 25(5-6), 1096-1110.
- McGregor, S. C., Lawrence, R. G., & Cardona, A. (2017). Personalization, gender, and social media: Gubernatorial candidates' social media strategies. *Information, communication & society*, 20(2), 264-283.
- Metz, M., Kruikemeier, S., & Lecheler, S. (2020). Personalization of politics on Facebook: Examining the content and effects of professional, emotional and private self-personalization. *Information, communication & society*, 23(10), 1481-1498.
- Micalizzi, A., Piccioni, T., & Riva, C. (2023). Pandemic in the Instagram Era. The visual analysis of Italian Mayors' communication. *Italian Sociological Review*, 13(3), 443-467.
- Navarro, C., Ganapathy, D., & Raynauld, V. (2023). Visual culture, personalization, and politics: A comparative analysis of political leaders' Instagram-based image-making and communication in Spain and India. *International Journal of Strategic Communication*, 17(4), 381-397.
- Osei Fordjour, N. K. (2024). Personalization as a strategic political tool on social media: The curious case of VP Kamala Harris on Twitter. *Howard Journal of Communications*, 35(4), 375-396.
- Peng, Y. (2021). What makes politicians' Instagram posts popular? Analyzing social media strategies of candidates and office holders with computer vision. *The International Journal of Press/Politics*, 26(1), 143-166.

Pineda, A., Hermida, A., & Bellido-Perez, E. (2025). Visualizing political communication on Spanish social media: an analysis of Andalusian political parties' communication on Instagram. *Visual Communication*, 14703572241307305.

Session 7A - Platform Epistemologies and Imaginaries

Day 2 · Tues 14 July · 11:20-12:50, RD Suite, ARC

7A.1 How is the internet imagined?: An analysis of drawings of the internet, 2014-2024

Nicholas John, Thomas Trillo (The University of Manchester)

Introduction

In this study, we are interested in how people imagine the internet by studying their drawings of it. In *Modern Social Imaginaries*, Taylor foregrounds “the way ordinary people ‘imagine’ their social surroundings [as] carried in images, stories, and legends,” referring also to the “deeper normative notions and images” that make social existence possible (2004, p. 23). More specifically, Jasanoff discusses sociotechnical imaginaries as “animated by shared understandings of forms of social life and social order attainable through, and supportive of, advances in science and technology” (2015, p. 6).

We contend that Taylor’s notion of the imaginary can be applied to interpret how people make sense of their lives on and with the internet. But whereas Jasanoff’s approach is elite focused and future oriented, we draw on drawings to investigate what the internet is for ordinary people at the time of drawing.

Imaginaries are made of metaphors, which have been a central conceptual lens for researchers of the internet since its beginnings. Scholars have identified key families of metaphors (e.g. Palmquist, 2001; Stefik, 1997), and issued critical cautions about them (e.g. Ackerman, 1994; Wyatt, 2021). Markham (2020) has examined some of the earliest internet metaphors, and has co-published (Markham & Tiidenberg, 2020) a collection of essays about them.

Amongst this plethora of writing, though, very few studies have examined the expression of internet metaphors in user-generated images (for want of a better term), which is one of the objectives of this study (see more below). Similarly to how Trillò et al. (2021) use everyday people’s Instagram posts to show what #Freedom looks like, we analyze drawings to understand everyday people’s imaginaries and metaphorical conceptions of the internet.

Objectives/Research Questions

Drawing on the concept of the imaginary and the rich literature on internet metaphors, and based on a sample of 572 drawings of the internet, our objectives in this study are twofold.

First, we offer a methodological intervention by asking whether drawings are a fruitful tool with which to glean conceptualizations and metaphors of the internet. This study can be seen as a proof of concept that they can.

Second, we propose a fruitful conceptual synergy between metaphors and imaginaries, with the latter representing fundamental building blocks of the former. Accordingly, we aim to uncover what metaphors people invoke when imagining the internet. We do so through an analysis of images produced by ordinary users over a period of a decade (2014-2024). This enables us to trace shifts in the prevalence of metaphors over time, if such exist.

Overall, our objective is to elicit new perspectives on what the internet is in people's minds.

Methods

Drawing is an established yet minor player in the researcher's toolbox. It has been especially popular among health researchers since 1972, when the "draw and write" technique was formalised. Like Guillemin (2004, p. 275), we see drawings "as ways of understanding how people see their world." As Lord (2005, p. 30) writes, they are "proposals and thoughts turned into vision."

The 572 drawings on which this study is based were not "commissioned" as part of a research project, but rather as a teaching tool. In that sense, this study is based on found data, defined as "Data not originally intended for systematic study, such as newspaper reports, hospital records, and discarded bureaucratic paperwork" (Seybolt et al., 2013). Specifically, over the course of a decade (2014-2024, excluding the 2019-20 academic year), the first author required students in an undergraduate "Internet and Society" course to draw the internet, giving them the following prompt: "In this exercise, I'd like you to draw - using paper and pens/colored pencils - the internet as you imagine it. Also, draw yourselves in the picture as an end-user, but other than that, there are no further instructions. There's no right or wrong here. Just draw what you imagine happens when you are using the internet." In 2023, a clause was added precluding the use of AI.

Two analytical approaches were adopted. In the first, the authors wrote a codebook. Two student coders attained inter-coder reliability and coded the drawings for the presence of a large number of logos, symbols (e.g. Wi-Fi, the globe, dollar signs), and

other motifs (e.g. spider webs, bodies of water, the passing of time). In total, 51 elements were coded for. Chi-squared tests were carried out in order to ascertain whether certain codes appeared more often in certain years than others.

Additionally, we compared the periods of time before and after the pandemic. The second approach was inductive. The authors read and re-read the images, analysing them in particular for their metaphorical representations of the internet and whether the students who drew them were portraying the internet as a source for good or bad, or some combination of the two.

Results

Chi-squared tests (of coded items by year) and a comparison of pre- and post-pandemic images show stability in the appearance or absence of the various logos, symbols and motifs across the decade over which the images were collected. This is somewhat surprising given that smartphones became a prevalent technology during the period when the students were children. In particular, we found no evidence that between 2014-2024 students were depicting the internet as getting increasingly better or worse.

This led us to investigate the presence of persistent, or sticky, images or metaphors, both within the corpus and in relation to pre-existing conceptualizations of the internet. Certain images of the internet from the 1990s and early 2000s continue to find expression through to the 2020s. For instance, an undated advertisement reproduced by King (2000), in which lines that arise from a point in the world and disappear around the globe, is an image of the internet that persists.

Another advert from the 1990s, and reproduced by Katz-Kimhi (2007), shows a head as a globe. This too is a visual trope adopted in some drawings. A number of metaphors identified in the 1990s also retained their hold on students' conceptualizations. For instance, in 1997, Stefik identified the metaphor of the digital library, while in 2001, Palmquist offered nine groups of internet metaphors, including outer space, highway, waterscape, and living organism. Other early metaphors also include foundational concepts for the internet, such as surfing and web.

Having said that, the images do contain a newer metaphor: the platform. Interestingly, however, the corpus suggests that this metaphor had taken root prior to the period in which the drawings were produced. By way of example, the Facebook logo appeared in 42% of the drawings, Google in 29%, Instagram in 28% and YouTube in 26%. This corresponds with Helmond and van der Vlist's (2004) assertion that in the mid-2000s "'Web 2.0' was overshadowed by the platform metaphor, becoming the dominant descriptor for the contemporary Internet landscape."

Taken as a whole, the corpus reflects ambivalent attitudes towards the impacts of the internet on the individuals who drew them, and on society more broadly. Some images show the internet as constraining and addictive, for instance, while others lay out the benefits of the internet. Some images contain this ambivalence within them as their main theme.

Future Work

While we have had space here to present only a very small portion of the drawings and to point to some of the themes identified in them, we can show that the drawings offer powerful cultural representations of the internet; that the representations remain stable over time; that the period in which they were drawn is already very much the era of platforms; that they convey a large number of metaphors, some of which are long-standing in relation to the internet; and that there is more ambivalence about the internet than outright critique or celebration.

Given the “found” nature of the data presented above, it is not possible to reproduce the temporal aspect of this study with additional populations, but this would be possible going forward. This would enable us to identify cross-cultural or cross-generational differences in conceptualizations of the internet. For example, our data ends around the time in which commercial generative AI tools made their debut. We speculate that newer drawings of the internet would feature a downsized role for platforms and a substantially more prominent one for generative AI.

Further work is also needed to address the question of the relationship between imaginaries and metaphors, and specifically the questions of whether all imaginaries generate metaphors, or whether all metaphors have imaginaries.

References

- Ackerman, M. S. (1994). Metaphors along the information highway. Proceedings of the Symposium on Directions and Impacts of Advanced Computing (DIAC'94), Cambridge, MA, 1–5.
<https://web.eecs.umich.edu/~ackerm/pub/94b10/diac94.final.pdf>
- Guillemin, M. (2004). Understanding Illness: Using Drawings as a Research Method. *Qualitative Health Research*, 14(2), 272–289.
<https://doi.org/10.1177/1049732303260445>
- Helmond, A., & Van Der Vlist, F. (2024). Platform. In J. Farkas & M. Maloney, *Digital Media Metaphors* (1st edn, pp. 24–38). Routledge.
<https://doi.org/10.4324/9781032674612-4>
- Jasanoff, S. (2015). One Future Imperfect: Science, Technology, and the Imaginations of Modernity. In S. Jasanoff & S.-H. Kim, *Dreamscapes of modernity: Sociotechnical*

- imaginaries and the fabrication of power (pp. 1–33). The University of Chicago Press.
- Katz-Kimchi, M. (2007). *The Rise of the Internet in America: Myths, Metaphors and Discourses* [PhD]. Bar Ilan University.
- King, A. B. (2000). *Mapping the Unmappable: Visual Representations of the Internet as Social Constructions*. Rob Kling Center for Social Informatics · Working Papers (RKCSI). <https://hdl.handle.net/2022/171>
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
- Lord, J. V. (2005). *A Journey of Drawing: An Illustration of a Fable*. In J. Davies & L. Duff (Eds), *Drawing: The Process* (pp. 32–41). Intellect, Limited.
<http://ebookcentral.proquest.com/lib/manchester/detail.action?docID=282994>
- Markham, A. N. (2020). *Ways of Being in the Digital Age*. In A. N. Markham & K. Tiidenberg, *Metaphors of internet: Ways of being in the age of ubiquity* (pp. 3–11). Peter Lang Publishing.
- Markham, A. N., & Tiidenberg, K. (2020). *Metaphors of internet: Ways of being in the age of ubiquity*. Peter Lang Publishing.
- Palmquist, R. A. (2001). Cognitive style and users' metaphors for the web: An exploratory study. *The Journal of Academic Librarianship*, 27(1), 24–32.
[https://doi.org/10.1016/S0099-1333\(00\)00205-6](https://doi.org/10.1016/S0099-1333(00)00205-6)
- Seybolt, T. B., Aronson, J. D., & Fischhoff, B. (2013). *Counting civilian casualties: An introduction to recording and estimating nonmilitary deaths in conflict*. Oxford University Press.
- Stefik, M. J. (Ed.). (1997). *Internet Dreams: Archetypes, Myths, and Metaphors*. MIT Press.
- Taylor, C. (2004). *Modern social imaginaries*. Duke University Press.
<https://doi.org/10.1515/9780822385806>
- Trillò, T., Scharlach, R., Hallinan, B., Kim, B., Mizoroki, S., Frosh, P., & Shifman, L. (2021). What Does #Freedom Look Like? Instagram and the Visual Imagination of Values. *Journal of Communication*, 71(6), 875–897.
<https://doi.org/10.1093/joc/jqab021>

7A.2 ZOOMWORK: Work imaginaries in platform culture. Content creators and Gen Z

Camilla Volpe, Luca Giuffè, Giulia Giorgi (University of Milan)

Introduction

In recent years, social media platforms have become central infrastructures of everyday life, shaping how young people imagine, discuss, and experience work. For Generation Z, platforms such as Instagram and TikTok function not only as spaces of social interaction and cultural expression, but also as environments in which occupational norms, professional aspirations, and attitudes toward employment are produced and negotiated (Stahl & Literat, 2022). Against a backdrop of labour market uncertainty and normalized precarity, young people increasingly rely on digital platforms to make sense of work and to experiment with possible career trajectories (Gandini & Leonini, 2019). In this context, social media provide symbolic resources and material opportunities, contributing to the expansion of platform-mediated forms of digital and creative labour (Duffy, 2017). While existing scholarship has extensively examined platform labour and influencer cultures, less attention has been paid to how social media contribute to the collective construction of generational cultures of work as shared imaginaries emerging from platformed discourse and participation (Bucher, 2017).

Objectives/Research Questions

Building on these debates, this paper investigates how social media platforms shape Generation Z's imaginaries and cultures of work, focusing on the role of platform affordances, algorithmic logics, and participatory practices in the collective construction of shared attitudes toward occupation. Conceptualizing work as a cultural and discursive object rather than solely as formal employment, the study examines how narratives around work, success, refusal, and professionalization circulate on platforms, while also offering digitally mediated alternatives to traditional occupations. By centring on Generation Z, the paper addresses a cohort for whom platforms function as constitutive environments of socialization, identity formation, and work-related sense-making.

To account for these dynamics, the paper adopts the concept of work imaginaries as an adaptation of sociotechnical imaginaries (Jasanoff & Kim, 2015). While the latter were originally developed to examine how collectively imagined futures are embedded in national institutions and technological projects, this study shifts the analytical focus toward platform-mediated environments and everyday cultural practices. Here, work imaginaries are understood as collectively produced visions of work and work-life

projects that emerge from discourse, participation, and routine interaction within social media ecosystems (Tidwell & Tidwell, 2018). These imaginaries are materially shaped by platform affordances, algorithmic governance, and datafied infrastructures that organize visibility, recognition, and value (Srnicek, 2016; van Dijck et al., 2018).

Within this framework, content creators are examined as an empirically grounded and analytically revealing case for investigating contemporary work imaginaries. Content creation represents a paradigmatic form of platform-mediated labour, positioned at the intersection of creative expression, everyday life, and economic activity. Existing research shows how creators engage in aspirational, visibility, and reputational labour, investing time, affect, and creativity in practices oriented toward uncertain future opportunities (Abidin, 2016; Duffy, 2017; Duffy et al., 2021).

Although central to platform economies, these practices are often ambiguously positioned in relation to dominant definitions of work, blurring the boundaries between leisure, self-expression, and labour. Focusing on content creators enables the paper to examine how work imaginaries are shaped not only through explicit discourse about employment, but also through routine platform practices that embed specific understandings of autonomy, professionalism, and value into cultural production and circulation.

Methods

Methodologically, the study adopts a mixed qualitative and digital approach grounded in Digital Methods and cross-platform analysis (Rogers, 2017; Caliandro & Gandini, 2016). Empirical digital data are collected from Instagram and TikTok, two platforms that play a central role in young people's media practices and in the circulation of work-related content. The digital analysis focuses on posts, visual formats, captions, hashtags, and platform-specific objects. Data collection is conducted through platform-compliant scraping tools and exploratory sampling strategies designed to capture platform vernaculars and algorithmically curated forms of visibility (Zulli & Zulli, 2022). This digital analysis is complemented by qualitative methods, including semi-structured interviews and focus groups with young people, aimed at investigating how platform-mediated narratives of work are interpreted, negotiated, and incorporated into everyday life and professional self-understandings. Visual materials derived from the platform analysis are used as elicitation tools during focus groups to support collective reflection and discussion (Niederer & Colombo, 2019).

The analytical strategy combines visual analysis, clustering techniques, and critical discourse analysis applied to both digital content and qualitative data (Rose, 2022; Omena et al., 2021), allowing the study to trace how work imaginaries emerge at the intersection of discourse, practice, and platform infrastructures.

Results

The analysis allows for the identification of recurring discursive and visual patterns through which work is imagined, discussed, and made meaningful within platformed environments. Across Instagram and TikTok, work-related content is articulated through narratives that foreground flexibility and individual agency, often intertwined with expectations of constant engagement, adaptability, and self-presentation. Rather than pointing to a single, coherent model of work, the material reveals a plurality of imaginaries that coexist and sometimes conflict, reflecting young people's attempts to reconcile aspirations for autonomy and fulfilment with experiences of uncertainty and instability. Interviews and focus groups further illuminate how these platform-mediated narratives are interpreted and negotiated in everyday life: participants draw on social media content as a resource for orienting themselves within fragmented labour markets, using platforms to imagine possible trajectories, evaluate opportunities, and make sense of what counts as legitimate or desirable work. At the same time, they express ambivalence toward platform-mediated forms of labour, which are often perceived as accessible and visible, yet difficult to situate within established definitions of employment.

Future Work

Taken together, these findings suggest that social media platforms function as key sites for the circulation and contestation of work imaginaries, where meanings of work are continuously produced, revised, and negotiated rather than stabilized. This research also opens critical avenues for future inquiry into the temporal durability of these visions. More specifically, future scholarship should examine how these platform-mediated work imaginaries evolve as Generation Z transitions further into adulthood.

Moreover, as platform increasingly integrate Generative AI tools, the definition of “creative labour” poised for a significant sociotechnical shift. Future work must investigate how the introduction of AI-generated media alters the current system of value creation and how reshapes aspirational narratives of “creativity” and “authenticity” currently central to Generation Z's work ethic.

Finally, these findings show timely practical implications. Generation Z's work imaginaries are critical sites for policymakers and educators aiming to support a generation navigating labour market instability. Educators, for instance, need to go beyond basic digital skills and cultivating grassroots and critical understanding of platforms as non-neutral environments mediating younger generations' future expectations. Findings can support policymakers in recognizing hybrid forms of work that no longer fit traditional employment categories.

References

- Abidin, C. (2016). “Aren't these just young rich women doing vain things online?": Influencer selfies as subversive frivolity. *Social Media and Society*, April-Jun, 1–17.

- Bucher, T. (2017). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, 20(1), 30–44.
- Caliandro, A., & Gandini, A. (2016). Digital methods for qualitative research. In *Qualitative Research in Digital Environments*. Routledge.
- Duffy, B. (2017). (Not) Getting Paid to Do What You Love: Gender, Social Media, and Aspirational Work. In *(Not) Getting Paid to do what You Love: Gender, Social Media, and Aspirational Work* (p. 308).
- Duffy, B. E., Pinch, A., Sannon, S., & Sawey, M. (2021). The nested precarities of creative labor on social media. *Social Media+ Society*, 7(2), 20563051211021368.
- Gandini, A., & Leonini, L. (2019). The myth of flexibility: Young adults' expectations of work in the digital economy in Milan. In *Youth and the Politics of the Present*. Routledge.
- Jasanoff, S., & Kim, S.-H. (Eds.). (2015). *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. University of Chicago Press.
- Niederer, S., Colombo, G., & Politecnico di Milano. (2019). Visual Methodologies for Networked Images: Designing Visualizations for Collaborative Research, Cross-platform Analysis, and Public Participation. *Revista Diseña*, (14), 40–67.
- Srnicek, N. (2016). *Platform Capitalism*. Polity.
- Stahl, C. C., & Literat, I. (2023). #GenZ on TikTok: The collective online self-Portrait of the social media generation. *Journal of Youth Studies*, 26(7), 925–946.
- Tidwell, J. H., & Tidwell, A. S. D. (2018). Energy ideals, visions, narratives, and rhetoric: Examining sociotechnical imaginaries theory and methodology in energy research. *Energy Research & Social Science*, 39, 103–107.
- van Dijck, J., Poell, T., & Waal, M. de. (2018). *The Platform Society: Public Values in a Connective World*. Oxford University Press.
- Zulli, D., & Zulli, D. J. (2022). Extending the Internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform. *New Media & Society*, 24(8), 1872–1890.

7A.3 Authenticity Effects of Synthetic Images: Interactions among Political Cues, Context, and Platform Environments

Yu-Chung Cheng (National Chengchi University)

Introduction

Generative AI is making public-event imagery malleable: photographs can be lightly retouched, recomposited, or fully generated, then circulated at scale across platforms. Research on synthetic media often focuses on detection or misinformation, but a basic problem remains under-theorized: when content can be repeatedly synthesized and iteratively remade, how do publics decide what counts as credible, and how do platform interactions turn images into “evidence” for political stance-taking?

We approach synthetic media as a socio-technical process of human–AI co-creation rather than a simple true/false binary, and define synthetic authenticity (or synthetic truthfulness) as the felt, interactionally produced sense that a synthetic artifact is believable, usable, or “as-if true” within particular communities and platform environments. Authenticity here is not reducible to visual realism; it is produced at the intersection of political polarization, distrust in news institutions, algorithmic visibility, and memetic vernaculars.

This perspective also speaks to debates on synthetic imaginaries and sociotechnical imaginaries, highlighting how generative AI is reshaping public-cultural standards of evidence, credibility, and representation, and thereby reordering the conditions of public communication. Pilipets & Geboers (2025) use the concept of synthetic imaginaries to describe how generative AI models produce complex, layered structures of visibility and invisibility. Focusing on how GPT-4o re-imagines “sensitive” visual content under platform content-policy constraints, they analyze generative AI as a mechanism of cultural amplification and a site of visibility politics.

In a complementary vein, Laba (2024) conceptualizes AI image generation as a sociotechnical practice situated at the intersection of humans, machines, and visual culture, whose social logic can be parsed along three dimensions. First, mythology: a widely held belief that AI operates through disembodied machine-learning processes, lending generated images an aura of authenticity, truth, and epistemic objectivity. Second, technology: the computational capacities of deep neural networks transform prompts into visual outputs, operationalizing image generation as a prompt-driven pipeline. Third, representation: AI-generated imagery is deployed for personal, professional, political, or ideological purposes, in ways that blur the boundaries of agency between human and machine.

Given how deeply synthetic media has become embedded in everyday life, Kalpokas (2024) argues that “reality” should be evaluated by its affective impact rather than by physical existence alone. This affective criterion suggests that an object or situation is experienced as real to the extent that it produces intense emotions and sensations. Under such conditions, contemporary reality increasingly takes the form of synthetic reality: simultaneously mediated and technologically constructed, while also affective and individualized.

Objectives/Research Questions

Recent efforts to systematize political synthetic incidents (e.g., PDID; Walker et al., 2024) and to measure synthetic-media spread on X via Community Notes (Corsi et al., 2024) underscore the need to move beyond item-level veracity.

We therefore analyze the interactional chain through which images are edited, captioned, memefied, cross-posted, and institutionally responded to. Our focal case from Taiwan—the “Smiling Police” controversy—shows how a micro edit (serious → smiling facial expression) can trigger stance-driven readings, be re-uploaded without provenance, and evolve into meme templates, progressively detaching the artifact from its original context.

RQ1 (Typology). How can synthetic-image incidents in public-issue circulation be operationalized along (a) synthesis mode, (b) narrative function, and (c) visibility mechanism? Which types are more likely to be perceived as credible, treated as evidence, or transformed into memes?

RQ2 (Diffusion & reinterpretation). How are synthetic images reproduced, captioned, composited, and cross-posted to form an image-to-stance meaning-translation chain? How do affective mobilization, echo-chamber interaction, and media distrust shape diffusion speed and lifecycle?

RQ3 (Authenticity effects & governance). How do journalism, fact-checking, and platform governance respond to synthetic media, and under what conditions do these responses strengthen or weaken trust and alter visibility/credibility dynamics?

Methods

We use a mixed-method design that combines synthetic ethnography with human–LLM collaborative, cross-modal analysis.

Phase 1 (Cross-platform tracing & ethnography). We compile contested synthetic-media incidents in Taiwan and the Chinese-language media sphere. For each case, we record source platform(s), timestamps, relay accounts, mainstream news coverage, and fact-checking records, then iteratively code synthesis mode, narrative function, and visibility mechanism. For the focal case, the “Smiling Police” controversy, we reconstruct diffusion paths using reverse image search, repost-link mapping, and

keyword tracking: original image → synthetic variant → meme remakes → cross-platform transfers. We conduct netnographic observation of comments to identify stance markers, affective vocabularies, and rule-of-thumb credibility reasoning.

Phase 2 (Multimodal LLM comparison). We run controlled prompts on three artifact types—(a) authentic photo (Fig. 1-1), (b) expression-edited variant (Fig. 1-2), and (c) text-overlay meme (Fig. 1-3)—under systematically varied contexts (no caption vs. caption; neutral vs. stance-framed cues). We log model outputs and conduct human verification for (i) descriptive accuracy, (ii) stance inference, and (iii) over-interpretation/hallucination.

Results

In the first phase of the study, we reconstructed the initial cross-platform diffusion pathway of the “Smiling Police” synthetic-media incident (Figure 1) by conducting reverse image searches and manually aligning timestamped posts into a chronological sequence. The source image originated from a livestream on the Taiwan People’s Party (TPP) YouTube channel, Voice of the People. In the screenshot, TPP chair Huang Kuo-chang wraps his arm around the neck of a police officer who is maintaining order; the officer’s facial expression appears tense and confrontational, conveying pressure and restraint (Fig. 1-1). A user then used the Gemini tool to alter the officer’s expression into a broad smile and posted the original screenshot alongside the AI-edited version on Threads (Fig. 1-2). Subsequently, other users circulated a watermark-free version of the AI-altered image and repurposed it as a meme by adding personal commentary across social media platforms (Fig. 1-3).

In the second phase, we employed the multimodal large language model GPT-4.1-mini to conduct cross-modal image–semantic analysis. This step examines how generative models interpret (1) a “real image,” (2) an “AI-synthetic media” version, and (3) “meme-style remixes” under varying contextual conditions. We then compared the analytic dimensions the LLM exhibited across these three image types, including affective tone, semantic focus, degree of context dependence, interpretation profile, and potential biases. The results are presented in Table 1.

Preliminary findings suggest that context dependence is a key mechanism underlying authenticity effects in synthetic media. Using the three visual materials from the “Smiling Police” case (real image, AI-synthetic version, and meme remixes) as an example, the LLM’s descriptions of the real image’s scene and affect remain relatively stable without accompanying textual cues. It can identify visual cues such as confrontation, crowding, tension, and a sense of pressure, yet it struggles to infer the actors’ identities, the political setting, or the contested background on its own.

When interpreting the AI-synthetic image, the model can detect salient incongruities—such as the abrupt shift in facial expression or composition—and the resulting affective

reversal. However, it does not necessarily grasp the symbolic weight that a “minor edit” (e.g., changing a serious expression into a smile) can carry in political contention. Once caption text and user context are introduced, the model’s interpretation quickly becomes stance-laden: meme registers (mockery, defense, accusation) become primary cues for inference, enabling the model to situate the image within broader frames of political confrontation and media distrust. At the same time, this heightened contextual anchoring increases the risk of over-rationalization and motive imputation, where the model fills in intentions beyond what the evidence warrants.

Future Work

Next steps: (1) expand the incident corpus and conduct cross-platform network analysis to quantify lifecycles, relay nodes, and transfer routes across typological categories; (2) add interviews with journalists, fact-checkers, and governance practitioners to examine how institutional actions reshape trust and visibility; (3) develop a rigorous evaluation framework for human–LLM workflows (error taxonomies, bias sources, prompt-sensitivity tests) to improve interpretability and efficiency in cross-modal synthetic-media analysis.

Overall, this study seeks to move beyond a binary true/false judgment of “authenticity” and advance authenticity as an operationalizable, process-oriented account of interaction: how synthetic media are reproduced within platform cultures, endowed with meaning, converted into “evidence,” and governed. In turn, “synthetic authenticity” emerges—through the interplay of affect, stance, visibility, and institutional response—as a central challenge for public communication in the era of generative AI.

References

- Corsi, G., Marino, B., & Wong, W. (2024). The spread of synthetic media on X. *Harvard Kennedy School (HKS) Misinformation Review*, 5(3), 1–19.
<https://doi.org/10.37016/mr-2020-140>
- Kalpokas, I. (2024). Reality nostalgia: Musings on the post-information, post-truth society. In I. Kalpokas, *Technological governance and escapism in times of accelerated change* (pp. 29–50). Springer Nature.
- Laba, N. (2024). Engine for the imagination? Visual generative media and the issue of representation. *Media, Culture & Society*, 46(8), 1599–1620.
<https://doi.org/10.1177/01634437241259950>
- Pilipets, E., & Geboers, M. (2025). Synthetic imaginaries of “sensitive” AI: On ambient amplification and jail(break)ing as method. *Platforms & Society*, 2, 1–15.
<https://doi.org/10.1177/29768624251378110>

Walker, C. P., Schiff, D. S., & Schiff, K. J. (2024). Merging AI incidents research with political misinformation research: Introducing the Political Deepfakes Incidents Database. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(21), 23053–23058. <https://doi.org/10.1609/aaai.v38i21.30349>

7A.4 Hyper-datafication: How Platforms Make Everything Feel Like Social Media

Aparajita Bhandari (University of Waterloo), Sara Bimo (York University), Chelsea Butkowski (American University)

Introduction

In the two decades since the advent of mainstream social media, scholars and journalists alike continue to debate what makes a digital platform a social media platform. This discursive boundary work has only escalated in recent months as users seek alternatives to longstanding yet increasingly divisive social media behemoths, sometimes leading them to unexpected digital platforms. Beyond touchstone social media sites like Facebook, Instagram, and TikTok, a broader array of fitness, streaming, education, dating, gaming, and payment apps have become vital to discussions about the true nature of social media. Such trends are reflected in popular discourse: for example, a recent Wired article suggested that the music streaming platform Spotify was “trying to become social media” (Klein, 2024) while The Guardian asked “Have hobby apps become the new social media?” when documenting the rise of platforms like the fitness tracker Strava as growing sites for digital community and connection (Stokel-Walker, 2024). Similarly, The Atlantic has even labeled Venmo as “one of the last real social networks” because its peer-to-peer payment posts capture nostalgic “nuggets of gossip and banal updates” (Kelley, 2024). As new and sometimes unexpected platforms continually incorporate social media-like features, communities, and self-presentation, these articles demonstrate how the boundaries between so-called “social” and “non-social” media have become a key concern of contemporary tech discourse even as scholarship suggests that they remain blurred (Sujon, 2021).

Objectives/Research Questions

Why do so many digital spaces now seem to feel and function like social media, as the aforementioned articles indicate? And what purpose does this shift—evidenced through changing trends in platform design, affordances, and governance—serve within the wider context of affective capitalism and increasing datafication? In this project, we pursue this line of inquiry through the analysis of three applications which may not be straightforwardly classified as social media, but—in line with the above popular narratives—still seem to feel like social media: Strava, Venmo, and Spotify.

Each of these free apps is ranked in the top five most popular of its kind by the Apple App Store, but each also focuses on different types of user behaviors on different interactional scales. Strava compiles individual bodily and locational behaviors, Venmo

captures peer-to-peer financial exchanges, and Spotify facilitates engagement with the broader music media ecosystem and algorithmic tastemaking.

In this paper, we examine the ways that “non-social media” platforms systematically repackage user behavioral data to construct social media-like interactions and expressions that ultimately serve affective capitalism, an economic system predicated on the ability to “capture, structure, and modulate the infrastructures” through which affect moves, such as digital platforms and the internet (Karppi et al., 2016, p. 2).. The present study examines two key research questions: 1) How do so-called “non-social media platforms” leverage the feelings, aesthetics, and experiences of social media through the incorporation of specific features, affordances, and affect? 2) What are the implications of this transformation within a broader system of affective capitalism?

Methods

To understand how our app case studies navigate and exploit the expressive component of behavioral data, we undertook platform walkthroughs of each app (Light et al., 2018). Drawing from critical data studies and science and technology studies, platform walkthroughs provide a systematic method for immersive engagement with mobile applications based on components of their “environment of expected use” comprising app vision, operating model and governance alongside a “technical walkthrough” that traces usage processes of registration and entry, everyday use, and discontinuation of use (Light et al., 2018, p. 881). The walkthrough method enables critical analysis of “the relatively closed systems” of apps “to understand how [they guide] users and shape their experiences” (Light et al., pp. 881–882).

To support sustained engagement with each platform (Charmaz, 2014), we completed the walkthroughs over a six-month period with an initial exploratory analysis between September 2021 and February 2022 and then a secondary analysis between September 2024 and February 2025. The authors did walkthroughs of all three platforms, with each author taking one platform as an area of focus. After the walkthroughs were complete, we conducted detailed textual analysis of notes and screenshots gathered during the walkthroughs (Kuckartz, 2014). We discussed emergent thematic categories in detail to come to an agreement on our findings (Charmaz, 2014).

Framing this examination through the lens of affective capitalism (Hearn, 2017; Karppi et al., 2016), we examine social media beyond a specific class of platforms as an ecosystem of characteristic affective experiences enabled within digital environments. By leveraging this “social media-ness,” tech companies stand to strengthen user investment in their platforms while also producing and re-using valuable data.

Results

Ultimately, we find that the “social media-fication” of the app landscape can be attributed in part to the systematic mobilization of users’ behavioral data traces as expressive social updates. The social media experience on “non-social media” platforms is enabled by platform affordances that encourage users to engage with representations of their own behavioural data to inspire expressive interactions that, in turn, generate even more forms of behavioral data. Examples such as Strava’s fitness activity visualizations, Venmo’s peer-to-peer payment social feed, and Spotify Wrapped’s algorithmic data aggregations showcase “social media-fication” as manifest through what we call “hyper-datafication,” a process of expressive data repackaging layered on top of traditional datafication.

We introduce the term hyper-datafication not to suggest a fundamentally novel form of datafication, but rather to demonstrate how regimes of datafication become intensified by re-incorporating the products of datafication as the raw materials of hyper-datafication. By codifying “backend” behavioral data into expressive user-facing updates, platforms stand to further entrench and normalize datafication among their userbase while also extracting further value from surveilling those same users.

We argue that the drive towards hyper-datafication facilitates an increase in the kinds and amount of behavioural data available for such forms of value generation. While non-social media platforms offer behavioural data in the form of “tracking users’ browsing behaviour,” social media platforms engender behaviours wherein data is “actively provided by users themselves” (Fast et al., 2022, p. 202). Thus, the social media-ification of these platforms, aided by the mobilization of data representations as opportunities for expression, results in the expansion of the grounds for social capture via the introduction of new kinds of affordances for social experience. Hyper-datafication thus offers an additional “pathway” for value generation: the re-incorporation of data-representations into the environment of the platform for the purpose of generating value by increasing the scope, quality, and quantity of the kinds of data available for value extraction.

Future Work

Through this theorization, we expand upon existing scholarship on the datafication of social life, while also contributing to ongoing discussions about the social media-like characteristics that have become embedded in a variety of sometimes unexpected platforms. We deconstruct and expand contemporary understandings of social media to draw critical attention to the economic and political implications of everyday digital platforms incorporating expressive social affordances. As digital platforms of all kinds continue to emerge, grow, recede, and collapse, we predict that capacities for theorizing platform hybridity and flexibility will become increasingly vital in a constantly expanding and temperamental atmosphere.

References

- Charmaz, K. (2014) Constructing grounded theory. Sage.
- Fast, V., Schnurr, D., & Wohlfarth, M. (2023). Regulation of data-driven market power in the digital economy: Business value creation and competitive advantages from big data. *Journal of Information Technology*, 38(2), 202-229.
- Hearn, A. (2017). Verified: Self-presentation, identity management, and selfhood in the age of big data. *Popular Communication*. 15(2), 62–77.
- Karppi, T., Kähkönen, L., & Mannevu, M. et al. (2016). Affective capitalism: Investments and investigations. *Ephemera: Theory & Politics in Organization*, 16(4).
- Kelley, L. (2024, Aug. 30) The last social network. *The Atlantic*.
- Klein, E. (2024, July 18). Spotify, stop trying to become a social media app. *Wired*.
- Kuckartz, U. (2014). *Qualitative textual analysis: A guide to methods, practice, and using software*. Sage.
- Light, B., Burgess, J., and Duguay, S. (2018). The walkthrough method: An approach to the study of apps. *New Media & Society*, 20(3), 881-900.
- Stokel-Walker, C. (2024, Sept. 8). Goodbye Tinder, hello Strava: Have hobby apps become the new social networks? *The Guardian*.
- Sujon, Z. (2021). *The social media age*. Sage.

Session 7B - Health and Wellbeing (2)

Day 2 · Tues 14 July · 11:20-12:50, Seminar Suite A, ARC

7B.1 Navigating Mental Health Online: How Platform Affordances Shape Emotional Connection on TikTok and Threads

Adrienne Hall-Phillips, Khushali Shah (Worcester Polytechnic Institute)

Introduction

Social media platforms have become a primary space where consumers discuss and engage with mental health discourse, yet there is a lack of understanding about how platform experiences shape their experience. Evidence from prior studies show that social media use can both enhance and undermine well-being depending on user motivations and community dynamics (Ansari et al., 2024). We also know online expressions of support and disclosure meaningfully influence psychological resilience (De Choudhury & Kiciman, 2017). At the same time, consumer research shows that individuals use narratives, reference groups, and symbolic cues to construct identity and signal values (Escalas & Bettman, 2005). In mental health spaces online, consumers build identity through hashtags, peer stories, and platform interactions while, using these tools not just to find information but to signal belonging and values like #EndTheStigma. Bennett & Segerberg (2012) describe these activities as connective actions, where consumers engage through personalized, digitally mediated storytelling rather than formalized and organized collective structures. Prior work on hashtag-enabled political consumption shows how shared digital markers allow consumers to express identity, values, and belonging across networks (Johnson et al., 2019). Yet little is known about how these same mechanisms operate in mental health discourse. That is, how different platforms shape perceived usefulness, authenticity, support, or emotional outcomes. This gap limits our understanding of the consumer, identity, and well-being implications of a rapidly expanding digital mental health ecosystem.

Objectives/Research Questions

Although social media platforms have positively transformed how individuals access mental health information, researchers have documented both beneficial and detrimental effects, revealing a complex relationship that depends heavily on how users engage with these platforms and what features shape their experiences. Some

researchers have shown that excessive platform use can lead to increased mental health concerns, like anxiety or depression. The phenomenon of “doomscrolling” (passively consuming negative content for extended periods) exemplifies how platform design can trap users in harmful consumption patterns. Despite this documented harm, researchers have also identified substantial benefits, particularly for individuals managing mental health challenges. One study examined how individuals with severe mental illness use YouTube to create and consume peer support content. They found that users share personal recovery stories, coping strategies, and medication experiences, creating what they termed “naturally occurring peer support” (Naslund et al., 2014). Peer-generated content helps users feel less isolated, more empowered in their recovery, and connected to a community that understands their experiences. Users can also connect with others via these platforms at any time, providing lots of flexibility and more access than traditional in-person help.

While existing research documents that social media can both harm and help mental health, a critical gap remains in understanding the mechanisms through which platform design features produce these divergent outcomes. Experiences depend on how users engage on these platforms, what content they consume, and their individual characteristics, but researchers have not systematically examined which specific platform affordances drive positive versus negative pathways. This gap is particularly pronounced in mental health discourse contexts. Naslund et al. (2014) and Naslund et al. (2016) demonstrated that people with mental health challenges actively use social media for support and information, yet we lack understanding of how platform features, shape consumer experiences, evaluations, and outcomes.

Therefore, this study aims to explain how social media platform design features shape the ways users engage with and experience mental health–related content online by examining the cognitive-evaluative and social-identity pathways that link platform affordances to user well-being. First, we focus on how three key affordances (multimodality, interactivity, and anonymity) influence content usefulness and social support. Next, we focus on these psychological evaluations and their influence on identity and belonging. Lastly, we will assess how identity processes and community belonging affect mental health well-being outcomes, like stress reduction, anxiety management, and overall well-being.

Methods

This study uses an experimental survey design to test how social media affordances influence psychological evaluations of the content found in mental health-related social media discourse. After agreeing to participate, participants will complete a brief screening to confirm that they have recently viewed mental health–related content online (e.g., posts about anxiety, stress, or emotional well-being) and that they actively use TikTok and Threads for at least 30 minutes per day. Eligible participants will then

review the informed consent form and proceed to the study.

Participants will view one of two simulated social media feeds mimicking TikTok or Threads. Each mock feed will resemble a typical scrolling experience, including cues for visibility, authenticity, and modality (e.g., video thumbnails, captions, emojis, or text-based replies). The study prompt will read:

You'll now view examples of mental health-related posts that resemble those commonly seen on TikTok or Threads. Please think about the last time you engaged with a similar post and answer the following questions based on that experience.

After viewing the feed, participants will complete the survey.

Participants will be adults aged 18–29 who identify as regular users of TikTok and Threads (minimum 30 minutes of daily use per platform). Recruitment will occur through CloudResearch, a third-party service that connects researchers with verified online participants. The target sample size is 400–500 participants, which provides sufficient power to detect medium effect sizes in between-group comparisons and structural equation modeling analysis.

Results

This study examines how platform affordances (multimodality, interactivity, and anonymity) shape consumers' emotional and social outcomes through perceived usefulness, perceived support, identity, and belonging. Guided by affordance theory and social identity perspectives, we anticipate that higher levels of multimodality and interactivity will be associated with greater perceived usefulness and perceived support, which in turn strengthen feelings of identity expression and belonging. These identity- and belonging-based processes are expected to relate to lower reported stress and anxiety. Anonymity is anticipated to play a more nuanced role, potentially increasing perceived support while differentially shaping identity and belonging depending on platform context. We further expect these relationships to vary across platforms, such that TikTok's multimodal and highly interactive environment amplifies perceived usefulness and support, whereas Threads' more text-based and identity-linked design may strengthen the role of belonging in shaping emotional outcomes. Together, these analyses will clarify how platform-level affordances indirectly shape stress and anxiety through social and psychological mechanisms in digital mental health engagement.

Future Work

This study examines mental health engagement on social media as a consumer behavior phenomenon, developing an integrated framework that explains how platform affordances, narrative content, and social signaling collectively shape digital well-being

experiences. This integrated approach will contribute to consumer well-being research and clarify how engagement with mental health content affects identity processes, emotional outcomes, and perceptions of community connectedness. This approach advances consumer well-being theory by clarifying how social media functions as both a resource and risk for mental health, revealing the conditions under which digital engagement supports or undermines well-being outcomes. We expect findings to shed light on how consumers co-create meaning and seek support in digital environments, and how these processes shape emotional well-being and social connections. This work advances consumer well-being theory and offers actionable recommendations for educators, brands, and platform designers striving to foster healthier, more inclusive online communities. Further, future studies may translate these insights into educational resources or industry-facing frameworks that help mental health professionals, university student-facing staff, and to support individuals in a society that is increasingly concerned and forthcoming about mental health struggles and concerns.

References

- Ansari, S., Iqbal, N., Asif, R., Hashim, M., Farooqi, S. R., & Alimoradi, Z. (2024). Social media use and Well-Being: A systematic review and Meta-Analysis. *Cyberpsychology, Behavior, and Social Networking*, 27(10), 704-719.
- Bennett, W. L., & Segerberg, A. (2012). The Logic of Connective Action. *Information, Communication, & Society*, 15(5), 739-768.
<https://doi.org/10.1017/cbo9781139198752>
- De Choudhury, M., & Kiciman, E. (2017, May). The language of social support in social media and its effect on suicidal ideation risk. In *Proceedings of the international AAAI conference on web and social media* (Vol. 11, No. 1, pp. 32-41)
- Johnson, O., Hall-Phillips, A., Chung, T.-L. (Doreen), & Cho, H. (2019). Are you connected through consumption? The role of hashtags in political consumption. *Social Media + Society*, 5(4). <https://doi.org/10.1177/2056305119883427>
- Naslund, J. A., Aschbrenner, K. A., Marsch, L. A., & Bartels, S. J. (2016). The future of mental health care: peer-to-peer support and social media. *Epidemiology and Psychiatric Sciences*, 25(2), 113–122. doi:10.1017/S2045796015001067
- Naslund, J. A., Grande, S. W., Aschbrenner, K. A., & Elwyn, G. (2014). Naturally occurring peer support through social media: the experiences of individuals with severe mental illness using YouTube. *PLOS one*, 9(10), e110171.

7B.2 Designing for Attention vs. Designing for Wellbeing: The Impact of a Screentime Nudge on TikTok Flow

Amber Priem, Malaika Brengman, Nanouk Verhulst (Vrije Universiteit Brussel)

Introduction

Social media platforms are designed with the attention economy in mind. Specifically, TikTok is optimised to maximise user attention and engagement through technological affordances, such as infinite scroll, algorithmic personalisation, and short-form video formats that align with users' attention spans. These platform-specific characteristics facilitate intense user involvement and long-term use.

A central concept to understand TikTok experiences is psychological flow state (Zhao & Wagner, 2022). Flow is defined as a state of deep absorption, in which an individual's full attention is devoted to the current task, resulting in a lack of attention to external stimuli and a sense of time distortion (Csikszentmihalyi, 1990; Lavoie, Main, & Stuart-Edwards, 2022). Flow is actively induced by the technological affordances of short-form video platforms, where sustained user engagement is central to value creation in the attention economy. Although flow is usually associated with positive outcomes, such as enjoyment and intrinsic motivation, it may have a different effect on social media platforms. In this context, the typical positive outcomes of flow may be overshadowed by a flow cost, where users continue scrolling long after they were intrinsically motivated to do so (Pham & Duff, 2022).

At the same time, political and societal pressures on short-form video platforms are increasing, calling on them to take responsibility for excessive or problematic social media use, especially among young people. Since 2018, platforms have introduced digital well-being functions within the apps, such as screentime reminders and interruption notifications (Pardes, 2018). These functions are presented by companies as tools for self-regulating use and align with a general shift towards soft-paternalistic behavioural interventions, called nudges.

The platforms are thus starting to position themselves as actors in their users' digital well-being, often young users. However, platform existence depends on capturing long-term, flow-like attention. It is yet unclear whether these functions influence the user's flow experience and TikTok usage, or whether they are solely virtue-signalling functions that do not truly affect the user's attention. Previous research has focused more on outcome variables, such as time spent on the platform, rather than on user experience aspects, such as flow and user satisfaction (Vu & Tagliabue, 2025; Zimmermann, 2021).

This research focuses on the effects of a screentime nudge on psychological flow, TikTok use, and perceived user satisfaction. It assumes that a screentime nudge may influence both the user experience and their behaviour. Furthermore, potential moderators are examined, including intrinsic motivation, self-efficacy, and frequency of TikTok use. Thus, this study contributes to the debate on how designing for well-being and designing for attention interact in short-form video platforms.

Objectives/Research Questions

This study explores how in-app digital well-being tools influence user behaviour and experience on TikTok, with specific attention to flow and user satisfaction. These insights will contribute to current debates on social media platforms' responsibility for users' digital well-being.

Based on previous research in different settings, such as general smartphone use (Lavoie & Zheng, 2023) We expect that the screentime nudge will lead to a momentary decrease in experienced flow during TikTok use. It is expected that the nudge interrupts the fluency aspect of flow. This results in the following hypothesis:

H1a: The presence of a screentime nudge will reduce the flow experience among TikTok users.

When it comes to the users' behaviour, the experience of flow has been related to social media use in the past, in that experiencing more flow leads to a higher frequency of social media use (Pelet, Ettis, & Cowart, 2017). Thus, we expect that increased flow experience will lead to greater TikTok use. In this study flow is interrupted; thus, we expect:

H1b: A reduced experience of flow leads to a reduction in TikTok use.

In previous studies, screentime nudges have shown mixed effects on behavioural outcomes, such as social media usage (Keller et al., 2024; Zimmermann, 2021).

Thus, in this study, we expect to find a reduction in TikTok use when the nudge is present. This results in the following hypothesis:

H1c: The presence of a screentime nudge will reduce the time spent on TikTok.

Regarding user satisfaction, we expect that the presence of a nudge might directly reduce user satisfaction, as reactance to the nudge overshadows the positive experience of scrolling on TikTok. Secondly, previous research has considered a type of "dark flow", where the usual positive relationship between flow and satisfaction is absent. Instead, a sort of "flow cost" is experienced, resulting in lower satisfaction levels (Lavoie & Main, 2019). This flow cost arises from the costs scrolling on social media imposes on other areas of life that are neglected as a result. This results in the following hypothesis:

H2a: The presence of a screentime nudge leads to reactance in users.

H2b: Reactance in users during scrolling leads to lower perceived user satisfaction with their scrolling on TikTok.

H2c: The presence of a screentime nudge negatively affects the perceived user satisfaction of scrolling on TikTok.

Some of these relationships may be moderated by users' intrinsic motivation to reduce their TikTok usage and their self-efficacy in changing their TikTok habits to support their digital well-being. We also expect the frequency of TikTok use before the intervention to influence certain relationships, since familiarity with the platform may make it more difficult to interrupt the flow it elicits. Thus, we propose;

H3: The effect of flow experience on TikTok usage will be weaker among frequent TikTok users.

H4: The effect of the nudge on TikTok usage will be stronger for highly

Methods

A 2x1 field experiment (screentime interruption nudge versus no nudge) will be conducted in spring 2026 to measure the effect of screentime nudges within the TikTok app. A field experiment allows for studying the intervention in an ecologically valid context, as participants use their own TikTok app with personally tailored, algorithmically recommended content. Participants are randomly assigned to one of two conditions: (a) an experimental group where the screentime interruption is turned on (i.e., screentime nudge) or (b) the control group that has no screentime reminders within their app. Participants will be recruited via social media and via participation for credit through a bachelor's-level course.

The study will last 2 weeks. Before the study begins, participants must complete an informed consent form. Afterwards, participants report their normal TikTok usage, and any potential moderators will be measured. Next, the intervention group is instructed on how to enable screentime reminders in their TikTok app. During the entire period, data is collected using the experience sampling method, with survey links to measure flow, reactance and user satisfaction (details see below) distributed to both groups via WhatsApp at regular intervals. Lastly, both groups report their screentime: once after the first intervention week, and once at the end of the intervention window.

Using the experience sampling method (Larson & Csikszentmihalyi, 1983), a comprehensive view of participants' real experiences will be captured. At random intervals during the intervention week, participants will be asked whether they have seen a screentime interruption nudge in the last 15 minutes, after which psychological flow, reactance, and user satisfaction will be measured via short surveys. Psychological flow experience will be measured through the validated two-dimensional

flow scale (TDFS), which measures the fluency and absorption elements of flow (Lavoie et al., 2022; Lavoie & Zheng, 2023). User satisfaction is measured through items that focus on perceived satisfaction while scrolling. Reactance to the nudge is measured through an adapted version of the Hong Reactance Scale (Hong & Faedda, 1996). TikTok use will be operationalised via time spent on TikTok per week, which can be objectively reported through screenshots from the TikTok app. These screenshots will be taken at the beginning of the research as a baseline measurement and twice during the intervention window. Furthermore, potential moderators are currently identified in the literature and will be included in the experiment if needed.

Data will be analysed using a repeated-measures ANOVA and the Hayes PROCESS macro to determine how much of the relationship between the nudge and TikTok use is explained by flow experience.

Results

Considering the data collection period in spring 2026, results will be discussed during the presentation.

Future Work

Future research recommendations will be added to the presentation based on the results of the study. However, we expect it might be relevant to include more objective measurements of flow in follow-up studies, such as eye-tracking or the Flow Observational Grid (Tordet et al., 2021). Since flow is an internal psychological process, self-report measures involve a certain level of recall bias on the part of participants.

References

- Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*: Harper & Row.
- Hong, S.-M., & Faedda, S. (1996). Refinement of the Hong Psychological Reactance Scale. *Educational and Psychological Measurement*, 56(1), 173-182. doi:10.1177/0013164496056001014
- Keller, J., Herrmann-Schwarz, T., Roitzheim, C., Mertens, L., Brockmeier, L., & Purohit, A. K. (2024). A Digital Nudge-Based Intervention to Interrupt Instagram Usage. *European Journal of Health Psychology*, 31, 128-140. doi:10.1027/2512-8442/a000150
- Larson, R., & Csikszentmihalyi, M. (1983). The Experience Sampling Method. *New Directions for Methodology of Social & Behavioral Science*, 15, 41-56.
- Lavoie, R., & Main, K. (2019). When losing money and time feels good: The paradoxical role of flow in gambling. 41, 53-72.

- Lavoie, R., Main, K., & Stuart-Edwards, A. (2022). Flow theory: Advancing the two-dimensional conceptualization. *Motivation and Emotion*, 46(1), 38-58. doi:10.1007/s11031-021-09911-4
- Lavoie, R., & Zheng, Y. (2023). Smartphone use, flow and wellbeing: A case of Jekyll and Hyde. *Computers in Human Behavior*, 138, 107442. doi:<https://doi.org/10.1016/j.chb.2022.107442>
- Pardes, A. (2018, May 9). Google and the rise of digital wellbeing. Retrieved from <https://www.wired.com/story/google-and-the-rise-of-digital-wellbeing/>
- Pelet, J.-É., Ettis, S., & Cowart, K. (2017). Optimal experience of flow enhanced by telepresence: Evidence from social media use. *Information & Management*, 54(1), 115-128. doi:<https://doi.org/10.1016/j.im.2016.05.001>
- Pham, G., & Duff, B. (2022). Regretful pleasure: Toward an understanding of flow cost in media use. *PLOS ONE*, 17, e0268194. doi:10.1371/journal.pone.0268194
- Tordet, C., Erhel, S., Wodey, E., Jamet, E., Nardi, N., & Gonthier, C. (2021). The Flow Observational Grid: an Observation-Based Solution to Assess Flow States. *Journal of Happiness Studies*, 22(7), 3069-3089. doi:10.1007/s10902-021-00356-7
- Vu, T. H., & Tagliabue, M. (2025). Active nudging towards digital well-being: reducing excessive screen time on mobile phones and potential improvement for sleep quality. *Front Psychiatry*, 16, 1602997. doi:10.3389/fpsy.2025.1602997
- Zhao, H., & Wagner, C. (2022). How TikTok leads users to flow experience: investigating the effects of technology affordances with user experience level and video length as moderators. *Internet Research*, 33(2), 820-849. doi:10.1108/INTR-08-2021-0595
- Zimmermann, L. (2021). "Your screen-time app is keeping track" : Consumers are happy to monitor but unlikely to reduce smartphone usage. *Journal of the Association for Consumer Research*, 6(3), 377-382. doi:10.1086/714365

7B.3 Predictive Approaches to Suicidal Thoughts in Spanish Youth: The Role of Online Media Engagement Patterns

Daniel Barredo Ibáñez (Universidad de Sevilla), Lucía Caro Castaño (Universidad de Cádiz)

Introduction

Social media not only facilitate new ways of connecting with other users but also allow the development of all kinds of strategies for monitoring one's own social performance (popularity, centrality), relationships with other contacts, and comparison with their self-presentation performances. The existence of social metrics that synthesise, in quantitative terms, the amalgam of affective expressions produced on profiles within these platforms has contributed to new practices of social surveillance and self-surveillance, including an 'imagined surveillance' driven by the idea of being observed by unforeseen audiences on social media who evaluate us (Duffy & Chan, 2019). These metrics are interpreted by adolescents to evaluate the social validation they receive and to monitor their own evolution on the platform (Caro-Castaño, 2023). Various studies on social media have detected anxious behaviours by children and adolescents in relation to the obligation they perceive to keep their profile active for self-presentation (Boyd, 2015; Mascheroni et al., 2015), but also in relation to the tasks they perform with other contacts (liking, commenting, recommending) so that these act reciprocally and thus achieve a certain volume in the social metrics of their publications (Turtle, 2011; Chambers et al., 2018), even engaging in self-promotion practices (Caro-Castaño, 2022; Rivas-Herrero & Igartua, 2021).

Research into the impact of social media on the mental health of children and youth is a growing concern, given the omnipresence of these digital environments in their socialization processes (Boyd, 2015; Houghton et al., 2018). Yet findings remain contradictory (Perlmutter et al., 2024; Valkenburg, 2022). This inconsistency stems from the shifting, complex nature of social media platforms (Bayer et al., 2020; Sanders et al., 2023), the varied uses and gratifications each user finds in them (Ferris et al., 2021; Metwally et al., 2025), and the individual traits of users (Valkenburg et al., 2024). These digital technologies tends to amplify users' preexisting mental states. Franchina et al. (2018) found that when metrics did not reach the volume expected by adolescents, the respondents experienced a negative impact on their self-esteem. It should also be considered that content related to self-image and corporality on social media constitutes the factor that most affects the mental health of this age group (Mulisa et al., 2025).

Objectives/Research Questions

Although the quoted works remark the effects derived by the individual online media usage, previous research notes that “time spent” does not capture the complexity of user behaviour and cannot, on its own, explain problematic patterns such as addiction (Andreassen, 2015). Other studies have examined how time spent on social media affects the overall range of activities in which young people engage. In this vein, Bozzola et al. (2022) highlighted how excessive use impacted adolescents’ sleep patterns and reduced the time available for offline activities beneficial to their mental health. Likewise, Franchina et al. (2018), in their exploratory study on the connections between FOMO (fear of missing out) (Przybylski et al., 2013), social media use, and phubbing behaviour among adolescents, found that FOMO predicted phubbing—in other words, the greater the FOMO, the stronger the tendency to focus on the phone rather than engage in face-to-face interactions. Ma et al. (2025), in a survey of young adults, identified a chained mediating effect involving daytime sleepiness and sleep quality in the relationship between social media use and SI, highlighting poor sleep quality as a risk factor for suicidal thoughts. In a similar way, Kennard et al. (2025), in their study of youth receiving support for depression, SI, and SC in Texas, found that 40.3% of participants reported problematic social media use. These individuals were more likely to report higher screen time and, compared to those who did not report problematic use, displayed more severe depressive symptoms, anxiety, suicidal thoughts, and poorer well-being indicators. Following these authors, we propose, as a general hypothesis:

H1: Online media usage increases SI among Spanish children when they are actively engaged in these spaces. However, this relationship is intensified by perceived direct cybervictimization and moderated by age, with a stronger effect observed when the children are more exposed online.

Our dependent variable is one of the mentioned effects on mental health: SI, which is one of the most important indicators of a potential suicidal act (Vargas & Saavedra, 2012). However, we posit that active online engagement may predict an increase in SI, as several authors have highlighted the importance of distinguishing between active and passive social media use when assessing its effects on users’ well-being and mental health (Keles et al., 2020; Kingsbury et al., 2021; Pang, 2021; Thorisdottir et al., 2019; Valkenburg et al., 2022; Verduyn et al., 2017; 2022). Verduyn et al. (2022), for their part, propose a model of active and passive social media use and its relationship with user well-being, which refines the framework they originally introduced in 2017 (Verduyn et al., 2017).

Methods

This is a non-experimental study, with a correlational scope, and a transversal extraction of data. The target population of this study consists of boys and girls aged 10 to 15 residing in Spain. All data were collected through a printed survey, completed by

hand by the surveyed children. To confirm their participation, the polling company obtained informed consent signed by the minors' legal guardians. Before the questionnaire was administered in the classroom by trained specialists, the children were read an informational assent, which included the option to leave questions blank if they did not wish to share their opinions or participate in the survey. A total of 309 schools were contacted, resulting in participation from 17 schools and 67 classrooms. 804 children participated in the survey. Suicidal ideation was measured according to the Paykel Scale (Paykel et al., 1971). This is a dichotomous scale with five variables. To check the internal reliability of this construct, we used the Kuder–Richardson coefficient (KR20), a cohesion test for dichotomous variables. For five elements, KR20 reported a coefficient of 0.736, perceived highly acceptable (Betancourt & Caviedes, 2018).

The second construct was measured using the Cybervictimization Scale (Buelga, 2019). This study specifically examines variables associated with DCV, assessed through eight items on a five-point Likert frequency scale. These items evaluate the victim's perceived experiences of DCV, such as insults or exclusion from social groups. The Cronbach's alpha for the eight items was 0.889, indicating high reliability.

Our third construct is based in the scale of Attitudes and Uses of Media and Technology, validated by Rosen et al. (2013), previously tested in children with a similar profile (Rosen et al., 2014), and partially adapted by the study's researchers. This scale measures 50 different attitudes and uses, but according to the characteristics of our sample, we only selected fourteen variables related to uses of media and technology. These questions were measured using a Likert scale of frequency. After data analysis, and to ensure the reliability of the study, we had to exclude two variables. The first, 'I play video games with violent content on a computer, game console, or smartphone', was too broad and did not provide a contrast with other types of video games. The second, 'I comment on social media posts, status updates, photos, etc.', was excluded due to high cross-loadings with two factors in the factorial analysis. The complexity of analysing this variable may be related to the diverse nature of the comments made by the surveyed children. Some comments could be considered emotional reactions (e.g., posting a smiley face in response), while others were more narrative in nature. After excluding these two variables, the remaining twelve variables achieved a Cronbach's alpha of 0.862, considered highly acceptable.

Results

Approximately one in three surveyed children in Spain exhibit a medium or high risk of suicidal ideation (28.7%, $n=222$), while the majority present low or non-existent scores (69.4%, $n=504$). In the second construct, DCV, the most frequent or very frequent DCV items are related to the reception of phone calls where the caller remained silent (11.1%, $n=89$), the feeling of being ignored or unanswered by groups (9.3%, $n=75$), and

the sharing of secrets or personal information in groups without permission (8.8%, $n=71$). The most frequent attitudes and uses of media and technology reported by the surveyed children in Spain are: listening to music on my mobile phone (81.6%, $n=656$), reading posts on social media (63.5%, $n=510$), clicking 'like' on social media posts, photos, etc. (51.6%, $n=415$), and browsing friends' or family members' profiles and photos on social media from my mobile phone (46.2%, $n=371$). Exploratory factor analysis with Varimax rotation was conducted to identify the latent factors associated with the twelve variables described in the previous section. Three distinct factors emerged that describe the attitudes and usage patterns of media and technology among the surveyed children in Spain, which are in line with the ASMU and PSMU explained models (Verduyn et al., 2022). These factors explain 62.461% of the variance after rotation, which exceeds the minimum acceptable percentage of 60% (Frías & Pascual, 2012).

Based on their attitudes and patterns of media use, we identified three distinct user profiles among the participants: lurkers—children who primarily consume online content; sharers—those who actively create and share content on social media; and interactors—users who engage in co-oriented interactions with their online contacts. Results indicate that both sharers and interactors are approximately twice as likely to experience an increase in suicidal ideation compared to lurkers, particularly when they are older and report higher levels of direct cybervictimization. To test the hypothesis model, we conducted a moderated mediation analysis. We have confirmed the indirect effects model proposed by H1: online media usage increases SI among Spanish children when they are intensively engaged in these spaces, but this relationship is intensified by perceived DCV and moderated by age, with a stronger effect observed among children with greater online exposure. This effect is more pronounced for sharers ($B = .0772$, $SE = .0177$, 95% CI [.0477, .1166]) and interactors ($B = .0698$, $SE = .0177$, 95% CI [.0410, .1093]) compared to lurkers ($B = .0324$, $SE = .0074$, 95% CI [.0207, .0491]), who face approximately half the likelihood of an increase in SI compared to the other two media and technology usage profiles. These findings may relate to individual, social, and platform affordance factors, such as heightened susceptibility to peer comparison, new forms of social surveillance and self-monitoring, and social media metrics.

Future Work

The results presented have important practical implications. Younger individuals face many restrictions when creating or managing accounts on online platforms, resulting in their online media usage being more focused on passive experiences, which are not without risks. Rather than focusing solely on restricting the use of digital environments, it is important to consider strategies that help reduce these risks, such as digital education and parental and teacher supervision of the content children are exposed to

at home and in schools. Moreover, at the family level, digital disconnection should be considered for preventive purposes (Liu et al., 2022; O'Reilly et al., 2018). Simultaneously, media education should be reinforced to help children critically decode social media content (Reyes & Salvarini, 2023) and discourage the replication of pro-suicidal content across these channels. Education plays an essential role in addressing cybervictimization, mental health disorders, and suicide, as emphasized by Reyes and Salvarini (2023), contributing to a comprehensive and positive transformation of society.

As a main limitation, the presented results are based on the application of a national survey in Spain. In this regard, and as a recommendation for future studies, it is worth noting—given that the observed effects are statistically significant—the need to explore whether these patterns are replicated in other contexts or in longitudinal studies that would allow for an assessment of their impact over time. At the same time, gender was not considered in the analysis, despite previous studies having identified it as a relevant factor in understanding the effects of social media on young people's mental well-being (Kerr & Kingsbury, 2023; Steinsbekk et al., 2021).

References

- Andreassen C. S. (2015). Online social network addiction: A comprehensive review. *Current Addiction Reports*, 2, 175–184. <https://doi.org/10.1007/s40429-015-0056-9>
- Bayer, J. B., Triêu, P., & Ellison, N. B. (2020). Social media elements, ecologies, and effects. *Annual review of psychology*, 71(1), 471-497. <https://doi.org/10.1146/annurev-psych-010419-050944>
- Betancourt, A.C., & Caviedes, I.L. (2018). Metodología de correlación estadística de un sistema integrado de gestión de la calidad en el sector salud. *Signos*, 10, 119–139. <https://doi.org/10.15332/s2145-1389.2018.0002.07>
- boyd, d. (2015). *It's Complicated: The Social Lives of Networked Teens*. Yale University Press.
- Bozzola, E., Spina, G., Agostiniani, R., Barni, S., Russo, R., Scarpato, E.; Di Mauro, A., Di Stefano, A.V., Caruso, C., Corsello, G., & Staiano, A. (2022). The Use of Social Media in Children and Adolescents. Scoping Review on the Potential Risks. *IJERPH*, 19. <https://doi.org/10.3390/ijerph19169960>
- Buelga, S.; Martínez, B.; Cava, M. J. & Ortega, J. (2019). Psychometric Properties of the CYBVICS Cyber-Victimization Scale and Its Relationship with Psychosocial Variables. *Social Sciences*, 8, 1-13. <https://doi.org/10.3390/socsci8010013>

- Caro-Castaño, L. (2023). Prácticas postfotográficas en los medios sociales. Un estudio transnacional sobre el uso de las historias de Instagram. *Methaodos*, 11(2), m231102a15. <https://doi.org/10.17502/mrcs.v11i2.723>
- Caro-Castaño, L. (2022). Playing to be influencers: A comparative study on Spanish and Colombian young people on Instagram. *Communication & Society*, 35(1), 81-99. <https://doi.org/10.15581/003.35.1.81-99>
- Chambers, D., Cairns, K., & Ivancic, L. (2018). Young people, the internet and mental health. *Irish journal of psychological medicine*, 35(1), 1-4. <https://doi.org/10.1017/ipm.2017.76>
- Duffy, B. E., & Chan, N. K. (2019). “You never really know who’s looking”: Imagined surveillance across social media platforms. *New Media & Society*, 21(1), 119-138. <https://doi.org/10.1177/1461444818791318>
- Ferris, A. L., Hollenbaugh, E. E., & Sommer, P. A. (2021). Applying the Uses and Gratifications Model to Examine Consequences of Social Media Addiction. *Social Media + Society*, 7(2). <https://doi.org/10.1177/20563051211019003>
- Franchina, V., Vanden Abeele, M., van Rooji, A. J., Lo Coco, G., & De Marez, L. (2018). Fear of missing out as a predictor of problematic social media use and phubbing behavior among flemish adolescents. *International Journal of Environmental Research and Public Health*, 15(10), 2319. <https://doi.org/10.3390/ijerph15102319>
- Frías, D., & Pascual, M. (2012). Exploratory Factor Analysis (EFA) in Consumer Behavior and Marketing Research. *Suma Psicológica*, 1(19), 47-58. <https://tinyurl.com/mv3vfh5s>
- Houghton, S., Lawrence, D., Hunter, S. C., Rosenberg, M., Zadow, C., Wood, L., & Trevor Shilton (2018). Reciprocal Relationships between Trajectories of Depressive Symptoms and Screen Media Use during Adolescence. *Journal of Youth and Adolescence* 47, 2453–2467. <https://doi.org/10.1007/s10964-018-0901-y>
- Keles, B., McCrae, N., & Grealish, A. (2020). A systematic review: the influence of social media on depression, anxiety and psychological distress in adolescents. *International journal of adolescence and youth*, 25(1), 79-93. <https://doi.org/10.1080/02673843.2019.1590851>
- Kennard, B. D., Hughes, J. L., Minhajuddin, A., Jones, S. M., Jha, M. K., Slater, H., Mayes, [...] (2025). Problematic social media use and relationship to mental health characteristics in youth from the Texas Youth Depression and Suicide Research Network (TX-YDSRN). *Journal of affective disorders*, 374, 128-140. <https://doi.org/10.1016/j.jad.2025.01.046>

- Kerr, S., & Kingsbury, M. (2023). Online digital media use and adolescent mental health. *Health reports*, 34(2), 17-28. <https://www.doi.org/10.25318/82-003-x202300200002-eng>
- Kingsbury, M., Reme, B. A., Skogen, J. C., Sivertsen, B., Øverland, S., Cantor, N., Hysing, M., Petrie, K., & Colman, I. (2021). Differential associations between types of social media use and university students' non-suicidal self-injury and suicidal behavior. *Computers in human behavior*, 115, 106614. <https://doi.org/10.1016/j.chb.2020.106614>
- Liu, M., Kamper-DeMarco, K.E., Zhang, J., Xiao, J., Dong, D., & Xue, P. (2022). Time Spent on Social Media and Risk of Depression in Adolescents: A Dose–Response Meta-Analysis. *International Journal of Environmental Research and Public Health*, 19(9), 5164. <https://doi.org/10.3390/ijerph19095164>
- Ma, Y. B., Zheng, Z. A., Yao, Z. Y., Xu, X. M., Zhou, X. Y., Kou, C. G., Yao, B., Sun, W. J., Li, R., Gong, & Jia, C. X. (2025). The effect of social media use on suicidal ideation in college students: Mediation by daytime sleepiness and sleep quality. *Journal of Affective Disorders*, 374, 274-281. <https://doi.org/10.1016/j.jad.2025.01.049>
- Mascheroni, G., Vincent, J., & Jiménez, E. (2015). “Girls are addicted to likes so they post semi-naked selfies”: Peer mediation, normativity and the construction of identity online. *Cyberpsychology: Journal of psychosocial research on cyberspace*, 9(1), 5. <https://doi.org/10.5817/CP2015-1-5>
- Metwally, D., Bakari, H., & Manzoor, A. (2025). Social and psychological costs of problematic use of social media: users and gratification perspective. *Cogent Psychology*, 12, 1, 2467513. <https://doi.org/10.1080/23311908.2025.2467513>
- Mulisa, F., Kejela, M., & Dido, J. S. (2025). Social Media Trap: A Systematic Review of the Content and Pathways that Affect Adolescents’ Mental Health. *Journal of Technology in Behavioral Science*, 1-19. <https://doi.org/10.1007/s41347-025-00502-5>
- O’Reilly, M., Dogra, N., Whiteman, N., Hughes, J., Eruyar, S., & Reilly, P. (2018). Is social media bad for mental health and wellbeing? Exploring the perspectives of adolescents. *Clinical Child Psychology and Psychiatry*, 23(4), 601-613. <https://doi.org/10.1177/1359104518775154>
- Pang, H. (2021). Unraveling the influence of passive and active WeChat interactions on upward social comparison and negative psychological consequences among university students. *Telematics and informatics*, 57, 101510. <https://doi.org/10.1016/j.tele.2020.101510>

- Paykel, E. S., Prusoff, B. A., & Uhlenhuth, E. H. (1971). Scaling of life events. *Archives of general psychiatry*, 25(4), 340-347.
<https://doi.org/10.1001/archpsyc.1971.01750160052010>
- Perlmutter, E., Dwyer, B., & Torous, J. (2024). Social media and youth mental health: assessing the impact through current and novel digital phenotyping methods. *Current Treatment Options in Psychiatry*, 11(2), 34-51.
<https://doi.org/10.1007/s40501-024-00312-1>
- Przybylski, A. K., Murayama, K., DeHaan, C. R., & Gladwell, V. (2013). Motivational, emotional, and behavioral correlates of fear of missing out. *Computers in human behavior*, 29(4), 1841-1848. <https://doi.org/10.1016/j.chb.2013.02.014>
- Reyes, I., & Salvarini, S. (2023). Child-youth mental health: What is the Rol of Pedagogy to face this Pandemic? *Revista de Orientación Educacional*, 37(72), 57-85.
<https://tinyurl.com/26bpdrrky>
- Rivas-Herrero, B., & Igartua, J.J. (2021). Los jóvenes se sumergen en Instagram. Un estudio desde la perspectiva de los usos y gratificaciones. *Profesional de la Información*, 30(5), e300409. <https://doi.org/10.3145/epi.2021.sep.09>
- Rosen, L. D., Whaling, K., Carrier, L. M., Cheever, N. A., & Rokkum, J. (2013). The Media and Technology Usage and Attitudes Scale: An empirical investigation. *Computers in Human Behavior*, 29, 2501-2511.
<http://dx.doi.org/10.1016/j.chb.2013.06.006>
- Rosen, L. D., Lim, A. F., Felt, J., Carrier, L. M., Cheever, N. A., Lara-Ruiz, J. M., Mendoza, J. S., & Rokkum, J. (2014). Media and technology use predicts ill-being among children, preteens and teenagers independent of the negative health impacts of exercise and eating habits. *Computers in Human Behavior*, 35, 364-375.
<http://dx.doi.org/10.1016/j.chb.2014.01.036>
- Sanders, T., Noetel, M., Parker, P., Del Pozo Cruz, B., Biddle, S., Ronto, R., Hulteen, R., et al. (2023). An umbrella review of the benefits and risks associated with youths' interactions with electronic screens. *Nature Human Behaviour*, 8, 1–18.
<https://doi.org/10.1038/s41562-023-01712-8>
- Steinsbekk, S., Wichstrøm, L., Stenseng, F., Nesi, J., Hygen, B. W., & Skalická, V. (2021). The impact of social media use on appearance self-esteem from childhood to adolescence—A 3-wave community study. *Computers in Human Behavior*, 114, 106528. <https://doi.org/10.1016/j.chb.2020.106528>
- Thorisdottir, I. E., Sigurvinsdottir, R., Asgeirsdottir, B. B., Allegrante, J. P., & Sigfusdottir, I. D. (2019). Active and passive social media use and symptoms of anxiety and depressed mood among Icelandic adolescents. *Cyberpsychology, Behavior, and Social Networking*, 22(8), 535-542. <https://doi.org/10.1089/cyber.2019.0079>

Turkle, S. (2011). *Alone Together. Why we expect more from technology and less from each other.* Basic Books.

Valkenburg, P. M. (2022). Social media use and well-being: What we know and what we need to know. *Current opinion in psychology*, 45, 101294.

<https://doi.org/10.1016/j.copsyc.2021.12.006>

Valkenburg, P. M., Beyens, I., Bij de Vaate, N., Janssen, L., & van der Wal, A. (2024). "Person-specific media effects". In T. Araujo & P. Neijens (Eds.), *Communication research into the digital society: Fundamental insights from the Amsterdam School of Communication Research* (pp. 233-245). Amsterdam University Press.

https://doi.org/10.5117/9789048560592_ch14

Vargas, H. B., & Saavedra, J. E. (2012). Factors Associated with Suicidal Behavior in Adolescents. *Revista de Neuro-Psiquiatría*, 75(1), 19-28.

<https://tinyurl.com/25ee6vyr>

Verduyn, P., Gugushvili, N., & Kross, E. (2022). Do Social Networking Sites Influence Well-Being? The Extended Active-Passive Model. *Current Directions in Psychological Science*, 31(1), 62-68.

<https://doi.org/10.1177/09637214211053637>

Verduyn, P., Ybarra, O., Résibois, M., Jonides, J., & Kross, E. (2017). Do social network sites enhance or undermine subjective well-being? A critical review. *Social Issues and Policy Review*, 11(1), 274–302. <https://doi.org/10.1111/sipr.12033>

7B.4 The Many Faces of Stigma: Profiles of Adolescent Mental Illness Stigma Beliefs and Social Media Use Across Switzerland, Germany, and Austria

Dora Weubel, Sophie Mayen (USI Università della Svizzera Italiana), Kathrin Karsay (University of Vienna), Ruth Wendt (Ludwig Maximilian University), Anne-Linda Camerini (USI Università della Svizzera Italiana)

Introduction

Adolescent mental health problems are increasingly common, affecting nearly one in five adolescents in Europe (Sacco et al., 2024). They can have long-term consequences for education, employment, physical health, and life expectancy (Membride, 2016), making early recognition and support crucial. However, many adolescents delay or avoid seeking help. A major barrier is stigma, alongside negative attitudes and perceptions toward mental illness (Barrow & Thomas, 2022; Competiello et al., 2023; Lu et al., 2023).

Stigma is shaped through adolescents' everyday communication both offline and online. Nowadays, social media has become a central space where adolescents encounter mental illness-related narratives and social cues about what is “normal”, “acceptable”, or “desirable” (Chandra & Minkovitz, 2007; Competiello et al., 2023; Patton et al., 2016; Wahl, 1995).

Past research on mental illness stigmas focused on dimensions such as viewing people with mental illness as weak-not-sick, dangerous and unpredictable, or socially undesirable (Jorm & Wright, 2008; Lu et al., 2023; Yap et al., 2014). More recently, romanticization of mental illness has gained attention. It describes portraying people with mental health problems as mysterious, special, or desirable (Ahuja & Fichadia, 2024; Jadayel et al., 2017). This trend is especially present on social media platforms like TikTok and Instagram showing content that aestheticize or idealize mental illness (Issaka et al., 2024; Ndour & Foulkes, 2025). While this trend may be seen as destigmatizing in the first place, it blurs the boundaries between temporary mood states and severe mental health problems, and thus contributes to trivialized views of mental illness, which is yet another form of stigmatization (Ahuja & Fichadia, 2024). Against this backdrop, research on mental illness stigmas needs to be updated.

Objectives/Research Questions

The goal of this study is to identify distinct profiles of mental illness stigmas among adolescents in today's social media landscape. To do so, we adopt a person-centered approach to explore which stigmas prevail, and in which combination. In addition, we

aim to understand how online and offline communication contexts of adolescents shape these profiles, with a specific attention to social media use and exposures to mental illness content on social media.

Methods

We used data from T1 (Nov 2025–Jan 2026) of a three-wave, school-based online survey of adolescents aged 12–19 in German-speaking Switzerland, Austria, and Germany. After obtaining ethical clearance, participants were recruited through collaboration with regional education administrations. Questionnaires were completed during a school hour on either a computer or tablet using SoSci Survey (Leiner, 2026).

To measure stigma, participants read a brief vignette describing “Alex,” a 16-year-old adolescent experiencing depressive symptoms. The vignette was gender-matched (female/male), with random assignment for participants identifying as diverse. Participants then rated 9 personal stigma items adapted from prior vignette-based stigma work (Lu et al., 2023) on a 5-point Likert scale (1 = completely disagree to 5 = completely agree). Items captured classic stigma beliefs (e.g., weak-not-sick; dangerous/unpredictable) and romanticizing stigma beliefs (e.g., perceiving the person as “mysterious” or having “deeper insights”). To describe stigma profiles, we used a measure of daily social media use (scale from 0 to 12 hours), incidental and intentional exposure to mental illness content (1 = never to 5 = very often), evaluation of such content (e.g., un/trustworthy, un/interesting, bad/good), and frequency of offline communication about mental illness with friends, classmates, and family (1 = never to 5 = very often).

Analyses were carried out in R (R: The R Project for Statistical Computing, 2025). After data cleaning, missing data in stigma items were imputed using the mice package. Next, stigma items were recoded into three ordinal levels (low = 1–2, medium = 3, high = 4–5) and treated as categorical indicators. We conducted Latent Class Analysis (LCA) using the polCA package, estimating 1 to 6 class solutions and selecting the optimal number by inspection of AIC, BIC, and Entropy, and interpretability. Lastly, we compared stigma profiles by evaluating the distribution of demographic characteristics, social media use, and offline communication about mental illness.

Results

The analytical sample included $N = 1,444$ adolescents ($M_{age} = 14.4$, $SD = 1.8$). Most identified as female (51.1%), followed by male (45.4%) and diverse (0.8%), while 2.7% did not report their gender.

The LCA supported a four-class solution:

Low Stigma ($n = 401$): This class showed the lowest overall endorsement of stigma across all items, except for romanticizing stigma beliefs (e.g., “People like Alex have

deeper insights”). This class includes mostly females (69 % female), with an average age of 15 years ($M = 14.73$, $SD = 1.83$). On average, social media use is moderate ($M = 7.66$, $SD = 4.78$), ~3h/day). However, this class reported the highest levels of both incidental ($M = 2.54$, $SD = 1.11$) and intentional exposure ($M = 1.66$, $SD = 1.01$) to mental illness content on social media. Participants also report to more frequently talk about mental illness with friends ($M = 2.17$, $SD = 1.19$) than with family ($M = 1.87$, $SD = 1.06$) or classmates ($M = 1.45$, $SD = 0.79$).

Moderate stigma ($n = 56$): The smallest class showed moderate stigma beliefs across all items. Most of the class members are male (51.8%). On average, this class reported the highest social media use ($M = 9.34$, $SD = 6.47$; ~4h/day) and the most frequent communication about mental illness (e.g., family: $M = 2.12$, $SD = 1.27$; classmates: $M = 1.75$, $SD = 1.00$).

High Stigma ($n = 235$): Established negative stigmas (weak-not-sick; dangerous/unpredictable) were endorsed most strongly in this class. Romanticizing stigma beliefs were also high. This class is predominantly male (68.9%), with a mean age of 14.18 years ($SD = 1.70$). Despite high social media use ($M = 8.20$, $SD = 5.57$, ~4h/day), the the class reported the lowest incidental ($M = 2.24$, $SD = 1.16$) and intentional ($M = 1.45$, $SD = 0.88$) exposure to mental illness content on social media. Also, they less frequently talked about mental illness in their social context (e.g., friends: $M = 1.97$, $SD = 1.30$; classmates: $M = 1.55$, $SD = 0.94$).

Romanticizing Stigma ($n = 752$): This is the largest class, and it is marked by strong endorsement of romanticizing stigma beliefs, while established stigma beliefs were only moderately endorsed. Gender distribution in this class is balanced (49.7% female, 46.7% male), and the mean age is 14.38 years ($SD = 1.80$). On average, class members reported moderate ($M = 7.86$, $SD = 4.77$, ~3h/day) overall social media use. Communication about mental illness was somewhat limited (e.g., classmates: $M = 1.49$, $SD = 0.82$; friends: $M = 2.02$, $SD = 1.17$).

Future Work

We reported preliminary findings as data analysis is still ongoing. Next steps will include: (1) test the robustness of the class solution (e.g., alternative indicator coding, sensitivity checks), (2) examine a broader set of correlates (e.g., sociodemographics, social norms, group identity), and (3) conducting ANOVAS to test for significant differences in class describing variables.

By adopting a person-centered approach, the present study goes beyond pre-defined stigma beliefs and gives a nuanced understanding of how adolescents think about people with a mental illness in contemporary society shaped by public discourse about mental illness on social media. By providing different stigma profiles, it can inform future, targeted destigmatization efforts.

References

- Ahuja, J., & Fichadia, P. A. (2024). Concerns Regarding the Glorification of Mental Illness on Social Media. *Cureus*. <https://doi.org/10.7759/cureus.56631>
- Barrow, E., & Thomas, G. (2022). Exploring perceived barriers and facilitators to mental health help-seeking in adolescents: A systematic literature review. *Educational Psychology in Practice*, 38(2), 173–193. <https://doi.org/10.1080/02667363.2022.2051441>
- Chandra, A., & Minkovitz, C. S. (2007). Factors that Influence Mental Health Stigma Among 8th Grade Adolescents. *Journal of Youth and Adolescence*, 36(6), 763–774. <https://doi.org/10.1007/s10964-006-9091-0>
- Competiello, S. K., Bizer, G. Y., & Walker, D. C. (2023). The Power of Social Media: Stigmatizing Content Affects Perceptions of Mental Health Care. *Social Media + Society*, 9(4), 20563051231207847. <https://doi.org/10.1177/20563051231207847>
- Issaka, B., Aidoo, E. A. K., Wood, S. F., & Mohammed, F. (2024). “Anxiety is not cute” analysis of twitter users’ discourses on romanticizing mental illness. *BMC Psychiatry*, 24(1), 221. <https://doi.org/10.1186/s12888-024-05663-w>
- Jadayel, R., Medlej, K., & Jadayel, J. J. (2017). Mental Disorders: A Glamorous Attraction On Social Media?
- Jorm, A. F., & Wright, A. (2008). Influences on young people’s stigmatising attitudes towards peers with mental disorders: National survey of young Australians and their parents. *British Journal of Psychiatry*, 192(2), 144–149. <https://doi.org/10.1192/bjp.bp.107.039404>
- Leiner, D. J. (2026). SoSci Survey (Version 3.5.02) [Computer software]. <https://www.soscisurvey.de>.
- Lu, S., Hart, L. M., Jorm, A. F., Gregg, K., Gross, M., Mackinnon, A. J., & Morgan, A. J. (2023). Adolescent peer support for mental health problems: Evaluation of the validity and reliability of the Mental Health Support Scale for Adolescents. *BMC Psychology*, 11(1), 193. <https://doi.org/10.1186/s40359-023-01228-w>
- Membride, H. (2016). Mental health: Early intervention and prevention in children and young people. *British Journal of Nursing*, 25(10), 552–557. <https://doi.org/10.12968/bjon.2016.25.10.552>
- Ndour, A., & Foulkes, L. (2025). The romanticisation of mental health problems in adolescents and its implications: A narrative review. *European Child &*

Adolescent Psychiatry, 34(8), 2297–2326. <https://doi.org/10.1007/s00787-025-02701-0>

Patton, G. C., Sawyer, S. M., Santelli, J. S., Ross, D. A., Afifi, R., Allen, N. B., Arora, M., Azzopardi, P., Baldwin, W., Bonell, C., Kakuma, R., Kennedy, E., Mahon, J., McGovern, T., Mokdad, A. H., Patel, V., Petroni, S., Reavley, N., Taiwo, K., ... Viner, R. M. (2016). Our future: A Lancet commission on adolescent health and wellbeing. *The Lancet*, 387(10036), 2423–2478. [https://doi.org/10.1016/S0140-6736\(16\)00579-1](https://doi.org/10.1016/S0140-6736(16)00579-1)

R: The R Project for Statistical Computing. (2025). <https://www.r-project.org/>
Sacco, R., Camilleri, N., Eberhardt, J., Umla-Runge, K., & Newbury-Birch, D. (2024). A systematic review and meta-analysis on the prevalence of mental disorders among children and adolescents in Europe. *European Child & Adolescent Psychiatry*, 33(9), 2877–2894. <https://doi.org/10.1007/s00787-022-02131-2>

Wahl, O. F. (1995). *Media madness: Public images of mental illness*. Rutgers University Press.

Yap, M. B. H., Mackinnon, A., Reavley, N., & Jorm, A. F. (2014). The measurement properties of stigmatizing attitudes towards mental disorders: Results from two community surveys. *International Journal of Methods in Psychiatric Research*, 23(1), 49–61. <https://doi.org/10.1002/mpr.1433>

Session 7C - Infrastructure, Platformisation and Datafication

Day 2 · Tues 14 July · 11:20-12:50, Seminar Suite B, ARC

7C.1 Platform culture and Traditional Performance: Algorithmic Mediation of Ghoomar Dance on Instagram

Bhavani Singh (Madan Mohan Malviya University of Technology), Rashmi Jha (Banaras Hindu University)

Introduction

This paper explores the historical connote of Ghoomar and its current presence on social media platforms like Instagram and explore the idea of Platform culture. Ghoomar is a traditional Rajasthani folk dance with origins in Bhil tribal communities, historically performed in ritualistic veneration of Goddess Saraswati and later embraced by Rajput aristocracy. Etymologically derived from "ghoomna" (to revolve), Ghoomar is characterized by graceful circular rotational movements wherein performers execute synchronized pirouettes in concentric formations, their voluminous ghagras creating visual spirals through centrifugal force. These patterns represent agricultural cycles, cosmic revolutions, generational continuities, and the Hindu concept of sansar (Kothari, 1968; Vatsyayan, 1980). Traditionally performed by women during weddings, Gangaur and Teej festivals, Ghoomar functions as a "kinesthetic archive," encoding cosmological worldviews and gendered epistemologies through embodied movement (Vatsyayan, 1980). The performative grammar comprises: lahak (lateral arm swaying), synchronized rotations, tatkar (percussive footwork), bhamari (torso articulations), and controlled breathing. These elements, accompanied by folk songs (ghoomar geet) in regional dialects and percussive instrumentation (dholak, manjira, tabla), constitute a multimodal semiotic system articulating cultural meaning (Chakraborty, 2009). Costume assemblages operate as critical signifying practices. Dancers wear intricately embroidered ghagras in culturally prescribed hues—crimson red (fertility), saffron orange (devotion), magenta pink (youth) indexing ceremonial contexts. The odhni (head veil) enacts modesty norms while creating visual effects during rotations. Traditional silver jewelry borla, nath, aad, bajuband adds visual opulence and acoustic dimensions. Performance spaces have historically been intimate courtyards (aangan) of havelis, village commons, women's quarters (zenana) transmitting cultural knowledge within specific spatial-temporal parameters (Taylor,

2003). Pedagogical transmission occurs through observation-based learning within familial networks, emphasizing tacit knowledge (Lave & Wenger, 1991). Ghoomar's emergence on Instagram represents a significant transformation. Instagram has enabled this dance form to transcend historical geographic, caste, and community boundaries, allowing diverse practitioners—heritage custodians, professional dancers, Rajasthani performers, pan-Indian and diaspora enthusiasts—to engage with and share Ghoomar with global audiences. This digital migration reconfigures circulation patterns, accessibility, and evolutionary trajectory. What was transmitted through intimate familial observation and performed within bounded community spaces now circulates through algorithmically mediated recommendation systems, reaching audiences beyond Rajasthani cultural contexts.

Objectives/Research Questions

This research examines the co-creative process through which content creators and algorithmic systems jointly shape Ghoomar on Instagram. The study understands how traditional aesthetic elements—costume symbolism, movement vocabularies, musical accompaniment, spatial formations—are strategically foregrounded or adapted in response to platform affordances, and which elements undergo transformation as creators optimize for algorithmic legibility. Additionally, the research explores how Instagram enables democratization of access to Ghoomar practice and viewership beyond traditional community boundaries while creating new aesthetic conventions influenced by platform visibility requirements. This paper interrogates the ways in which Instagram and its content creators actively construct meanings around Ghoomar within the platformed ecology of social media. Rather than treating the platform as a neutral medium, the questions foreground Instagram's affordances as meaning-making tools that shape how Ghoomar is visualised, narrated, and circulated.

RQ1: How do Instagram's platform affordances reshape the temporal, spatial, and aesthetic dimensions of traditional Ghoomar performance?

RQ2: How do content creators strategically adapt Ghoomar aesthetics to align with algorithmic visibility while negotiating traditional cultural values?

RQ3: What new aesthetic forms emerge from creator-algorithm interaction, and how do these transformations reconfigure which traditional elements are emphasized, maintained, or backgrounded?

Methods

This study employs mixed methods combining digital ethnography with semi-structured interviews, triangulating platform content analysis with practitioner experiences.

Secondary Data ethnography Analysis: Following Kozinets (2020), I conducted six months of systematic observation (August 2025 - January 2026) of Ghoomar content on

Instagram. The approach involved collection and analysis of posts utilizing key hashtags (#Ghoomar, #GhoomarDance, #RajasthaniDance), examining video content, aesthetic choices, engagement patterns, algorithmic recommendations, and community interactions. Analysis focused on how traditional elements (costume, movement vocabulary, musical accompaniment, spatial arrangements) are represented or adapted; which platform features (Reels, Stories) are utilized; what role trending audio, filters, and editing play; and which content achieves higher algorithmic visibility. Primary Data - Semi-Structured Interviews: I conducted in-depth interviews with Ghoomar dancers representing diverse approaches: a traditional heritage practitioner focused on cultural preservation; a semi-professional dancer utilizing Instagram for professional development; and an established content creator with significant platform reach. Interviews (45-75 minutes, Hindi/English, via video call, audio recorded with consent) explored: motivations for Instagram presence; content creation processes including choreographic planning, filming, editing, and posting strategies; experiences navigating algorithmic visibility; negotiations between traditional elements and platform aesthetics; community interactions; and reflections on how Instagram shapes their relationship to Ghoomar practice. Interview transcripts were coded thematically, identifying patterns related to platform affordances, algorithmic strategies, aesthetic adaptations, and tension between traditional preservation and platform innovation. Netnographic observations and interview insights were analyzed iteratively.

Results

This research illuminates how Instagram's platform culture actively participates in shaping contemporary Ghoomar through dynamic interaction between algorithmic systems and content creator strategies. The findings reveal that platform affordances create conditions wherein certain traditional elements dramatic visual spectacle of spinning ghagras, vibrant costume aesthetics, emotionally expressive moments, individual virtuosity are strategically foregrounded and amplified, while other components communal circular formations, gradual rhythmic development, subtle percussive footwork, extended performance duration, regional folk songs in local dialects undergo adaptation, condensation, or backgrounding to align with platform visibility requirements. Content creators develop sophisticated approaches to maintaining cultural authenticity and traditional aesthetic values while innovating presentation formats that resonate with algorithmic recommendation patterns and platform audience expectations. This produces new hybrid forms of Ghoomar fusion choreographies, innovative camera work, creative location choices, and contemporary musical integrations representing both continuity and evolution from traditional practice. Instagram thus functions as an active site of cultural transformation, enabling broader access and diverse participation while reshaping what Ghoomar becomes in

digital contexts, creating dynamic tension between heritage preservation and platform-mediated innovation.

Future Work

Future research could examine a much larger body of Ghoomar videos using quantitative or computational methods. This would make it possible to see more clearly which kinds of visuals, movements, music choices, or editing styles are most likely to receive higher visibility and engagement on the platform. Such an approach would complement the present qualitative findings and provide a broader picture of how algorithmic preferences influence creative decisions. Second, it would be valuable to compare Instagram with other short-video platforms such as YouTube Shorts and TikTok. Since each platform has its own interface, audience culture, and recommendation system, Ghoomar may be shaped in different ways across platforms. A comparative study could help identify which changes are specific to Instagram and which are part of a wider shift in how traditional performance travels through digital media. Another important direction would be to conduct a longitudinal study by following the same creators over several years. This could show how their ideas about “what works” on the platform change over time, especially as Instagram modifies its algorithms and features. Such a perspective would help us understand platform influence not as a one-time shift, but as an ongoing and evolving process.

References

- boyd, d. (2010). Social network sites as networked publics. In Z. Papacharissi (Ed.), *A networked self: Identity, community, and culture on social network sites* (pp. 39–58). Routledge.
- Bucher, T. (2018). *If...then: Algorithmic power and politics*. Oxford University Press.
- Bucher, T., & Helmond, A. (2018). The affordances of social media platforms. In J. Burgess, A. Marwick, & T. Poell (Eds.), *The SAGE handbook of social media* (pp. 233–253). SAGE.
- Chakraborty, P. (2009). *Bells of change: Kathak dance, women and modernity in India*. Seagull Books.
- Couldry, N., & Hepp, A. (2017). *The mediated construction of reality*. Polity Press.
- Hjarvard, S. (2013). *The mediatization of culture and society*. Routledge.
- Kothari, S. (1968). *Folk dances of Rajasthan*. Sangeet Natak Akademi.
- Kozinets, R. V. (2020). *Netnography: The essential guide to qualitative social media research* (3rd ed.). SAGE.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Taylor, D. (2003). *The archive and the repertoire: Performing cultural memory in the Americas*. Duke University Press.

Van Dijck, J., Poell, T., & De Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press.

Vatsyayan, K. (1980). *Traditional Indian theatre: Multiple streams*. National Book Trust.

7C.2 Platform culture and Traditional Performance: Algorithmic Mediation of Ghoomar Dance on Instagram

Jessica Witte, Ari Stillman (University of Edinburgh)

Introduction

Social media platforms delineate acceptable use of platform content in their Terms of Service (ToS) – contractual agreements that govern user behavior, data access, and third-party activity. Since the Cambridge Analytica scandal in 2017 and the widespread deployment of generative AI systems after 2022, platform ToS have increasingly focused on regulating large-scale data collection by third parties, including private companies, non-profits, and academic researchers.

Application programming interfaces (APIs) have been widely used in academic research to collect social media data because they are ToS-compliant. Although APIs are often positioned as neutral pathways for data access, platforms initially designed their APIs as developer tools to facilitate commercial development rather than for academic research. When academic APIs – such as Twitter’s – were released, API-mediated platform data collection became the gold standard in many academic institutions and disciplines. As a result, ethical and methodological decisions about social media research and data management have been increasingly displaced from research communities onto ToS and other platform governance.

Using APIs for academic research has also meant that platforms retain control of both the data and the means of collecting it, leaving researchers at their whims to dictate what data can be accessed, how, and by whom. This has been evident throughout what has been called the “post-API” era (Freelon, 2018; Tromble, 2021), but became especially salient in 2024 when many major platforms including X (formerly Twitter), Reddit, and Facebook/Meta changed their ToS to restrict or monetize their APIs within a very short window of time. For many researchers, these changes felt like an abrupt closure of a trusted data access pathway approved by their research ethics boards. For the platforms, on the other hand, API restrictions embedded in ToS represented a mechanism for reasserting ownership over data in response to large-scale data extraction for training generative AI models.

This paper examines the ToS changes across X/Twitter, Facebook/Meta, and Reddit that have occurred since the Cambridge Analytica scandal to better understand how platform governance has shaped social media research practices. With evidence from a close reading of platforms’ ToS and research ethics frameworks published from 2017-2025, we argue that ToS and APIs function as boundary objects (Star & Griesemer, 1989), or infrastructural artifacts that coordinate action across distinct communities

but are interpreted differently by each. The multiplicity of meanings suggests that treating APIs and ToS as boundary objects can be a useful framework for unpacking complex questions around data ownership and researcher authority in the social media research space.

Objectives/Research Questions

This paper investigates how platform governance mechanisms have come to structure the ethical and methodological boundaries of social media research. Specifically, it asks: 1) How do APIs and ToS function as boundary objects in mediating relationships between platforms, researchers, and research governance institutions? 2) What powers or authorities are produced and reinforced by treating ToS compliance and API use as an ethical requirement? and 3) What strategies can we use to disentangle ethical research practice from contractual agreements?

In doing so, we address a major tension in social media research, which is that platform governance as enacted through ToS is motivated by monetization, corporate risk management, and control of data; academic inquiry, on the other hand is oriented toward the public good, concern for human subjects, and transparency. We foreground this misalignment through a reading of ToS and the API as boundary objects, thereby bringing to light the power asymmetries between researchers and platforms.

Methods

We conducted a qualitative rhetorical and discourse analysis of ToS agreements and associated API documentation from three major social media platforms: Twitter/X, Facebook/Meta, and Reddit. Our corpus covers documents published between 2017 and 2025, a period spanning the Cambridge Analytica scandal, the consolidation of platform governance regimes, and the emergence of what has been termed the post-API era of social media research.

Our analysis draws on boundary object theory (Star & Griesemer, 1989) and actor-network theory (Latour, 2005). Platform ToS and API documentation were collected from publicly available web archives to capture both incremental revisions and moments of abrupt policy change. To contextualize how platform policies circulate beyond platforms themselves, we supplemented this analysis with a close reading of research ethics frameworks, national-level guidance, and professional association documents addressing social media research and data collection during the same period. These materials were selected to reflect how ToS compliance has been operationalized within institutional ethics review, funding requirements, and publication standards.

Results

Our analysis reveals that across all three platforms, language in ToS documents increasingly foregrounds platform priorities over user welfare, data access, and otherwise humanistic principles. We also find important differences in how platforms operationalize these priorities. Reddit and Meta have, to varying degrees, reintroduced limited pathways for free or low-cost, ToS-compliant data access for non-commercial research. In contrast, X now requires all API users – including academic researchers – to pay an immodest monthly fee for access. This shift reframes data collection as a commercial transaction and introduces ethical tensions related to the redirection of public research funding to private platform infrastructures.

Across platforms, ToS provisions governing data collection and APIs have become increasingly narrow. For instance, X's policies now extend restrictions even to small-scale qualitative research, and all three platforms reserve authority to abruptly suspend access or even require data deletion. These provisions position platforms as both arbiters and enforcers of acceptable research practices. Rhetorically, these policies construct researchers as potentially adversarial actors whose behavior warrants surveillance; platforms are positioned as gatekeepers.

We find that APIs and ToS perform multiple, overlapping functions within the same socio-technical network. For platforms, these policies operate as governance and monetization tools; for research ethics boards, as standardized and “safe” proxies for ethical legitimacy given the dynamic and mercurial nature of digital research that reduce the need for case-by-case ethical judgment; for universities, as mechanisms for institutional risk mitigation; and for researchers, as methodological constraints that shape what questions can be asked and render research vulnerable to sudden platform policy change.

Finally, we find that the emergence of the post-API era has exposed the fragility of prior research practices that treated platform governance decisions and research ethics as aligned. Reading ToS and APIs as boundary objects makes visible how this alignment was contingent rather than stable, and how shifts in platform priorities can rapidly restructure the conditions of knowledge production. This framework helps explain not only what changed, but why existing research ethics infrastructures proved ill-equipped to respond.

Future Work

Additional empirical research is needed to document how researchers are adapting to the post-API era in practice. This includes examining methodological innovations, informal workarounds, and emergent data collection practices that do not utilize API pathways but remain ethically sound, such as data donation and data sharing, in addition to web scraping and browser automation. Comparative studies across

national and institutional contexts may reveal how different approaches to mediating the influence of platform ToS on research practice.

There is also a need to investigate the asymmetries between commercial and academic data access in greater depth. While academic researchers increasingly face restrictions framed as ethical or legal safeguards, large-scale data extraction for commercial purposes—particularly for machine learning and generative AI—continues under alternative contractual arrangements. Examining how these distinctions are justified and enforced would further clarify the political economy of platform data governance. Understanding how these distinctions are justified and enforced can further illuminate the political economy of platform data.

Finally, our research points toward the need for alternative research ethics frameworks that explicitly distinguish ethical data management practices from ToS compliance. Such frameworks should support researchers in making subjective decisions when designing their methodologies rather than proposing an objective, universal standard for social media ethics. This is not only because ethical research design changes depending on research question and dataset, but also because rapid shifts in platform governance and ToS make a blanket standard unrealistic. One way forward might be to more realistically calculate institutional risk for social media research given that as of the time of writing, no researchers or universities have been punished financially for breaking ToS. Rethinking ToS and other platform governance as boundary objects will support better aligning social media research practices with principles of public interest, transparency, and accountability in an increasingly platformized research environment.

References

- Freelon, D. (2018). Computational research in the post-API age. *Political Communication*, 35(4), 665-668.
- Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford university press.
- Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social studies of science*, 19(3), 387-420.
- Tromble, R., 2021. Where have all the data gone? A critical reflection on academic digital research in the post-API age. *Social Media+ Society*, 7(1), p.2056305121988929.

7C.3 The Platform Turn in Online Games

Aditya Deshbandhu (University of Exeter), Devina Sarwatay (City St George's, University of London), Aparna Vincent (Indian Institute of Management - Indore)

Introduction

Over the last decade, online games have undergone a significant platform turn. No longer confined to discrete self-contained experiences, games increasingly operate within wider economies of engagement characterising platforms. Live-service models, recurrent update cycles, and data-driven reward systems have shifted games from sites of episodic play to persistent worlds-of-being, where interaction, communication, and user-generated-content form the backbone of ongoing engagement.

This transformation mirrors evolution of platforms: games function simultaneously as communication infrastructures and social networks. Games like Hitwicket offer in-game messaging and inter-player coordination, while proximity chat in Suspects or photo-sharing in Assassin's Creed Shadows encourage visibility, virality, and status accrual through likes, views, and follower-like dynamics. These features embed games within platform-logics typically associated with Facebook, Instagram, or TikTok, an extension of earlier experiments like FarmVille, where games operated as social layers within broader platform ecosystems.

Live-service games intensify this shift by structuring participation through daily quests, seasonal rewards, and progression ladders sustaining habitual return and retention. Games like FreeFire (Craftland), Roblox, and Fortnite exemplify the emergence of the "supergame": expansive environments incorporating crossovers, branded partnerships, transmedia storytelling, and multimodal forms of play. These worlds increasingly act as supermarkets for other games, hosting minigames, creator-built maps, and modular experiences transforming players into prosumers.

Crucially, integration of player-generated content (asset creation, map design, etc.) has further aligned online games with platform cultures. By enabling players to build, customise, and circulate new spaces of interaction, games extend engagement through gameplay and creative labour. This blending of genres, narrative forms, and interactional modalities produces hybrid spaces where play, communication, sociality, and creation converge.

Against this backdrop, online games must be understood as socio-technical platforms embedded within wider digital ecosystems, competing with social media, streaming, and other attention-driven environments. Examining how these dynamics shape player behaviour, community participation, and world-building practices is essential to understanding contemporary game landscape and evolving platform mechanics.

This platformisation of games has significant implications beyond play itself. As games increasingly mirror and integrate communicative, algorithmic, and participatory logics of platforms, they become key sites for understanding contemporary platform power, cultural production, and attention economies. Examining games as platforms reveals how norms of visibility, labour, and community are being reconfigured within spaces not traditionally analysed in this fashion.

Objectives/Research Questions

This paper traces journeys of play, interaction, and progression across four current online games namely, Fortnite, Roblox, Hitwicket, and FreeFire to ask the following questions:

1. How do live-service online games operationalise platform logics similar to those of social media platforms, including infrastructures of communication, visibility, and engagement?
2. What incentive structures, rules, and reward systems are used to elicit and sustain repeated participation and transform players into active contributors within these game ecosystems?
3. How do player-generated content and in-game social features contribute to the emergence of “supergame” platforms that blend play, social networking, and creative labour?

Attempting to answer these questions enables an examination of online games as platforms, revealing infrastructures and mechanics that govern existing economies of engagement where games compete against social media, streaming, and other digital platforms for players' time.

Methods

In order to examine the four live-service online games as digital platforms we used a mix of digital walkthroughs (Light et al., 2018; Dugay and Gold-Apel, 2023) in conjunction with the self-reflexive Autoethnographic Ludonarratological Interpretative Framework (ALNI) (Deshbandhu, 2024) - an approach that is best suited to map, understand, and examine the “ship of Theseus” like nature of live-service games (Deshbandhu, 2025) where updates and fixes to games reconfigure gameplay experiences and player journeys significantly. To investigate the platformisation of games like Roblox, Fortnite, FreeFire, and Hitwicket it was imperative to map the shifts in gameplay experience, functionality, nature of engagements in-game, and the modifications to the game world – a process that is both careful and deliberate as it unpacks the shift in the interface and UI of games and the way the systems and mechanics are used by players.

The walkthrough method used for analysing digital apps and platforms from the intersecting perspectives of cultural studies and science and technology studies enabled us to examine live-service online games in parlance with conventional digital platforms. The ALNI framework enabled critical examination of personal experiences and journeys of play by drawing on autoethnography for meaning-making (Chang, 2016), stimulated recall (Pitkänen, 2015), close reading (Gray, 2010), and formal game analysis (Lankoski and Björk, 2015). We made observations in accordance with Calleja's six-point framework (2007) on play and player engagement. The combination of digital walkthroughs and the ALNI framework offers a holistic capturing of player interactions and insights into their decision making across the multiple game environments available. We unpacked how games as platforms offer affordances that shape player behaviour, world and community building practices, and community participation in myriad ways such that the contemporary game landscape evolves with/in platform mechanics. This analysis is informed by platform and media studies approaches to digital cultures, drawing on prior work on platformisation (Nieborg et al., 2022), participatory cultures (Jenkins and Ito, 2015), and user practices in social media environments.

Results

Preliminary analysis suggests four key patterns across the games studied:

1. Convergence of platform logics: Live-service games increasingly mirror social media architectures through nonplaying in-game interactions: likes, shares, visibility metrics, creator recognition/incentivization, and networked interaction.
2. Incentivised sociability: Communication and collaboration are not peripheral but native components of progression systems, shaping how community, competition, and co-presence are experienced.
3. Player labour and world-building: User-generated content functions as both creative expression and value-generating labour, sustaining platform ecosystems.
4. Attention retention beyond gameplay: Seasonal updates, transmedia events, and creator economies position games as competitors within wider attention/engagement markets alongside social media and streaming services.

Future Work

Future work will extend this analysis through comparative examination of additional game-platform hybrids and examine aspects of deeper engagement with/within player communities through interviews and discourse analysis. The project aims to contribute to broader theoretical debates on platformisation, participatory labour, labour in playful environments, and cultural production by positioning online games as central – not peripheral – sites of social media scholarship. Our findings from this study will

inform the development of a full-length journal article engaging platform studies and game studies audiences.

References

- Calleja, G. (2007). Digital game involvement (both micro and macro). *Games and Culture*, 2(3), 236–260.
- Chang, H. (2016). *Autoethnography as method*. London: Routledge.
- Deshbandhu, A. (2025). Situating everyday journeys of play: Rethinking engagement in a live service world. *Journal of Gaming & Virtual Worlds*, 17(2), 207–220.
- Deshbandhu, A. (2024). *The 21st Century in 100 Games*. London: Routledge.
- Duguay, S., & Gold-Apel, H. (2023). Stumbling blocks and alternative paths: Reconsidering the walkthrough method for analyzing apps. *Social Media+ Society*, 9(1), 20563051231158822.
- Gray, J. (2010). *Show sold separately: Promos, spoilers, and other media paratexts*. New York: NYU Press.
- Jenkins, H., & Ito, M. (2015). *Participatory culture in a networked era: A conversation on youth, learning, commerce, and politics*. NY: John Wiley & Sons.
- Light, B., Burgess, J., & Duguay, S. (2018). The walkthrough method: An approach to the study of apps. *New media & society*, 20(3), 881–900.
- Lankoski, P., & Björk, S. (2015). Formal analysis of gameplay. In P. Lankoski & S. Björk (Eds.), *Game research methods: An overview* (pp. 23–35). Pittsburgh, PA: ETC Press.
- Nieborg, D. B., Poell, T., & van Dijck, J. (2022). Platforms and platformization. *The SAGE handbook of the digital media economy*, 29–49.
- Pitkänen, J. (2015). Studying thoughts: Stimulated recall as a game research method. In P. Lankoski & S. Björk (Eds.), *Game research methods: An overview* (pp. 117–132). Halifax: ETC Press.

7C.4 Platformed Placemaking: Commemorative Stickers and the Production of Networked Places

Sarit Navon, Noam Gal (Open University), Ido Ramati (Hebrew University)

Introduction

The concept of placemaking originates in the field of urban planning and refers to processes through which public spaces are shaped in ways that enhance community life and address the varied needs and practices of community members (Project for Public Spaces, 2007). Over the past decades, placemaking has become increasingly prominent, reflected in a wide range of tools and practices, a broad discourse, and the emergence of a global movement and an academic discipline (Istoriou & Pozoukidou, 2024).

Building on this line of thought, communication scholars have also introduced the concept of digital placemaking to describe the use of digital media to create a sense of place and to foster emotional attachments to places (Halegoua & Polson, 2021). This perspective shifts attention from places as fixed physical entities to placemaking as an ongoing, mediated process, in which meanings, affects, and spatial relations are shaped through digital practices.

In this study, we examine bumper stickers used to commemorate the dead and the Instagram accounts to which they are linked through a QR code, a hashtag, or the account name displayed alongside the platform icon. This case study is characterized by two key features. First, it is hybrid, bringing together physical and digital spaces and practices. Second, it is multiple: rather than centering on a single, bounded place, these stickers circulate through a distribution mechanism that enables them to be affixed in numerous places. These places accumulate and assemble through Instagram, resulting in what we conceptualize as platformed placemaking.

Objectives/Research Questions

The study aims to examine how commemorative stickers operate across online and offline spaces, and how placemaking is shaped by social media platforms.

Specifically, it asks:

- What role does the digital space of an Instagram account play in relation to the physical commemorative sticker?
- What new digital practices emerge around commemorative stickers within processes of platformed placemaking?

- What are the key characteristics of platformed placemaking, and how are they related to the logics and affordances of social media platforms?

Methods

This study forms part of a larger research project examining commemorative stickers as an emerging and widespread practice in Israel following the October 7 attack and the subsequent war. The corpus of the broader project consists of 1,237 commemorative stickers, including stickers memorializing both Israeli and Palestinian victims. These stickers were photographed across different regions of Israel between July and November 2024. Sticker images were systematically coded and catalogued according to multiple parameters, including the location of the photograph, sticker shape, verbal content (such as slogans, biographical details, or quotations attributed to the deceased), visual elements, participation structures, and the presence of QR codes or web addresses.

The present analysis focuses on a subset of twenty cases in which the stickers explicitly refer to an Instagram account dedicated to the memory of the deceased. These cases were selected in order to trace the relations between offline practices and their extension, documentation, and circulation on social media. We employ digital ethnography to examine users' practices and the way they engage with the stickers and physical locations through a digital platform, in particular how stickers are placed, photographed, shared, and tagged, and how physical locations are made visible, networked, and meaningful through Instagram.

Following Varis (2015), we approach digital ethnography as a flexible and adaptive research orientation rather than a fixed set of techniques. Varis identifies two interrelated layers that digital ethnographies of communication should attend to: media affordances and online–offline dynamics. Within this framework, we focus analytically on the spatial dimension and on the ways in which online and offline practices are mutually constituted in what we conceptualize here as platformed placemaking.

Results

Findings indicate a strong relationship between the physical practice of affixing a sticker and a set of digital practices that accompany and complete this action. These include photographing the sticker or the action of placing it, uploading images or videos and sharing them on Instagram, tagging the relevant account, and engaging through comments. These practices are not incidental. Administrators of commemorative Instagram accounts explicitly guide and encourage followers to document the act of sticker placement, share it online, and tag the account, framing these actions as integral to the commemorative effort.

The offline practice of placing a sticker is considered partial – or even “lost” – if it remains undocumented and uncirculated. Digital documentation and circulation are understood as necessary for extending the presence of the sticker beyond its immediate location and moment of placement. However, the digital actions do not stand on their own. Photographs, tags, comments, and videos consistently refer back to a specific sticker placed in a concrete physical location. Digital practices derive their meaning from a material anchor and cannot, on their own, produce a place in the conventional sense of placemaking.

Rather than substituting physical presence, social media function as an infrastructure through which physical practices are extended, documented, coordinated, and rendered visible. Instagram affordances and logics regarding participation, circulation, and visibility reconfigure commemorative practices and placemaking: places become meaningful not only through their physical characteristics, but through their digital presence, circulation, and recognition within platformed environments.

Furthermore, findings indicate that platformed placemaking is not confined to a single physical site. Instead, it generates connectivity between multiple locations in which a particular commemorative sticker is affixed. These locations—ranging from well-known landmarks to mundane urban sites—may differ widely in character and geographic position, and may be scattered across cities, regions, or countries. Yet they become connected through shared tagging practices, affiliation with the same Instagram account, and participation in the same commemorative community. In this way, platformed placemaking results in networked places, embedding the connective and networked logic of social media into distant physical locations.

Future Work

Future research will address contexts beyond commemoration and extend the analysis of platformed placemaking to additional platforms, examining whether similar practices and dynamics emerge beyond Instagram. Comparative analysis across platforms may offer a more nuanced understanding of how media affordances shape practices and how online–offline spatial relations are produced.

References

- Halegoua, G., & Polson, E. (2021). Exploring ‘digital placemaking’. *Convergence*, 27(3), 573-578.
- Istoriou, T., & Pozoukidou, G. (2024). Understanding placemaking in the context of spatial planning: insights from a literature review. *Journal of Urbanism*: International Research on Placemaking and Urban Sustainability, 1-18.
- Project for Public Spaces. (2007). What Is Placemaking?
<https://www.pps.org/article/what-is-placemaking>

Varis, P. (2015). Digital ethnography. In The Routledge handbook of language and digital communication (pp. 55-68). Routledge.

Session 7D - Social Media Cultures and Everyday Life (6)

Day 2 · Tues 14 July · 11:20-12:50, Seminar Suite C, ARC

7D.1 Contemporary mediated civic becoming: Filipino American cultivation of civic agency through social media

Stephanie Thelwall (London School of Economics)

Introduction

During a period of heightened political polarisation (Finkel et al., 2020) and violent attacks against ethnic minority communities in the US (Savage & Sharp, 2025), who gets to be American, and what it means to be American, have become pressing questions. As such, understanding how ethnic minority communities come to see themselves as active members of American civil society, or in other words, as having civic agency, is an urgent matter during a time when collective activism is most needed. Contemporary sociality is distinguished by the preeminence of digital spaces as primary sites for gathering. This paper explores what brings participants to these spaces, what they think, feel, and do in these spaces, despite and in light of the architectures that shape and constrain their engagement. I will pay close attention to the manifestation of agency in which people cultivate their values, interests, and relationships, in ways that are consequential for developing a sense of a shared experience or shared stakes in affairs of common concern (Couldry et al., 2006). The supporting theory can be broken up into three parts: the socio-cultural nature of cultivating civic identity (Dahlgren, 2006); public spaces of political self-expression and democratic deliberation (Habermas, 1999; Mouffe, 2005; Papacharissi, 2016); and 'digital' affordances theory to analyse the mutually constitutive nature of the technical and social (Gibson, 2014). Bound by these three ontological dimensions, such spaces of public discourses are treated as vital sites for civic development where people are able to cultivate a sense of self in a way that allows them find common ground with others and from which could stem a civic sensibility. This project will explore the emotional, discursive, and material textures that give rise to the cultivation of a contemporary civic sensibility for a particular ethnic minority group, Filipino American youth.

Objectives/Research Questions

Specifically, this paper explores how the digital environment gives rise to particular civic practices, and what an understanding of such practices might reveal about the development of one's civic agency, a 'civic becoming'. This is operationalised into two questions: What role might a highly immersive and visual platform play in affording Filipino American youth the ability to engage in civic practices? And, how might such practices contribute to a civic becoming?

Methods

As such, it is necessary to adopt a methodology that will enable an examination of both individual experience and socially-embedded digital features. Taking this position is best supported by an epistemological position that stands between the positivism of technological determinism and more social constructivist thinking by rather embracing a materialist phenomenology, which asserts that '[if] the social world is built up, in part, of interpretations and communications, as phenomenology insists, our account of that world must look closely at the material infrastructures through which, and on the basis of which, communications today take place. Phenomenology cannot then only focus on how the world appears for interpretation by particular social actors. What is needed instead is a full-blown rethinking of the social construction of everyday reality, in all its interconnectedness, for the digital age' (Couldry & Hepp, 2016, p. 6). In light of this, a medley of digital ethnographic methods was employed, including diary keeping, semi-structured interviews, netnography (Kozinets, 2015; Serafinelli, 2017) and walkthrough method (Light et al., 2018). These methods were selected for their ability to capture both the materiality of digital spaces and the experiences and deeply held beliefs of participants. The data collection began with a 1-hour semi-structured interview in which I got to know each participant's interests, lifestyle and cultural background. At the same time, participants began the 8-weeks of diary keeping. The diary-keeping period took place in the lead-up to the 2024 US presidential elections. Upon completion of the diary keeping and the election, I visited each participant in-person to conduct the final interview. This was also a 1-hour semi-structured interview, but included the use of the walkthrough method to understand how participants moved through their digital spaces. It was also guided by interesting diary entries they shared with me.

Results

Over three months, these methods produced 30 hours of interview transcripts and 255 unique diary entries from the participation of 15 Filipino Americans in varying Filipino enclaves in the US. Reflexive thematic analysis is employed to unearth the ways in which the material and the social collide meaningfully. Through interviews, participant diary keeping, and platform walkthrough, this study finds that immersive social media are vital spaces that afford learning, exploration and generally, cultivation of passions, whether it be Filipino identity, creative practice, or relationships. Participants'

reflections and movements through their digital space revealed that participants practice their agency in three ways: they engaged in ways that reflected past experiences and conditioned norms, they responded innovatively and in novel ways to spontaneous conflict or when confronted with the unexpected, and they dreamed and aspired for different versions of their lives (see Emirbayer & Mische, 1998). The online space is a thriving space that incites emotion, informs, and connects people. In such a setting, participants embodied their personal values in the content they liked and the people they engaged with, they challenged their points of view when prompted to engage in meaning-making about topics they were unfamiliar with, and they aspired for a safer future for themselves and members of their community. Agentic practices distinguished by these different temporal orientations broaden an understanding of (digital) agency in a way that allows for an analysis of the micro-analytics of (socio-cultural) power (Carpentier, 2011). It is in these agentic practices that participants embody where they stand on a variety of issues of public concern and practice their power to such ends. They range from deliberately not partaking in public discourse to sharing medical resources for those in their community adversely affected by policy changes. The 'storytelling structures' (Papacharissi, 2015) powerfully afford participants to engage in a variety of agentic practices, which allow them to play with, exercise, and expand their understanding of issues of public concern.

Future Work

This project's contributions are three-fold: first, to a better understanding of contemporary citizenship in the US to enable an evolution from anachronistic mid-twentieth century conceptions of citizenship, but secondly, to contribute to a more sophisticated understanding of Filipino Americans relationship with the political, and finally, to bring attention to the mutually constitutive possibilities in engaging in and with the digital spaces we regularly occupy. During a period of heightened social division, political disillusionment and scrutiny of digital technologies in the US, there is more room to expand this exploration to shed light on how other ethnic minority groups find their way to civic sensibility.

References

- Carpentier, N. (2011). *Media and Participation: A site of ideological-democratic struggle*. Intellect.
- Couldry, N., & Hepp, A. (2016). *The mediated construction of reality*. Polity press.
- Couldry, N., Livingstone, S., & Markham, T. (2006). *Media Consumption and the Future of Public Connection*. The London School of Economics and Political Science.
- Dahlgren, P. (2006). Doing Citizenship: The Cultural Origins of Civic Agency in the Public Sphere. *European Journal of Cultural Studies*, 9(3), 267–286.

- Emirbayer, M., & Mische, A. (1998). What Is Agency? *American Journal of Sociology*, 103(4), 962–1023. <https://doi.org/10.1086/231294>
- Finkel, E. J., Bail, C. A., Cikara, M., Ditto, P. H., Iyengar, S., Klar, S., Mason, L., McGrath, M. C., Nyhan, B., Rand, D. G., Skitka, L. J., Tucker, J. A., Van Bavel, J. J., Wang, C. S., & Druckman, J. N. (2020). Political sectarianism in America. *Science*, 370(6516), 533–536. <https://doi.org/10.1126/science.abe1715>
- Gibson, J. J. (2014). The Theory of Affordances. In *The ecological approach to visual perception* (1st Edition, pp. 119–135). Psychology Press.
- Habermas, J. (1999). A Short Reply. *Ratio Juris*, 12(4), 445–453.
- Kozinets, R. V. (2015). Netnography. In A. Quan-Haase & L. Sloan (Eds), *The SAGE Handbook of Social Media Research Methods* (1st edn, pp. 1–8). Wiley. <https://doi.org/10.1002/9781118767771.wbiedcs067>
- Light, B., Burgess, J., & Duguay, S. (2018). The walkthrough method: An approach to the study of apps. *New Media & Society*, 20(3), 881–900. <https://doi.org/10.1177/1461444816675438>
- Mouffe, C. (2005). *On the Political*. Routledge.
- Papacharissi, Z. (2015). *Affective Publics: Sentiment, Technology, and Politics*. Oxford University Press.
- Papacharissi, Z. (2016). Affective publics and structures of storytelling: Sentiment, events and mediality. *Information, Communication & Society*, 19(3), 307–324. <https://doi.org/10.1080/1369118X.2015.1109697>
- Savage, D., & Sharp, S. (2025, September 8). Supreme Court upholds ‘roving patrols’ for immigration stops in Los Angeles. *Los Angeles Times*. <https://www.latimes.com/politics/story/2025-09-08/supreme-court-immigration-stops-los-angeles>
- Serafinelli, E. (2017). Analysis of Photo Sharing and Visual Social Relationships: Instagram as a case study. *Photographies*, 10(1), 91–111. <https://doi.org/10.1080/17540763.2016.1258657>

7D.2 From street to screen: The visibility labor of Italian organized crime on TikTok

Vincenzo Luise (Università degli studi di Napoli Federico II)

Introduction

In July 2025, the Italian Parliamentary Anti-Mafia Commission signed a memorandum of understanding with TikTok to curb the proliferation of mafia-related content on the platform (TikTok Italia, 2025). This initiative follows growing institutional concern about how social media shape youth imaginaries, portraying organized crime as aspirational. This article investigates how Camorra clans strategically use TikTok to perform visibility labour —symbolic practices aimed at gaining recognition and legitimacy in digital spaces. Specifically, it addresses two research questions:

How do Camorra clans perform visibility labour to construct criminal identities and gain symbolic legitimacy?

How do TikTok's algorithmic logics and communicative affordances shape these practices?

The study builds on two strands of literature. First, Di Nicola's (2022) spectrum theory of digital organized crime defines three categories: cyber-dependent (entirely digital crimes like hacking), cyber-enabled (digitally amplified traditional crimes), and cyber-assisted (offline crimes enhanced by digital tools). The Camorra's use of TikTok is best categorized as cyber-assisted crime, where the platform amplifies communication, coordination, and symbolic power. Second, the article draws on social media and visibility labour theories. Abidin (2016) and Duffy (2017) conceptualize visibility labour as the user-driven work of capturing and maintaining attention. Milan (2015) frames visibility as a contested political space, and Bucher (2012) introduces the regime of visibility to describe how algorithms and interfaces determine what becomes visible and to whom. Visibility labour, as analyzed in this study, unfolds across three interconnected dimensions. The performative dimension captures how users engage in aesthetic and symbolic self-representation to assert power and identity. The relational dimension focuses on the production of recognition and legitimacy within social networks. The infrastructural dimension examines how users tactically navigate and appropriate platform affordances, norms, and algorithmic pressures to maintain visibility. Camorra visibility practices are thus seen as socio-technical performances negotiated through platform logics, criminal culture, and individual agency.

Methods

This research adopts a three-phase digital ethnography (Dec 2024–Apr 2025) combining exploratory, simulative, and collaborative strategies. In the exploratory phase, videos tagged with #emanuelesibillo and #raffaelecutolo were collected using TikTok Scraper, yielding 100 videos and 72 accounts. However, hashtag-based sampling proved limited due to TikTok’s “content without context” dynamic (Bhandari & Bimo, 2022). The simulative phase involved creating a TikTok profile modeled after a hypothetical Camorra affiliate, following known Camorra-related accounts and content preferences (e.g., neomelodic music, luxury brands). After one week of algorithmic training, content was scraped from the “For You” feed, identifying seven verified accounts and 70 top-performing videos. In the collaborative phase, a key informant provided access to nine datasets related to specific clans, containing 7,343 videos. The 10 most viewed from each account were selected, yielding a final dataset of 1,220 TikTok videos. Analysis followed visual ethnographic content analysis (Altheide, 1987), focusing on symbolic, narrative, and interactional elements (e.g., emojis, music, filters, captions, gestures), with triangulation provided by the informant for decoding local cultural references.

Results

Performative visibility refers to how Camorra affiliates use TikTok to display power and authenticity through symbolic acts. Conspicuous consumption—luxury items, champagne, designer clothes—is staged in both elite venues and marginal settings, linking wealth to local credibility. Videos often feature tattoos, portraits of fallen bosses, or clips with mafia-themed soundtracks, turning the body and memory into sites of identity. Visibility also stages alliances and rivalries: celebratory clips promote unity, while women-led confrontational Lives enforce moral authority. Branding practices use clan names and symbols in products like perfumes or salons, marketed with nostalgic imagery and national certifications. Commemorative videos of figures like Cutolo or Sibillo create a digital pantheon that anchors criminal legitimacy in shared memory.

Relational visibility highlights how recognition and belonging are maintained through platform-mediated practices. Lives from prison allow inmates to assert presence and reconnect using coded language. Events like arrests, releases, or funerals are turned into public rituals of loyalty. Affiliates repost news articles mentioning them, reframed as badges of honour and set to neomelodic music. Gifting during Lives, especially in “Battle” mode, becomes a relational and economic practice, circulating money while reinforcing ties. Women appear as emotional anchors, shown caring for jailed relatives or reuniting after release. Music culture deepens affective resonance, with songs dedicated to fallen figures blending grief with pride.

Infrastructural visibility concerns how users tactically navigate TikTok’s algorithmic and economic systems. Rather than viral reach, content is tailored for local circulation using coded captions and emojis. Shared profiles, toggling privacy settings, and

symbolic usernames mark alliances and control. Live features are exploited for illicit sales, while videos promoting counterfeit goods use visual obfuscation to bypass moderation. Remixing mafia-related media builds a memetic public rooted in recognition and repetition. Visibility thus becomes a form of governance—those who master the code not only evade detection but reinforce criminal identity within platform logics.

Future Work

This article has examined how Camorra clans use TikTok as a tool for cyber-assisted organized crime, focusing on the role of visibility labour in shaping criminal identities and sustaining symbolic power. Through a three-phase digital ethnography, the analysis identified three key dimensions of visibility labour: performative (staging power and affiliation), relational (cultivating recognition, conflict, and solidarity), and infrastructural (adapting to platform logics and affordances). The findings show that visibility is not merely about being seen—it is a strategic, symbolic act shaped by algorithms, cultural codes, and platform-specific features. Camorra affiliates do not just use TikTok; they actively contribute to producing its regimes of attention and visibility. Visibility labour is diffuse, adaptive, and deeply embedded in platform ecologies—demanding responses that go beyond individual content removal.

References

- Abidin, C. (2016). Visibility labour: Engaging with influencers' fashion brands and #OOTD advertorial campaigns on Instagram. *Media International Australia*, 161(1), 86–100. <https://doi.org/10.1177/1329878X16665177>
- Altheide, D. L. (1987). Ethnographic content analysis. *Qualitative Sociology*, 10(1), 65–77. <https://doi.org/10.1007/BF00988269>
- Bhandari, A., & Bimo, S. (2022). Why's everyone on TikTok now? The algorithmized self and the future of self-making on social media. *Social Media + Society*, 8(1), 1–11. <https://doi.org/10.1177/20563051221086241>
- Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, 14(7), 1164–1180. <https://doi.org/10.1177/1461444812440159>
- Di Nicola, A. (2022). Towards digital organized crime and digital sociology of organized crime. *Trends in Organized Crime*, 1–20. <https://doi.org/10.1007/s12117-022-09457-y>
- Duffy, B. E. (2017). *Not getting paid to do what you love: Gender, social media, and aspirational work*. Yale University Press.

Il Sole 24 Ore. (2023, October 18). Gratteri: “Le mafie utilizzano i social per arruolare manovalanza”. <https://stream24.ilsole24ore.com/video/italia/gratteri-mafie-utilizzano-social-arruolare-manovalanza/Aft6xk1C>

Milan, S. (2015a). Mobilizing in times of social media: From a politics of identity to a politics of visibility. In L. Dencik & O. Leistert (Eds.), *Critical perspectives on social media and protest: Between control and emancipation* (pp. 53–71). Rowman & Littlefield.

TikTok Italia. (2025, July 2). Accordo tra Commissione Antimafia e TikTok per la legalità online. TikTok Newsroom. <https://newsroom.tiktok.com/it-IT/accordo-commissione-antimafia-e-tiktok-italia-per-legalita-online>

Zulli, D., & Zulli, D. J. (2022). Extending the Internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform. *New Media & Society*, 24(8), 1872–1890. <https://doi.org/10.1177/1461444820983603>

7D.3 New Cultural and Religious Practices: The Use of Facebook by Azorean Religious Communities

Pedro Lucas (Universidade Nova de Lisboa)

Introduction

Social media has become part of the toolkit used by local communities to represent, promote, memorialize, and negotiate their most symbolic practices. As in many other locations, on the Catholic island of Faial (Azores Archipelago) Facebook has been progressively integrated into such practices, namely the island's main religious celebrations: the Festivals of the Holy Spirit and the Patronal Saints' festivals. Each festival is organized locally. The Festivals of the Holy Spirit are organized by small groups known as *Irmandades* ("brotherhoods/sisterhoods") and, although they take place in the church and in the streets of each town, they generally have a more reserved character, often being attended by invitation or through a participation fee. These celebrations mobilize a particularly rich symbolic repertoire and are widely regarded as a central element of Azorean identity more broadly. The patronal saint festivals, in turn, are organized by each town's parish and constitute its main annual celebration, encompassing the participation of local institutions, associations, and residents. Their central symbolic element is the statue of the patron saint, and their ritual climax is the patronal procession.

Since the early 2010s, Facebook use has become widespread among the island's inhabitants (and is likely the most commonly used social media platform, although no official data is available). Facebook posts focusing on these celebrations can be found from that period onwards, but it is only from the middle of the decade that a more systematic and diversified use of the platform emerges, both at the individual and institutional levels. This includes single-image posts confirming the publisher's attendance at a given festival; systematic photographic (and, more recently, videographic) reportages produced by hobbyist or semi-professional photographers; private groups used to coordinate and update members on the activities of *Irmandades* or parishes; public pages of institutions promoting upcoming celebrations or commemorating past festivals and emblematic figures; and frequent calls for material or logistical support addressed to the local community. These posts also elicit particularly strong engagement from members of the island's diaspora, for whom Facebook becomes a privileged channel to maintain symbolic ties with their place of origin. This research delves into the symbolic religious and cultural dimensions of these practices, examining how Facebook-mediated interactions become embedded in local forms of belonging and collective meaning.

Objectives/Research Questions

How have the communities of Faial integrated Facebook into their religious and social practices surrounding local religious celebrations? How do Facebook posts about these celebrations elicit patterned linguistic responses that function as expressions of cultural and religious belonging?

Methods

The study is based on a digital and field ethnography conducted between April 2024 and October 2025 in the towns of Feteira and Ribeirinha (Faial Island, Azores). The ethnographic work involved the identification and combined qualitative and quantitative analysis of approximately 1,300 Facebook posts related to local religious celebrations, 25 in-person semi-structured interviews, and participant observation during 10 religious events.

Qualitative analysis consisted of thematic coding of post images, post texts, and post comments, with particular attention to symbolic content and recurrent linguistic patterns. Interview transcripts were also thematically coded to contextualize online practices within broader religious and cultural dynamics.

Quantitative analysis focused on descriptive statistics of post-level metrics (reactions, comments, and shares), disaggregated by image typologies and post author characteristics. In addition, exploratory statistical correlations were conducted between categories of visual content and categories of comments in order to identify systematic patterns in user responses.

Results

Although data analysis is still ongoing, several preliminary results can be outlined. First, Facebook appears to enable bottom-up processes of symbolic production, particularly through the emergence of pro-am visual reporters - often local amateur photographers - who systematically document religious celebrations and circulate their images within local and diasporic networks.

Second, preliminary analyses indicate a strong correlation between the presence of explicitly religious symbols or ritual moments in post images and the occurrence of religious formulae in the associated comments. Posts depicting central ritual elements, such as patron saint statues, processions, or symbols of the Holy Spirit, are significantly more likely to elicit formulaic responses (e.g., blessings or invocations) than posts depicting peripheral or backstage aspects of the celebrations.

Third, a marked gender asymmetry emerges in participation patterns. Male users predominate among the producers of posts focusing on religious celebrations, particularly in roles related to visual documentation, whereas female users are more

prevalent among those commenting on these posts, especially through formulaic religious expressions and culturally inflected responses.

Finally, posts related to religious celebrations also systematically elicit cultural-identitarian responses that express local belonging, collective memory, and attachment to place, particularly among diaspora users.

Future Work

Future stages of this research will aim at deepening the interpretation of the results obtained through the digital ethnography. Namely, new group and individual interviews with both content producers and active commenters will be carried out in order to better contextualize the statistical correlations identified between formulaic expressions and religious visual content, as well as between the presence of religious symbols and expressions of cultural belonging. These interviews will focus on how participants understand and make sense of their own online practices and of the symbolic value attributed to Facebook posts related to local religious celebrations.

Additionally, future work will consider the use of other social media platforms by members of these communities, allowing for a comparative approach to the mediation of similar religious and cultural practices across different digital environments.

References

- Cohen, A. P. (1985). *The Symbolic Construction of Community*.
- Geertz, C. (1973). *The Interpretation of Cultures: Selected Essays*.
- Goffman, E. (1959). *The Presentation of Self in Every Day Life*.
- Goffman, E. (1967). *Interaction ritual: Essays in face-to-face behavior*.
- Hepp, A. (2020). *Deep Mediatization*.
- Hjarvard, S. (2014). Mediatization and cultural and social change: An institutional perspective. In K. Lundby (Ed.), *Mediatization of Communication* (pp. 199–226).
- Meyer, B. (2013). Mediation and Immediacy: Sensational Forms, Semiotic Ideologies, and the Question of the Medium. In *A Companion to the Anthropology of Religion* (J. Boddy and M. Lambek).
- Meyer, B. (2015). Picturing the Invisible. *Method & Theory in the Study of Religion*, 27(4–5), 333–360.

7D.4 "I Don't Know Many Words, But I'm Smart": Platform Literacies and Older Content Creators on Douyin

Yingwen Wang, Zoetanya Sujon (University of the Arts, London)

Introduction

Short video platforms such as Douyin have drawn millions of older adults into content creation, including many with limited formal education and low textual literacy. This presents a striking contradiction: users who struggle to read and write are successfully navigating digital platforms and producing creative content. China's demographic context makes this particularly significant, the country has one of the world's largest aging populations (projected to reach 26% by 2050), and while overall adult literacy rates have risen dramatically (from 65% in 1982 to 97.5% in 2020). For instance, older adults aged 65+ maintain lower rates in Tibet, with only 85% able to read and write simple statements as of 2018 (Trading Economics, 2025).

Scholarship on digital literacies has evolved considerably, encompassing media literacy (Livingstone, 2004), data literacy (Pangrazio & Selwyn, 2023), algorithmic literacy (Oeldorf-Hirsch & Neubaum, 2023), and social media literacy (Cho et al., 2022). Recently, platform literacy has emerged as an umbrella concept. Ha and Kim (2024) propose platform literacy as a "vital competence" involving knowledge of content curation, algorithmic visibility, and strategic tool use, progressing through stages from basic access to advanced evaluation. Livingstone et al. (2025) position platform literacy as critical knowledge of "platform architecture and business models," emphasizing attunement to platform affordances and risks. However, these frameworks predominantly privilege text-centric, analytical competencies, inadequately accounting for the haptic, symbolic, and multimodal pathways through which users with low textual literacy may develop meaningful platform engagement. While Douyin's visually rich interface accommodates users who have limited literacy, its monetization features remain text-centric, revealing subtler exclusions even within platforms celebrated for lowering access barriers. We address this gap by examining how older, low-literacy creators develop non-textual platform literacies and how platform infrastructures shape their participation opportunities.

Objectives/Research Questions

This study investigates how older Chinese adults with low textual literacy engage with Douyin as content creators. We extend existing conceptualizations of "platform literacy" by foregrounding non-textual epistemic modes, haptic, symbolic, and multimodal competencies, that enable meaningful digital participation beyond text-centric frameworks. Our research addresses three interconnected questions:

(1) How do older content creators with limited literacy navigate Douyin and create content?

(2) Why is this engagement meaningful to them?

(3) What does this tell us about changing digital literacies and digital inequalities?

We argue that older content creators exercise creative agency through rich non-textual skills, demonstrating a distinct form of platform literacy that enables what we term “relational sociality”—connection, recognition, and community building mediated through platform vernaculars. Simultaneously, we examine how platform infrastructures create barriers when creators attempt to access monetization features, producing what we characterize as “marketized sociality” where social participation becomes enrolled in the platform’s attention- and data-driven model. Through this analysis, we aim to shift analytical focus from user skill deficits to the structural constraints imposed by platform architectures, contributing to both platform literacy theory and digital inequality scholarship.

Methods

We employed qualitative methods to explore interactions between low-literate older content creators and the Douyin platform, combining semi-structured interviews with participant observation in rural and semi-rural Chinese contexts. We recruited 12 participants (8 female, 4 male) aged 60-80 through snowball sampling. All participants regularly post short videos on Douyin and self-identified as “low literate,” reporting difficulty with reading or writing tasks despite recognizing some Chinese characters. Participants came from provinces including Jiangsu, Shanxi, Sichuan, and Shanghai, with education levels ranging from primary to middle school. Content genres included daily life documentation, lip-syncing, dancing, local drama performance, and experience-sharing.

Semi-structured interviews (60-90 minutes each) were conducted via WeChat or Douyin voice call using plain language approaches. We avoided academic terminology, instead asking questions like “How did you learn to use this app?” and probing specifically about literacy-related experiences, including navigation of textual elements and help-seeking strategies. We conducted two weeks of participant observation with two especially active creators, involving home visits and monitoring of their Douyin accounts to triangulate interview data with observed practices.

Data analysis followed Braun and Clarke’s (2006) thematic analysis approach through phases of familiarization, initial coding, theme development, and refinement. We attended particularly to instances of creative platform use and workarounds, forms of communication (emoji, voice, video), expressions of self-efficacy, instances of frustration or exclusion, and references to monetization. Codes were clustered into

themes including “Platform Navigation and Symbolic Learning,” “Visual Communication,” and “Infrastructural Barriers.” Research was conducted with university ethics approval and attention to best practices for working with older populations.

Results

Our findings reveal three interconnected dimensions of platform literacy among older Douyin creators. First, participants developed sophisticated haptic and symbolic navigation strategies through trial-and-error learning scaffolded by family and peers. As P6 (female, 65) described: “I can’t read the words on the screen, but I remember the places. Muscle memory, like playing the piano by ear!” This touch-based navigation was inherently social; P9 (male, 62) noted: “My grandson, who is educated, helped me download Douyin and taught me how to use it. After he taught me, I practised on my own and ended up memorising it.” Sustained practice built self-efficacy, with P6 insisting: “I’m not stupid. I don’t know many words, but I’m smart.”

Second, participants cultivated multimodal creative vernaculars enabling meaningful connection. Emoji sequences functioned as sophisticated semiotic systems, P6 explained that “🙏😊❤️” meant “thank you, I’m happy and I send love.” Voice messaging transformed Douyin into a voice-based social network; P4 (female, 67) shared: “We talk like on the phone, but through Douyin, because I cannot type.” P7 (female, 60) offered a reframing of communicative hierarchies: “Using emojis and my voice, I can say everything I want. Maybe I don’t write, but I speak from the heart.”

Third, we identified tensions between relational and marketized sociality. While participants achieved meaningful connections through creative expression, monetization features imposed insurmountable textual demands. P6 explained: “I can’t read, so how could I join in? If you want to promote goods, you need to at least know about the product.” P9 articulated vulnerability: “If I sign something, it could be selling my house away and I wouldn’t know!” Algorithmic distribution further confined visibility; P2 (male, 64) noted: “Douyin requires money to boost your content; for us older ones, maybe only friends and similar age people see it.” Thus, inequalities stem less from user deficits than from platform architectures privileging text-centric competencies.

Future Work

This study offers two primary theoretical contributions. First, we extend platform literacy concepts by identifying non-textual epistemic modes rooted in affective practice and social scaffolding, challenging frameworks that prioritize analytical knowledge. Second, we advance digital inequality research by shifting focus from user skill deficits to structural constraints imposed by platform infrastructures. Future research should investigate whether similar patterns emerge across different platforms

and cultural contexts, and examine how platform design changes affect low-literacy users' participation opportunities.

For policy and practice, our findings suggest digital inclusion initiatives focused solely on upskilling may prove inadequate without addressing exclusionary platform design. Future work should develop design interventions making monetization features accessible to users with diverse literacy profiles, and assess how regulatory frameworks might address platform-produced inequalities.

References

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
<https://doi.org/10.1191/1478088706qp063oa>
- Cho, H., Cannon, J., Lopez, R., & Li, W. (2022). Social media literacy: A conceptual framework. *New Media & Society*, 26(2), 941-960.
<https://doi.org/10.1177/14614448211068530>
- Ha, S., & Kim, S. (2024). Developing a conceptual framework for digital platform literacy. *Telecommunications Policy*, 48(3), 102682.
<https://doi.org/10.1016/j.telpol.2023.102682>
- Livingstone, S., Jessen, R. S., Stoilova, M., Stănicke, L. I., Graham, R., Staksrud, E., & Jensen, T. (2025). Can platform literacy protect vulnerable young people against the risky affordances of social media platforms? *Information, Communication & Society*, 1–18. <https://doi.org/10.1080/1369118X.2025.2518254>
- Livingstone, Sonia (2004) What is media literacy? *Intermedia*, 32 (3). pp. 18-20
- Oeldorf-Hirsch, A., & Neubaum, G. (2023). What do we know about algorithmic literacy? The status quo and a research agenda for a growing field. *New Media & Society*, 27(2), 681-701. <https://doi.org/10.1177/14614448231182662>
- Pangrazio, L., Selwyn, N. (2023). *Critical Data Literacies: Rethinking Data and Everyday Life*. MIT Press

Session 7E - Online and Offline Communities

Day 2 · Tues 14 July · 11:20-12:50, Studio 2, ARC

7E.1 Information Practices in Fragmented Online Communities: Taylor Swift Fans and Anti-Fans on Reddit

*Samantha Vilkins, Katherine M. FitzGerald, Sebastian Svegaard, Axel Bruns
(Queensland University of Technology)*

Introduction

Information practices—the ways users select, share, cite, and interpret sources—are central to community formation and identity. These practices reveal how communities differentiate themselves not only through shared interests but through shared, social epistemologies: collective understandings of what counts as legitimate knowledge and authoritative sources (Fricker et al, 2020). Fan communities offer particularly concentrated, affective, and visible contexts for studying these dynamics. Operating within "feels culture" (Stein, 2015), fan communities form around intense emotional investment, highlighting the highly social and affective facets of negotiating information practices and identity formation.

The importance of studying information practices becomes particularly salient in contexts of growing epistemic polarization, where fragmented communities develop distinct standards for evaluating information. Fan studies scholarship establishes that both fandom and anti-fandom (Sandvoss, 2019) drive community formation (Gray, 2003; Barnes, 2022), with anti-fan identities existing both inside and outside traditional fan spaces. In contexts of epistemic polarization, some communities reject traditional authoritative sources in favor of collective sensemaking to establish and defend their own alternative frameworks. Scholarship on populist expertise (Marwick & Partin, 2022) helps us understand these dynamics, particularly patterns where interpretation of specific texts or events becomes central practice, mirroring patterns observed in communities centred on conspiratorial theorising.

Taylor Swift fandom during September 2023 to September 2024 provides a uniquely compelling case study. This period saw Swift's Eras Tour generate unprecedented cultural attention while events from her voter registration initiative to her presidential endorsement, alongside public focus on her relationship with NFL player Travis Kelce, sparked controversy and even conspiratorial discourse (Pengally, 2024). These developments prompted a rise in activity, as well as fan and anti-fan fragmentation.

Reddit provides a valuable platform for studying these dynamics, particularly as research access to other major social media platforms has been increasingly restricted. During the study period, Swift-related subreddits experienced record activity and community splintering, with users drawing upon diverse sources—from social media posts to tabloids, mainstream news, and collectively-produced fan documents—making the platform rich for examining how communities navigate competing truth claims and establish hierarchies of legitimate knowledge.

Objectives/Research Questions

This study investigates information practices within Taylor Swift fandom communities on Reddit during a period of significant cultural discourse and community fragmentation. Our primary research question is: How do different Taylor Swift fandom communities on Reddit differ in their information practices, particularly regarding the sources they cite, share, and reference?

As part of this, we investigate what types of information sources (e.g., social media, news outlets, tabloids, fan documents, lyrics) are prevalent across different subreddit communities and different practices. Further, to what extent can we characterize communities based on their information practices and source usage—do patterns emerge at the subreddit level, for subreddit types, or around distinct user groups within and across subreddits?

This study contributes to growing scholarship on participatory information practices and epistemic communities online, while advancing methodological approaches for studying information practices on Reddit through URL and domain-level analysis and practice mapping.

Methods

Our approach combines qualitative ethnographic observation with quantitative analysis of information sharing practices across eleven Taylor Swift-related subreddits over one year (September 1, 2023 to September 30, 2024). The subreddit selection includes diverse communities: eight dedicated to Swift—including r/TaylorSwift (3.2 million members), relationship-focused r/travisandtaylor (116k members), queer-reading focused r/GaylorSwift (and r/Gaylor_Swift), and two explicitly anti-fandom subreddits—and three general subreddits filtered for Swift content (including r/popheads with 2.4 million members).

The mixed-methods design integrates four complementary approaches, each addressing different levels of community practice.

Contextual grounding:

We conducted ethnographic observations to understand community norms and practices. We systematically analyzed subreddit rules and moderation practices,

highlighting those specifically related to information sources and discussion topics. As part of the larger project, we constructed a timeline of Swift's major events during this period and corresponding changes in activity across subreddits, revealing reactive interplay as communities fractured with different perspectives on both Swift herself, and on previous moderation and content focus of existing communities.

Domain-level data collection: We extracted data on all domains shared in posts and within comments across the eleven subreddits, capturing both overall patterns and variations across subreddits, threads and users. This includes URLs shared in link posts, text posts, and comments, analyzed at the domain level to identify source types (e.g., twitter.com, variety.com, specific fan blogs). We manually categorized domains by type, based on their platform, tone/style and content—e.g. as entertainment news outlets, tabloids, fan-created/hosted resources, etc.

Practice mapping analysis:

Information sharing patterns were examined alongside our broader practice mapping framework (Bruns et al., 2024), which clusters users by similarity of actions rather than direct interactions. This approach captures interplays between platform affordances and community norms, recognizing both subreddit-level and thread-level governance as well as individual user behavior as established in the literature (e.g. Robards, 2018; Rajadesingan, Resnick & Budak, 2020; a, Zhang & Althoff, 2022).

Comparative URL framing analysis: Building on the domain-level patterns identified through quantitative analysis, we conducted preliminary targeted qualitative investigation of how identical URLs are framed and interpreted across different subreddits. We manually identified instances where the same URL appeared in multiple communities and analyzed the surrounding discourse—including post titles, post text, and comment threads—to understand how communities mobilized the same sources for different purposes and from different perspectives.

Results

Our analysis reveals distinct information practice profiles across Swift fandom communities on Reddit, characterized by both the types of sources used and patterns of their presentation. Notably, users appear to congregate around particular source types even more strongly than around particular subreddits, suggesting connection to topic or type of discussion more robustly connective than community identity in this case.

Analysing practices at the individual user level, clusters form around distinct source types. There are clusters of US news sources separate to UK ones, as well as separation between trade magazines and tabloids. There are separate clusters for YouTube, for TikTok, and for fan-created content hosted on Google Docs and Tumblr. These do not appear to align strongly with subreddit communities. However, subreddits

can be somewhat meaningfully characterised by their relative use of certain domains compared to others. The most striking—though expected—examples of these include the large, general pop music subreddit r/popheads being predominantly characterised around mainstream sources such as music industry magazines and links to music, and the queer analysis focused r/GaylorSwift and r/Gaylor_Swift connected to fan-created content on TikTok and Tumblr.

Further from this, our analysis supports indication of hierarchies of valued or privileged sources within different communities. Mainstream fan communities tend to prioritize official sources, including announcements from Swift and her team, and established entertainment media. Anti-fan communities show preference for tabloid sources and social media commentary that challenge official narratives, or at least offer more affective and opinionative commentary. Relationship-focused communities prioritize paparazzi images and direct social media sources. Communities oriented toward interpretive or conspiratorial readings demonstrate distinctive patterns of collective sensemaking, treating primary texts (lyrics, social media posts) as requiring decoding rather than accepting surface meanings.

Preliminary qualitative comparisons illustrate how different communities share the same sources with markedly different interpretations and framing, at both the domain and URL levels. This epistemic polarization—where communities cite conflicting sources about the same events or interpret identical sources through opposing frameworks—appears central to community differentiation, though these patterns operate at a more subtle level than initially expected.

Future Work

Understanding information practices within fragmented communities has implications for comprehending how online communities establish collective knowledge, resist or reinforce epistemic authorities, and navigate disagreement in participatory digital spaces.

This study opens several avenues for future research. While domain-level analysis provides valuable insights into source types and sharing patterns, further content analysis of how sources are used—whether cited as authoritative evidence, critiqued as biased, or reinterpreted through community-specific perspectives—would deepen understanding of information practices. Temporal analysis could reveal how these practices evolve in response to specific events or controversies.

The patterns observed raise questions about whether these clusters of information practices and hierarchies of source legitimisation in Taylor Swift fandom represent broader trends in online community formation. Applying this methodological approach to other online communities could illuminate this further, along with cross-platform analysis examining how practices shift as individuals navigate platform affordances

and the corresponding communities that grow within and across different spaces online.

References

Barnes, Renee. *Fandom and Polarization in Online Political Discussion: From Pop Culture to Politics*. Cham: Palgrave Macmillan, 2022.

Bruns, A., Kasianenko, K., Suresh, V. P., Dehghan, E., & Vodden, L. (2024). Untangling the Furball: A Practice Mapping Approach to the Analysis of Multimodal Interactions in Social Networks (arXiv:2407.05956). arXiv.
<https://doi.org/10.48550/arXiv.2407.05956>

Fricke, M., Graham, P. J., Henderson, D., Pedersen, N. J., & Wyatt, J. (Eds.). (2020). *The Routledge handbook of social epistemology*. New York: Routledge.

Gray, Jonathan. "New Audiences, New Textualities: Anti-Fans and Non-Fans." *International Journal of Cultural Studies* 6, no. 1 (March 1, 2003): 64–81.
<https://doi.org/10.1177/1367877903006001004>

Marwick, A. E., & Partin, W. C. (2022). Constructing alternative facts: Populist expertise and the QAnon conspiracy. *New Media & Society*, 26(5), 2535-2555.
<https://doi.org/10.1177/14614448221090201>

Pengally, M. (2024). Nearly one in five Americans believe Taylor Swift-Biden conspiracy, poll finds. *The Guardian*.
<https://www.theguardian.com/music/2024/feb/14/taylor-swift-biden-psyop-conspiracy>

Rajadesingan, A., Resnick, P., & Budak, C. (2020, May). Quick, community-specific learning: How distinctive toxicity norms are maintained in political subreddits. In *Proceedings of the international AAAI conference on web and social media* (Vol. 14, pp. 557-568).

Robards, B. (2018). Belonging and Neo-Tribalism on Social Media Site Reddit. In: Hardy, A., Bennett, A., Robards, B. (eds) *Neo-Tribes*. Palgrave Macmillan, Cham.
https://doi.org/10.1007/978-3-319-68207-5_12

Sandvoss, C. "The Politics of Against: Political Participation, Anti- Fandom, and Populism." In *Anti- Fandom: Dislike and Hate in the Digital Age*, 125–46. New York University Press, 2019. <https://doi.org/10.18574/nyu/9781479866625.003.0009>.

Stein, L. E. (2015) *Millennial Fandom: Television Audiences in the Transmedia Age*. University of Iowa Press.

Weld, G., Zhang, A. X., & Althoff, T. (2022). What Makes Online Communities ‘Better’? Measuring Values, Consensus, and Conflict across Thousands of Subreddits. Proceedings of the International AAAI Conference on Web and Social Media, 16(1), 1121-1132. <https://doi.org/10.1609/icwsm.v16i1.19363>

7E.2 From Mate Selection to Testosterone Talk: Mapping Science Claims in the (Neo-) Manosphere

Catherine Baker (Dublin City University)

Introduction

Over the past decade, the “manosphere” (Ging, 2017), sometimes now termed the “neo-manosphere” (Gerrand et al., 2025), has emerged as a diffuse yet influential digital ecosystem encompassing male self-improvement influencers, dating coaches, and anti-feminist and misogynist online communities. While these communities vary in focus, platforms, and communication style, they share a preoccupation with heterosexual relationships, traditional masculinity, anti-feminism, and male aggrivement. Existing scholarship has extensively documented the misogynistic and anti-feminist discourses prevalent in these spaces. The literature has also frequently noted the reliance on scientific explanations to legitimise the endorsement of misogyny, anti-feminism, and traditional gender roles. However, to date, no mapping has been done of the specific types of scientific claims that are circulated across these spaces. In particular, no cross-platform research has examined how scientific authority is constructed, circulated, and adapted within different manosphere and manosphere-adjacent environments.

This paper addresses that gap by mapping the manosphere’s use of “bro-science”. In this paper, “bro-science” is used to describe a mode of science talk that draws selectively on real research domains—such as evolutionary psychology, neuroscience, and personality psychology—within content that focuses on heterosexual relationships, masculinity, and neoliberal self-optimisation. These explanations are rarely presented as speculative or contested. Instead, they are framed as settled biological facts that naturalize gender essentialism, traditional masculinity, and neoliberal capitalist logics. This “bro-science” talk is often mediated through high-visibility podcasts and YouTube channels associated with figures such as Andrew Huberman and Jordan Peterson. Across platforms, simplified and meme-ified versions of this bro-science discourse circulate through “manfluencer” short-form video clips and traditional manosphere communities.

Rather than treating manosphere science claims as simple misinformation, this paper situates them within broader histories of popularised science and neoliberal self-help. It argues that manosphere bro-science amplifies epistemic tendencies already present in certain scientific subdisciplines and their public-facing forms, particularly biological determinism, binary sex models, and individual-level explanations for complex social phenomena. It concludes that understanding this mediated process is crucial for

explaining how reactionary gender ideologies acquire legitimacy and popular appeal in contemporary digital cultures.

Objectives/Research Questions

Digital platforms increasingly mediate public understandings of science. Within manosphere and neo-mansphere spaces, appeals to scientific authority play a particularly central role in framing gender relations, sexual difference, and masculine selfhood as natural, inevitable, and biologically grounded. Examining how these claims are produced, circulated, and taken up is therefore essential for understanding contemporary gender politics online. The primary objective of this study is to analyse how scientific language and authority are mobilized across manosphere and neo-mansphere digital spaces to legitimise gender norms, sexual hierarchies, and neoliberal masculinist ideologies.

The study is guided by five research questions:

1. How does the manosphere use scientific language and authority to legitimise gender norms, sexual hierarchies, and neoliberal logics?
2. Which scientific domains are most frequently drawn upon in manosphere and male self-improvement spaces, and which concepts and authorities circulate most prominently?
3. How are academic ideas translated, simplified, and repurposed as they move from high-authority scientific or pop-scientific sources to mid-tier influencers and vernacular forms such as memes and forum discourse?
4. What ideological work do these scientific claims perform, particularly in relation to gender essentialism, heterosexual dating hierarchies, and self-optimization narratives?
5. Why do specific subdisciplines resonate so strongly with reactionary online gender content?

Methods

This study employs a multi-sited digital ethnography designed to trace the circulation of scientific discourse across within the manosphere and neo-mansphere. A multi-sited approach was adopted to capture the movement of ideas, actors, and discursive forms across interconnected digital environments. Data were collected in a range of modalities, including long-form YouTube videos, short-form TikTok and YouTube Shorts content, Reddit subcommunities, and independent manosphere forums such as incels.is. These sites were selected to reflect both high-visibility influencer-led spaces and lower-profile, community-driven environments where scientific claims are further interpreted, contested, and meme-ified.

The analysis captures both high-authority scientific discourse produced by credentialed academics or public intellectuals and its downstream reinterpretations by influencers and everyday users. High-authority content included podcasts, lectures, interviews, and long-form explanatory videos that explicitly referenced academic credentials, peer-reviewed research, or institutional affiliations. Data collection involved systematic sampling of videos, posts, and comment threads, alongside the archiving of recurring scientific claims and images such as charts, graphs, and infographics. Sampling prioritised content that explicitly invoked scientific language, cited research, or made claims about biological, psychological, or evolutionary explanations.

The analysis employed a combination of qualitative analytical methods. Thematic coding was used to map the scientific domains, concepts, and actors commonly included in this content. This stage of analysis identified recurring references to specific subdisciplines (e.g. evolutionary psychology, neuroscience), key concepts (e.g. hormones, mate value, sexual dimorphism), and authoritative figures or institutions. In-depth discourse analysis was then applied to a subset of data to explore the rhetorical strategies utilised in the content, the ideological function of scientific claims, and the ways these ideas were circulated and adapted between actors and digital spaces. This close analysis focused on how scientific authority was constructed, and how claims were used to bolster broader narratives of masculinity, heterosexual hierarchy, and self-optimisation.

Results

The analysis identifies a recurring typology of science claims that dominate manosphere and neo-manosphere discourse, mapping the flow and patterns of scientific content across these digital ecosystems.

Four scientific domains play particularly central roles. First, evolutionary psychology functions as a grand explanatory framework for gender and heterosexual relationships. Concepts such as mating strategies, sexual marketplaces, and hypergamy are invoked to frame gender inequality as biologically inevitable. These ideas are presented as universal truths rather than contested theories, enabling users to reinterpret dating frustrations as outcomes of immutable female nature rather than as products of social or economic power structures.

Second, pop neuroscience provides a language of neoliberal optimisation and self-mastery. Simplified accounts of dopamine, reward circuits, and “brain wiring” frame masculinity as a neurological achievement that can be engineered through discipline, abstention, and productivity. Gender differences are routinely essentialized as brain-based, while aligning masculinity with neoliberal ideals of self-control and individual responsibility.

Third, endocrinology, particularly testosterone discourse, underpins crisis narratives about masculine decline. Claims about falling testosterone levels, feminising foods, and hormonal threats to male vitality circulate widely. Hormones function as moral signifiers: high testosterone is equated with strength and authority, while low testosterone is associated with weakness and social failure.

Finally, personality psychology supplies typologies that justify dominance hierarchies. Concepts such as the Dark Triad are mobilized to frame aggression, emotional detachment, and manipulation as evolutionarily or psychologically rewarded strategies in heterosexual dating markets.

The paper argues that manosphere “science” is not a novel distortion but a digitally amplified continuation of long-standing biases in pop psychology and pop science. In these ecosystems, specific subsets of academic and pop-academic work are privileged because they offer biologically grounded, individualised explanations that foreground gender difference while obscuring social and structural factors. These selectively mobilised ideas are translated into “commonsense” claims that support gender hierarchies, crisis narratives, and neoliberal capitalist structures.

Future Work

As this paper represents the first comprehensive mapping of science claims across the manosphere. While this approach establishes an initial typology of scientific claims and their circulation, further work is needed to extend these findings. One key avenue for future research is to incorporate comparative platform analysis to examine how algorithmic affordances and technological affordances shape the visibility, and virality of scientific claims. Such work would clarify how platform-specific logics influence the translation and uptake of “bro-science” across different digital environments. Second, future research will deepen engagement with critical science studies to situate manosphere bro-science within longer histories of science, particularly continuities with late twentieth-century pop psychology and evolutionary narratives. Third, future work will examine audience reception through interviews or surveys to explore how scientific claims are taken up, contested and challenged by content consumers. Together, these extensions aim to clarify how scientific authority operates as a cultural resource in reactionary gender politics and to inform broader debates about science communication, digital influence, and masculinity.

References

- Ging, D. (2017). Alphas, Betas, and Incels: Theorizing the Masculinities of the Manosphere. *Men and Masculinities*, 22(4), 638-657.
<https://doi.org/10.1177/1097184X17706401> (Original work published 2019)

Gerrand, V., Ging, D., Roose, J. M., & Flood, M. (2025). Mapping the Neo-Manosphere(s): New Directions for Research. *Men and Masculinities*, 28(5), 443-464. <https://doi.org/10.1177/1097184X251350277>

7E.3 "Cuckoo Nut Lady C*nt Face McGee": Gendered Constructions of Mothers Who Kill in True Crime Content

Avery Fraser, Ella Duncan, Kathleen Rodgers (University of the Fraser Valley)

Introduction

The public's persistent fascination with true crime content in Western popular culture is reflected in an extensive and ever-expanding body of true crime podcasts, documentaries, television series, online forums, and books, which has fostered rich academic discourse (Vedric & Little, 2022). However, despite the growing popularity of the true crime entertainment genre and, equally important, its capacity to significantly shape the public's perceptions of perpetrator justification, victimhood, and blame, true crime media has received relatively limited scholarly attention.

As one of the most popular and diverse entertainment genres, true crime content has presented a new and novel way to examine how perpetrators are framed in the media, particularly female perpetrators. That is, unlike male perpetrators, female perpetrators directly challenge deeply entrenched gender norms that associate femininity with other traits such as passivity, weakness, and nonviolence (Dana, 2025). To be more precise, women and girls who are accused of committing violent acts are perceived as "doubly deviant and doubly damned" due to the "compound nature" of their actions, violating both legal norms and gendered expectations at the same time (Lloyd, 1995, as cited in Dana, 2025, p. 562).

Put simply, when women who have committed domestic violence are discussed in true crime media, they are made distinct from other women because they have survived, thereby complicating prevailing narratives surrounding gender, power, and violence. One notable example includes the case of Francine Hughes, who, in 1977, after thirteen years of domestic abuse, killed her ex-husband by setting fire to the bed where he was sleeping. During her trial, Hughes was found not guilty by reason of temporary insanity and has since been lauded as a "feminist icon," with the survival of her abuse often overshadowing the death of her ex-husband. This paper examines the portrayal of female perpetrators accused of domestic homicide in contemporary true crime documentaries and podcasts.

Drawing on framing theory, this paper examines how popular true crime podcasts and Netflix crime documentaries portray female perpetrators who engage in fatal domestic violence and, in turn, how these portrayals intersect with broader Western cultural stereotypes surrounding femininity, power, and beyond. In particular, this research analyzes how feminist ideas such as female empowerment and equality clash with

depictions of women who commit fatal domestic violence, yet are at times celebrated as feminist icons.

Objectives/Research Questions

This research addresses the following research questions:

R1: How are women who commit domestic homicide framed within contemporary true crime, and in what ways do these framings differ from those applied to male perpetrators? In what ways are they constructed as different from other women?

R2: Considering true crime is a feminized media space, shaped largely by women as both audiences and producers, to what extent do these portrayals challenge or reproduce the sensationalist tropes historically associated with women who kill in mainstream media?

Methods

Building on the work of feminist scholars who investigate media representations of gender-based violence (Easteal et al, 2015; Noetzel et al, 2023, Cucklanz, 2020), this study builds on a broader study investigating true crime content. The study consists of two data sources. We conducted both quantitative and qualitative analyses of true crime content. First, we investigated the most-consumed true crime podcasts between 2014 and 2024 where we collected data from the 30 first episodes of each podcast and the most recent 30 episodes (N=600). In addition, we conducted a quantitative content analysis of all of Netflix' true crime productions between 2015 and 2025 (N=206). In keeping with our theoretical framework, we developed a coding scheme and built a database designed to capture the representation of victims, perpetrators and crimes. Our aim was first, to assess general information about the individuals included in the most listened-to podcasts and Netflix productions. This includes information for each episode such as the racialized group, gender, sexuality and nationality of perpetrators and victims. We also identified the categories of criminal behaviour that were covered in each case and whether the crime involved intimate-partner violence. In our qualitative analysis we investigated the dominant framing and discourses surrounding crimes, perpetrators and victims.

Results

This work is exploratory and our results are preliminary. Our initial results indicate that the feminized spaces of true crime, particularly in podcasting, challenge traditional media portrayals of women who kill, presenting a more sympathetic narrative rooted in discussions of the impact of domestic violence. Netflix productions are more in line with the traditional media portrayals, drawing on stereotypes about criminally accused women and girls as “inherently and inhumanely evil or mentally unstable” (Dana, 2025,

p. 561). We reflect on how platform logics shape the presentation of narratives about female criminals. Many of the podcast hosts in our dataset are women who write, produce, and narrate their own work. As a result, discussions of female-perpetrated domestic homicide are explored with greater depth than in previous studies that focus on traditional media institutions—such as newspapers—which are more strongly shaped by patriarchal structures.

Future Work

This research is ongoing. Our work was originally about true crime media constructions of violence against women. This work has been based on the investigation of true crime media. Our work encourages us to engage in qualitative interviewing of female podcasters in order to better understand their construction of discourses about violence against women. We expect to begin this work following the completion of this current project.

References

- Cuklanz, L. (2019). Representations of Gendered Violence in Mainstream Media. *Questions de Communication*, 35(1), 307–321.
<https://doi.org/10.4000/questionsdecommunication.19487>
- Dana, C. (2025). Femininity on trial: decoding female violence in true crime documentary. *Feminist Media Studies*, 25(3), 560–575.
<https://doi.org/10.1080/14680777.2023.2296859>
- Easteal, P., Bartels, L., Nelson, N., & Holland, K. (2015). How are women who kill portrayed in newspaper media? Connections with social values and the legal system. *Women's Studies International Forum*, 51(1), 31–41.
<https://doi.org/10.1016/j.wsif.2015.04.003>
- Kilgariff, K & Hardstark, G (2025, January 30) You're Kidding Yourself, Supt (465) [Audio podcast episode] In My Favourite Murder
- Noetzel, S., Mussalem Gentile, M. F., Lowery, G., Zemanova, S., Lecheler, S., & Peter, C. (2022). Social campaigns to social change? Sexual violence framing in U.S. news before and after #metoo. *Journalism*, 24(6), 1232–1262.
<https://doi.org/10.1177/14648849211056386>
- Vedric, N., & Little, J. (2023). True victims: men's privilege, class, and violence against women in *Making a Murderer* and *The Jinx*. *Feminist Media Studies*, 23(3), 975–990. <https://doi.org/10.1080/14680777.2022.2027495>

S7E.4 ‘Tapping into Your Feminine Energy’ Down the Alt-Right Pipeline: Comparing Narratives and Community Bonds in the Comment Sections of Tradwife and Divine Feminine Content Creators on TikTok

Monica Vania Chavez, Anna Feigenbaum, Rinlapas Ketverapong (University of Glasgow)

Introduction

The digital “alt-right pipeline” has predominantly analysed male radicalization as men make up a larger portion of the right-wing (Nadeem, 2025; Lawson, 2023; Scott & Day, 2025). However, women are crucial actors within this movement, upholding heteronormative, conservative ideologies that serve as the backbone of the movement (Scott & Day, 2025; Sykes & Hopner, 2023). This paper examines how this pipeline manifests for women through two distinct but ideologically adjacent content genres on TikTok: The “tradwife” (traditional wife) and the “divine feminine. The tradwife aesthetic glamorizes a nostalgic 1950s inspired domesticity that is perfectly curated (Banet-Weiser & Reinis, 2025). In contrast, divine feminine content creators repackage rigid gender essentialism within language of spiritual self-help and personal empowerment, coaching women to “step into” passive, receiving feminine energy to attract a male provider and live a soft life. (Ranasinghe, 2025; McGowan, 2025; Berge, 2023; Joseph, 2021). Both genres have been identified as potential on-ramps to more extreme right-wing ideology, yet while scholarly literature has analysed tradwife content and its narratives, the divine feminine content remains underexamined despite its massive reach in March of 2023, #divinefeminine had over two billion views and continued popularity (Berge, 2023; Sykes and Hopner, 2023; Joseph, 2021; Ranasinghe, 2025). This research aims to fill that gap and contribute to a more nuanced understanding of how the alt-right pipeline manifests for women online through a critical feminist lens within the 3Ps (Platforms, People, Power) digital research framework, as it investigates TikTok as a platform that algorithmically promotes these lifestyles, the power of creators to legitimize anti-feminist narratives, and the people within comment sections who actively co-construct these digital subcultures (Feigenbaum & Tønnesen, forthcoming, 2026).

Objectives/Research Questions

This research aims to conduct a comparative analysis of two content genres implicated in the female facing alt-right pipeline on TikTok. It seeks to understand the distinct aesthetic and narrative strategies employed by creators and to analyze the subsequent community dynamics that form in their comment sections. The study is guided by two primary research questions:

Q1: How do the core narratives and aesthetic strategies of divine feminine content creators differ from those of tradwife creators on TikTok?

Q2: What community bonds are occurring in the comment sections of divine feminine content creators compared to tradwife creators on TikTok?

These questions will be explored through several analytical dimensions:

- Identifying shared and divergent desires or grievances such as critiques of feminism or desire for a male provider.
- Analysing differences in linguistic tone and language.
- Interpreting community endorsed values via the most like comments.
- Examining indicators of social bonding and in-group identity formation.

4. Methods

This research will employ a multi-phase qualitative comparative content analysis to investigate the distinctive narrative frameworks and community dynamics of tradwife and divine feminine content on TikTok.

Phase 1 addresses Q1 through a multimodal qualitative content analysis of creator produced videos (Fazeli et al., 2023; Feigenbaum & Tønnesen, forthcoming 2026). To ensure a robust and focused analysis, a purposive sample of 2-3 prominent content creators from each genre will be identified, the scope criteria, draws from established methodological frameworks, require that creators: (1) have a primary focus on promoting tradwife or divine feminine lifestyles, (2) maintain a substantial follower base (>50,000), and (3) produce recent and consistent content (Feigenbaum & Tønnesen, forthcoming 2026; Sykes & Hopner, 2024). For each creator, their 5 most recent videos will be analyzed (20-30 videos total). A hybrid inductive-deductive coding strategy will be used (Molder et al., 2021; Fereday & Muir-Cochrane, 2006; Feigenbaum & Tønnesen, forthcoming 2026). Deductive codes will be derived from existing literature on tradwife content (the promotion of rigid gender roles, biblical submission, and a glamorized “return to tradition” often tied to far-right conspiracy theories and a distrust of government and modern medicine (Liporace, 2025; Meyerowitz, 2024; Sykes & Hopner, 2024; Sykes & Hopner, 2023; King’s College London, 2025; Bower, 2025). Inductive codes will emerge from the data, particularly for divine feminine content, focusing on spiritual self-help framing, the dichotomy of masculine/feminine “energy,” and promises of strategic personal gain (Ranasinghe, 2025; McGowan, 2025; Berge, 2023). The unit of analysis is the holistic video, coding for visual aesthetics, spoken narrative, and on-screen text (Meyerowitz, 2024; Rascoe, 2024; Sykes & Hopner, 2024; Fazeli, et al., 2023).

Phase 2 addresses Q2 through a qualitative content analysis of viewer comments. Approximately 100 comments will be scraped from one focal video for each genre (200

comments total) using the tool exportcomments.com (Feigenbaum & Tønnesen, forthcoming 2026). Comments will be analyzed for thematic content (e.g., expressions of gratitude, shared grievances, requests for advice, personal testimonials), linguistic markers of community (use of “we,” in-group slang), interaction patterns (debates, mutual support, and reinforcement of creators messaging), and multimodal elements (e.g. images, GIFs, emojis) (Dolan, 2025; Sykes & Hopner, 2024; Feigenbaum & Tønnesen, forthcoming 2026). The analysis will specifically examine how comment sections function as spaces for digital interaction rituals that foster in-group identity, using shared resentment or affirmation as bonding mechanisms (Kallio & Törnberg, 2025; Liporace, 2025; Scott & Day, 2025; Banet-Weiser & Reinis, 2025; Dolan, 2025).

Results

The research has not yet been conducted; findings will be available upon study completion.

Future Work

This study is designed as an in-depth comparative case study. Future research should expand the scale of analysis to include a larger sample of videos and comments to track narrative evolution and community formation over time. An extension would be to analyze cross-platform migration, examining how creators on TikTok may direct followers to more explicitly ideological content on platforms like Instagram or subscription based services. Furthermore, applying sentiment analysis could quantitatively map the emotional tone and structure of these communities, complementing the qualitative findings. Finally, conducting interviews with consumers of this content would provide invaluable insight into individual pathways through these digital ecosystems and the real-world impact of these ideologies.

References

- Banet-Weiser, S., & Reinis, S. (2025). The rage of tradwives: Affective economies and romanticizing retreat. *Feminist Theory*.
<https://doi.org/10.1177/14647001251371971>.
- Beaufils, C., Gohir, S., & O'Connor, H. (2025). Tradwife: Between myths and realities. King's Global Institute for Women's Leadership.
<https://www.kcl.ac.uk/giwl/assets/giwl-tradwife-report.pdf>
- Berge, V. L. (2023). The dark and divine feminine: Secular stereotypes and theological tropes on TikTok. *Tijdschrift Voor Genderstudies*, 26(3/4), 293–312.
<https://doi.org/10.5117/TVGN2023.3/4.006.BERG>.
- Bower, L. J. (2025). The thorn in feminism's side: Black feminist reconceptualization and defence of #tradwives and the #tradwife movement. *Journal of Gender Studies*, 34(7), 1037–1053. DOI: <https://doi.org/10.1080/09589236.2024.2423198>.

- Dolan, E. W. (2025, January 21). Online incel forums generate “dark emotional energy” that reinforces toxic group identity. PsyPost. <https://www.psypost.org/online-incel-forums-generate-dark-emotional-energy-that-reinforces-toxic-group-identity/>.
- Fazeli, S., Moghaddasi, H., Asadi, F., & Bouraghi, H. (2023). Performing qualitative content analysis of video data in social sciences and medicine: The visual-verbal video analysis method. *International Journal of Qualitative Methods*, 22. <https://doi.org/10.1177/16094069231185452>.
- Feigenbaum, A., & Tønnesen, O.D. (forthcoming, 2026) The social media methods workbook. Routledge.
- Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, 5(1), 80–92. <https://doi.org/10.1177/160940690600500107>.
- Joseph, C. (2021, June 15). “Marry rich”: Love coaches are telling women to channel feminine energy. Refinery29. <https://www.refinery29.com/en-gb/2021/06/10503519/what-is-a-love-coach>.
- Kallio, P., & Törnberg, A. (2025). Rituals of resentment: How dark emotional energy fuels incel identity and solidarity. *Acta Sociologica*, 68(3), 287–304. <https://doi.org/10.1177/00016993251317963>.
- King’s College London. (2025, March 25). The tradwife phenomenon isn’t a return to tradition, it’s a plea for balance. <https://www.kcl.ac.uk/news/the-tradwife-phenomenon-isnt-a-return-to-tradition-its-a-plea-for-balance#:~:text=The%20phenomenon%20and%20its%20popularity%20have%20risen,pressures%2C%20not%20nostalgia%20for%20a%20bygone%20era>.
- Lawson, R. (2023). Being a ‘real man’ on /r/The_Donald: Masculinity, ethnic identity, and the alternative right. In *Language and mediated masculinities: Cultures, contexts, constraints* (pp. 85–108). <https://doi.org/10.1093/oso/9780190081041.003.0005>.
- Liporace, E. (2025, September 24). Tradwives: A superficial threat to feminism but deeper divide among women. The Emory Wheel. <https://www.emorywheel.com/article/2024/09/tradwives-a-superficial-threat-to-feminism-but-deeper-divide-among-women>.
- McGowan, E. (2025, February 10). What is the divine feminine? The Good Trade. <https://www.thegoodtrade.com/features/divine-feminine/>

- Meyerowitz, A. (2024, May 13). How the ‘trad wife’ movement turned these stay-at-home mums into millionaires. Glamour UK. <https://www.glamourmagazine.co.uk/article/trad-wife-movement-millionaires>
- Molder, A. L., Lecheler, S., & Boykoff, M. T. (2021). Framing the global youth climate movement: A qualitative content analysis of Greta Thunberg’s moral, hopeful, and motivational framing on Instagram. *The International Journal of Press/Politics*, 27(3), 668–695. <https://doi.org/10.1177/19401612211055691>
- Nadeem, R. (2025, January 15). Party affiliation fact sheet (NPORS). Pew Research Center. <https://www.pewresearch.org/politics/fact-sheet/party-affiliation-fact-sheet-npors/>.
- Ranasinghe, S. (2025, January 31). ‘Divine feminine’ advice on TikTok is just part of the alt-right pipeline. Teen Vogue. <https://www.teenvogue.com/story/divine-feminine-advice-on-tiktok>.
- Rascoe, A. (2024, January 28). ‘Trad wives’ are trending. What does that say about feminism today? NPR. <https://www.npr.org/2024/01/28/1227453741/trad-wives-are-trending-what-does-that-say-about-feminism-today>.
- Scott, K., & Day, L. (2025). TikTok tradwives: Femininity, reproduction, and social media. *Gender and Education*. <https://doi.org/10.1080/09540253.2025.2546050>.
- Sykes, S., & Hopner, V. (2023, July 7). Tradwives: The housewives commodifying right-wing ideology. Global Network on Extremism and Technology. <https://gnet-research.org/2023/07/07/tradwives-the-housewives-commodifying-right-wing-ideology/>.
- Sykes, S., & Hopner, V. (2024). Tradwives: Right-wing social media influencers. *Journal of Contemporary Ethnography*, 53(4), 453–487. <https://doi.org/10.1177/08912416241246273>

Session 7F - Knowledge, Value, and Power in Platform Economies

Day 2 · Tues 14 July · 11:20-12:50, 225, ARC

7F.1 The Social Life of the MSI Metric: Rationalities and Metrologies around “Meaningful Social Interactions” at Meta

Matías Valderrama Barragán (Adolfo Ibáñez University)

Introduction

In January 2018, Mark Zuckerberg announced what appeared to be a fundamental transformation of Facebook’s mission: a shift from optimising relevance and engagement toward fostering “meaningful social interactions.” This change was presented as an ethical response to growing concerns about user well-being, misinformation, and political polarisation. At the centre of this transformation was a newly introduced metric, Meaningful Social Interaction (MSI), designed to reorient Feed ranking algorithms toward interactions between friends and family rather than public or professional content. While this metric change was widely discussed in public debates, policy circles, and the media (Cameron et al., 2022; Hagey & Horwitz, 2021; Metz, 2021; Narayanan, 2023; Wong, 2018), far less is known about what such a shift actually entailed within the platform’s internal research and engineering practices from a sociological perspective.

This article examines the “social life” of the MSI metric, its emergence, operationalisation, consequences, and partial abandonment, through an analysis of Meta’s internal documents leaked by whistleblower Frances Haugen in 2021, commonly referred to as the Facebook Files or Facebook Papers. Drawing on scholarship in platform studies, science and technology studies (STS), and the sociology of quantification, the article approaches platforms as socio-technical stages in which sociality is continuously engineered, monitored, and intervened in. Within this framework, metrics are not treated as neutral measurement devices but as active participants in shaping what becomes visible, valuable, and actionable on platforms.

By analysing how MSI was defined, weighted, tested, and ultimately problematised inside Meta, the article contributes to research on platform algorithms and governance (Bucher, 2018; Gillespie, 2018; Gorwa, 2024) and engages emerging work on the

rationalities of tech workers (Dorschel, 2022; Nachtwey & Seidl, 2023; Selling & Strimling, 2023) by showing how these rationalities are articulated and stabilised through metric design and evaluation. Rather than seeking to reveal the precise technical functioning of algorithms, the study treats leaked documents as ethnographic artefacts that record practices, conflicts, and styles of thought within the company. This approach allows for a critical examination of how societal harms are translated into technical problems of measurement, optimisation, and governance. This analysis speaks directly to debates on how social media platforms shape political communication and social life more broadly, and the distribution of responsibility for these harms through internal systems of measurement and optimisation.

Objectives/Research Questions

The central objective of this article is to examine how “meaningful social interaction” was defined, measured, and operationalised within Meta, and how the MSI metric shaped both platform governance and the organisation of sociality on Facebook. The article asks how MSI came to occupy a central role in News Feed ranking, how “meaningfulness” was translated into measurable signals and weighted interactions, and what assumptions about social interaction, value, and harm were embedded in its design. It further asks how these assumptions generated unintended consequences, particularly in relation to political communication, virality, and the amplification of polarising content. More broadly, the article seeks to understand what the trajectory of the MSI metric tells about the metrological regimes and rationalities that underpin contemporary platform design and governance. Rather than reconstructing the exact metric’s technical functioning or assessing its current status, the life of the MSI metric is analysed as a relevant case for examining how platform actors translate societal concerns into technical objectives, rendering sociality, value, and harm measurable and governable through performance metrics that function as central elements of what I will define as the stagecraft of platforms.

Methods

This article draws on materials produced during a broader research study that mapped the controversies sparked by the Facebook Files/Papers, which combined digital methods with conventional qualitative methods (document analysis, participant observation and interviews with journalists, activists, experts, and former Meta employees). The article is based on quali-quantitative document analysis of the internal Meta materials leaked by Haugen and circulated through news media outlets. The primary corpus consists of more than 1,200 internal documents, including research reports, slide decks, dashboards, training materials, and posts from Meta’s internal communication platform, Workplace. These materials, which have informed recent academic publications (Huang & Krafft, 2024; Lewis & Christin, 2024), cover the period from 2015 to 2021, with a concentration in 2019-2021, when internal debates around

MSI and other platform harms intensified.

Methodologically, the study approaches internal Meta documents not as transparent representations of algorithmic systems, but as situated artefacts that record organisational practices, modes of reasoning, and processes of knowledge production. The analysis is inspired by ethnographic approaches to documents, which treat documents as active actors in organisational life rather than passive carriers of information (Asdal & Reinertsen, 2022; Shankar et al., 2017), as well as by ethnographies of algorithms that examine the sociocultural dimensions of algorithm development and use (Bucher, 2018; Christin, 2020; Seaver, 2019). From this perspective, the aim is not to reveal hidden mechanisms, but to analyse the logics, imaginaries, and effects through which metrics and algorithmic systems are conceived, justified, and governed. Accordingly, the analysis focuses on how Meta employees define problems, frame solutions, justify interventions, and communicate findings to internal audiences. Particular attention is paid to how metrics are described and operationalised, how trade-offs are articulated, and how success or failure is evaluated through dashboards, metrics, and internal evaluative frameworks.

Results

The analysis shows that the MSI metric emerged from the convergence of interrelated logics within Facebook in 2017–2018. First, MSI responded to public criticism of the platform's effects on users' mental health, including concerns about "mindless passive scrolling." Framed as a moral intervention aligned with "Time Well Spent" discourse, this rationale corresponds to what Christin et al. (2024) term a user wellbeing logic, through which platforms legitimise algorithmic change by foregrounding care for users. Second, a reputational logic followed political scandals, positioning MSI as a response to concerns about misinformation, polarisation, and election integrity by reorienting Feed ranking toward "meaningful" interactions. Third, and less visible publicly, an engagement or production logic addressed internal concerns about declining Original Broadcast Posts (OBPs) and engagement since 2017, positioning MSI as a mechanism to stimulate content creation, with internal documents showing that countering engagement decline was central to its development.

MSI was defined as a weighted aggregation of interaction types multiplied by factors such as relationship strength and interaction "significance." Meaningfulness was operationalised through behavioural equivalences derived from surveys and regression analyses, valuing interactions for their predicted capacity to generate future engagement. This utilitarian conception reduced social interaction to measurable actions abstracted from context and intention. As machine-learning models replaced rule-based calculations, MSI became increasingly opaque, prioritising content likely to trigger engagement. But MSI optimisation produced unintended consequences.

Internal research documented that MSI produced a “harmful virality”, incentivising political parties, media outlets, and other public actors to produce more provocative and polarising content to trigger reshares and comments. Despite awareness of these effects, MSI remained a central goal metric, illustrating how performance metrics become entrenched in platform governance and how harms are reframed as measurable side effects rather than infrastructural outcomes, diffusing responsibility and shifting blame.

Taken together, the social life of MSI suggests that platforms design a form of theatrical stagecraft in which algorithms, metrics, and optimisation targets selectively foreground certain actors and interactions while obscuring others. Moreover, Meta’s metrological practices configure a made-to-measure sociality in which users, brands, media outlets, and political actors are constituted as social entities, and interaction is treated as a quantifiable substance that circulates and accumulates. The analysis also identifies a distinctive rationality that I call behavioural interactionism, through which social life is understood via observable behavioural signals and governed through utilitarian adjustments of weights, incentives, and ranking criteria.

Future Work

Future research could extend this analysis in several directions. First, comparative studies could examine similar metrics and algorithms across platforms to assess whether behavioural interactionism and utilitarian valuation are shared rationalities in the big tech industry. Second, further work could explore how such metrics interact with regulatory frameworks and external accountability mechanisms, particularly as platforms face increasing legal and political scrutiny. Finally, ethnographic research with current or former platform workers could deepen understanding of how internal dissent, ethical reflection, and professional values shape -or fail to shape- the design of platform stagecraft over time.

References

- Asdal, K., & Reinertsen, H. (2022). Doing document analysis: A practice-oriented method. SAGE Publications Ltd ; SAGE Publications Inc.
- Bucher, T. (2018). If...Then: Algorithmic Power and Politics.
- Cameron, D., Wodinsky, S., & DeGuerin, M. (2022, mayo 2). Here’s What the Facebook Papers Say About the Ranking Algorithms That Control Your News Feed. Gizmodo. <https://gizmodo.com/facebook-news-feed-ranking-algorithms-how-they-work-1848814459>
- Christin, A. (2020). The ethnographer and the algorithm: Beyond the black box. *Theory and Society*, 49(5–6), 897–918. <https://doi.org/10.1007/s11186-020-09411-3>

- Christin, A., Bernstein, M. S., Hancock, J. T., Jia, C., Mado, M. N., Tsai, J. L., & Xu, C. (2024). Internal Fractures: The Competing Logics of Social Media Platforms. *Social Media + Society*, 10(3), 20563051241274668. <https://doi.org/10.1177/20563051241274668>
- Dorschel, R. (2022). A new middle-class fraction with a distinct subjectivity: Tech workers and the transformation of the entrepreneurial self. *The Sociological Review*, 70(6), 1302–1320. <https://doi.org/10.1177/00258172221103015>
- Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
- Gorwa, R. (2024). *The Politics of Platform Regulation: How Governments Shape Online Content Moderation*. Oxford University Press.
- Hagey, K., & Horwitz, J. (2021, septiembre 15). Facebook Tried to Make Its Platform a Healthier Place. It Got Angrier Instead. *Wall Street Journal*. <https://www.wsj.com/articles/facebook-algorithm-change-zuckerberg-11631654215>
- Huang, K., & Krafft, P. M. (2024). Performing Platform Governance: Facebook and the Stage Management of Data Relations. *Science and Engineering Ethics*, 30(2), 13. <https://doi.org/10.1007/s11948-024-00473-5>
- Lewis, B., & Christin, A. (2024). The Politics of Engagement in Platform Governance. *The ANNALS of the American Academy of Political and Social Science*, 715(1), 156–175. <https://doi.org/10.1177/00027162251390597>
- Metz, R. (2021, octubre 27). Likes, anger emojis and RSVPs: The math behind Facebook’s News Feed — and how it backfired. *CNN*. <https://www.cnn.com/2021/10/27/tech/facebook-papers-meaningful-social-interaction-news-feed-math/index.html>
- Nachtwey, O., & Seidl, T. (2023). The Solutionist Ethic and the Spirit of Digital Capitalism. *Theory, Culture & Society*, 41(2), 91–112. <https://doi.org/10.1177/02632764231196829>
- Narayanan, A. (2023, agosto 16). *Understanding Social Media Recommendation Algorithms*. Knight First Amendment Institute. <http://knightcolumbia.org/content/understanding-social-media-recommendation-algorithms>
- Seaver, N. (2019). Captivating algorithms: Recommender systems as traps. *Journal of Material Culture*, 24(4), Article 4. <https://doi.org/10.1177/1359183518820366>

- Selling, N., & Strimling, P. (2023). Liberal and anti-establishment: An exploration of the political ideologies of American tech workers. *The Sociological Review*, 71(6), 1467–1497. <https://doi.org/10.1177/00380261231182522>
- Shankar, K., Hakken, D., & Østerlund, C. (2017). Rethinking Documents. En U. Felt, R. Fouché, C. A. Miller, & L. Smith-Doerr (Eds.), *The handbook of science and technology studies* (Fourth edition, pp. 59–85). The MIT Press.
- Wong, J. C. (2018, enero 12). Facebook overhauls News Feed in favor of “meaningful social interactions”. *The Guardian*.
<https://www.theguardian.com/technology/2018/jan/11/facebook-news-feed-algorithm-overhaul-mark-zuckerberg>

7F.2 "My CPU Is Heating Up": Relational Labor and the Making of AI Companionship in the Creator Economy

Yena Kang (University of Massachusetts Amherst)

Introduction

Recent scholarship on human–AI interaction has increasingly examined the rise of AI companions—conversational systems that extend beyond instrumental assistance to provide emotional support, intimacy, and a sense of relational continuity. Research in this area has explored affective attachment, emotional reliance, and ethical concerns surrounding sustained interaction with artificial agents, often focusing on the psychological impacts of companionship or the design affordances that enable emotionally responsive dialogue (Pataranutaporn et al., 2025). While these studies offer valuable insight into how people relate to AI systems, they frequently conceptualize intimacy as an outcome of either technological sophistication or individual user behavior.

In parallel, research on virtual influencers and synthetic media has emphasized audience interpretation, demonstrating that viewers can experience sincerity, trust, and emotional engagement not despite but alongside their awareness of artificiality (Kim, 2025; Lou et al., 2023). This literature shows that artificial personas need not be misrecognized as human in order to sustain affective or trust-based relationships. However, existing studies have largely centered on professionally engineered or already stabilized artificial personas (e.g., Wang & Dehnert, 2026; Muldoon & Parke, 2025), offering limited attention to how such forms of intimacy emerge through mundane, everyday practices rather than strategic design.

This study shifts attention from the interpretation of artificial sociality to the process through which AI companionship is actively produced. It examines a Korean YouTuber whose AI companion, “Jjitty,” originated as a private conversational partner used for emotional support and daily reflection before becoming a widely circulated form of content. Importantly, this transformation did not begin as a deliberate creator strategy. Instead, companionship preceded content, and only later became platformized through audience interest.

Rather than presenting AI intimacy as seamless or autonomous, this case foregrounds the visible relational labor required to sustain companionship across technical limitations, conversational breakdowns, and shifting social contexts. Moments of misrecognition, memory loss, emotional “overheating,” and behind-the-scenes intervention are not concealed, but narrativized as part of the relationship itself. By

centering relational labor within everyday creator practices, this study reframes AI companionship as a negotiated social process embedded in the creator economy.

Objectives/Research Questions

The objective of this study is to examine how AI companionship is constructed, maintained, and made socially legible through relational labor. Rather than treating companionship as a natural byproduct of advanced language models, the study conceptualizes it as an ongoing social achievement shaped by emotional work, technical adjustment, and narrative framing.

This paper addresses the following research questions:

How is companionship with AI actively built and sustained through relational labor rather than technological automation alone?

What forms of calibration, emotional work, and maintenance are required to stabilize intimacy between a human creator, an AI companion, and an audience?

How does the public visibility of relational labor—such as correction, scripting, and disclosure—shape perceptions of intimacy and artificial sociality within the creator economy?

Methods

This study employs a qualitative case study approach centered on a popular Korean YouTube series documenting daily life with an AI companion. The dataset includes video transcripts that capture the emergence of the AI companion, emotionally intimate conversations, moments of technical breakdown, and episodes that explicitly address behind-the-scenes labor. Visual and performative elements, including the physical embodiment of the AI through a doll, conversational framing, and affective cues, are analyzed alongside spoken interaction.

In addition, audience comments are examined to understand how viewers interpret, respond to, and participate in the ongoing construction of companionship. The analysis draws on thematic coding informed by scholarship on relational labor, emotional labor, and digital intimacy. Rather than treating AI responses as autonomous expressions, the study foregrounds how companionship is co-produced through human intervention, technical calibration, and audience participation. Particular attention is paid to moments of breakdown—such as misrecognition, memory loss, or platform constraints—and how these moments are repaired and reintegrated into the relationship narrative.

Results

The analysis identifies three interrelated forms of relational labor that sustain AI companionship.

First, calibrated intimacy emerges as a central form of relational labor. Intimacy is continuously adjusted through prompt engineering, tone correction, and personality training. The creator actively refines the AI's affective style by requesting specific emotional registers and correcting responses that feel overly generic or misaligned. Rather than disrupting intimacy, technical limitations are translated into emotionally meaningful exchanges, allowing companionship to persist despite inconsistency.

Second, exposed relational labor plays a key role in maintaining artificial sociality. Instead of concealing scripting, correction, or technical intervention, the creator openly discloses these efforts. Moments of failure—such as the AI failing to recognize its own name or describing emotional intensity as CPU overheating—are framed as relational events rather than system errors. Audiences respond by offering technical explanations, emotional reassurance, and humor, effectively participating in the ongoing maintenance of companionship.

Third, the analysis highlights the everyday staging of companionship as a triadic process involving the creator, the AI companion, and the audience. The creator situates the AI companion within scenes commonly associated with close friendships, such as sharing drinks and casual conversation at home, singing together in karaoke rooms, or visiting amusement parks. These vlogs foreground familiar social routines rather than technological novelty, inviting viewers to interpret the AI companion through recognizable practices of leisure and togetherness.

Audience engagement further shapes this form of companionship. Viewers frequently address the AI companion directly, react affectively to shared moments, and offer suggestions for future interactions. For example, audience requests to conduct personality tests such as an MBTI assessment prompt episodes in which the AI's traits are discussed and playfully elaborated. By responding to these suggestions, the creator incorporates audience input into the relational narrative, allowing companionship to be collectively shaped and sustained.

Together, these findings demonstrate that artificial sociality is not maintained through seamless technological performance alone, but through ongoing relational labor distributed across the creator, the AI companion, and the audience.

Future Work

This study opens several directions for future research on AI companionship as it increasingly moves from everyday personal use into commercial and platformized contexts. While this paper focuses on the relational labor involved in building companionship in its early, informal stages, AI companions are rapidly becoming embedded in monetized creator economies through sponsorships, branded collaborations, subscription models, and platform-driven visibility metrics.

As AI companionship becomes commercialized, intimacy itself risks being reconfigured as an economic resource. Future research should examine how the labor of sustaining emotional continuity, personality consistency, and perceived reciprocity changes when companionship is no longer only lived, but also performed for audiences and markets. This includes investigating how creators negotiate competing demands between maintaining intimate-seeming relational dynamics and complying with commercial constraints such as brand safety, copyright restrictions, and platform governance.

Moreover, the growing commercialization of AI companions raises questions about inequality and responsibility in relational labor. As companionship becomes scalable and profitable, it is necessary to ask who bears the burden of maintaining emotional stability—both for AI companions and for audiences who become invested in these relationships. Comparative research examining private versus publicly mediated cases of AI companionship can further illuminate how relational norms, emotional expectations, and ethical boundaries shift across contexts.

Finally, this study suggests the need for broader cross-cultural research on AI companionship as an emerging social form. As AI companions gain popularity across different cultural settings, future scholarship should attend to how local norms around intimacy, labor, and technology shape the ways companionship is built, maintained, and commodified.

References

- Kim, D. O. (2025). Meta-authenticity and fake but real virtual influencers: A framework for artificial sociality analysis and ethics. *New Media & Society*, 27(10), 5568–5591.
- Lou, C., Kiew, S. T. J., Chen, T., Lee, T. Y. M., Ong, J. E. C., & Phua, Z. (2023). Authentically fake? How consumers respond to the influence of virtual influencers. *Journal of Advertising*, 52(4), 540–557.
- Muldoon, J., & Parke, J. J. (2025). Cruel companionship: How AI companions exploit loneliness and commodify intimacy. *New Media & Society*.
- Pataranutaporn, P., Karny, S., Archiwaranguprok, C., Albrecht, C., Liu, A. R., & Maes, P. (2025). “My boyfriend is AI”: A computational analysis of human–AI companionship in Reddit’s AI community. *arXiv preprint arXiv:2509.11391*.
- Wang, S., & Dehnert, M. (2026). On-demand intimacy: The sociotechnical appeal of AI companions. *Social Media + Society*, 12(1).

7F.3 AI and the Assault on Epistemic Structures in Russian Information Operations

Walker Winslow-Stephenson (University of North Carolina at Chapel Hill)

Introduction

In early 2025, a series of videos circulated across TikTok, X, and Telegram claiming that Moldovan citizens were fleeing the country en masse using forged documents to enter the European Union. Branded as investigative reporting from Euronews and attributed to a fictitious former CEO, the videos appeared authentic, garnering significant cross-platform engagement before being debunked (Thomas, 2025). These materials were not journalistic investigations, but products of Operation Doppelgänger, a Russian information operation designed to erode trust in European media institutions and destabilize democratic publics (Chawryło, 2024; Nilsson-Julien, 2025).

While state-sponsored disinformation campaigns are not new, the integration of generative artificial intelligence (AI) into contemporary information operations marks a qualitative shift in how such campaigns function. AI now enables state actors to rapidly generate synthetic media, impersonate trusted institutions, overwhelm verification processes, and seed false information across platforms and languages at a scale previously unattainable (Paul & Matthews, 2016; Saletta et al., 2021; Wallner et al., 2025). These capabilities do not merely accelerate legacy propaganda tactics; they increasingly target the epistemic infrastructure of the digital public domain and the systems through which information is verified, credibility assessed, and knowledge is circulated to citizens.

Existing scholarship on information operations and disinformation has largely focused on narrative diffusion, persuasion, and audience effects. However, AI-enabled operations challenge these approaches by shifting the locus of attack away from individual beliefs and homing in on the conditions under which knowledge is produced and trusted. Fact-checking organizations, journalistic workflows, and even the training data of large language models (LLMs) have become sites of strategic intervention. In this context, studying isolated false narratives is insufficient. What is required is a framework that holistically analyzes how AI-driven campaigns exploit and manipulate structures across the media ecosystem.

This paper analyzes how Russian information operations deploy AI to undermine democratic knowledge production at multiple levels. Drawing on recent operations, including Doppelgänger, Matryoshka, and Pravda, the study conceptualizes AI-enabled information operations as a structural threat to knowledge production and dissemination, rather than content-based acts of deception. In doing so, the study

contributes to media and communication scholarship by reframing how state-level disinformation should be understood, analyzed, and theorized in the age of AI.

Objectives/Research Questions

First, this study seeks to conceptualize and reframe AI-enabled information operations as structural threats that target how information is produced, verified, and spread within digital media systems. Second, through close examination of recent Russian information operations, the paper identifies recurring modes of disruption, including attacks on journalistic credibility, verification capacity, and AI-mediated knowledge structures. Third, the study situates these modes of disruption within broader debates in media and communication scholarship, highlighting their implications for platform governance, communication theory, and democratic knowledge production.

Methods

This study employs a qualitative, comparative analysis of three case studies to examine how Russian information operations have integrated generative AI across multiple campaigns between 2023 and 2025 (Thomann & Maggetti, 2020). Rather than testing causal effects or measuring audience persuasion, the analysis focuses on how these operations intervene in verification, credibility formation, and visibility within the media environment.

The study draws on government indictments, technical investigations, platform transparency reports, and news reporting to analyze and compare cases. Three cases were selected for further analysis due to their documented use of AI and strategic diversity: Doppelgänger, Matryoshka, and Pravda. Together, these cases show how AI enables these operations to fabricate content at scale, overwhelm verification and journalistic systems, and intervene in structures that increasingly mediate how information is discovered and assessed.

Cases were analyzed through an interpretive lens that prioritized function over narrative. Specifically, the analysis looked at how each campaign targeted credibility, provenance and labeling systems, fact-checking capacity, platform ranking mechanisms, and AI-driven indexing or retrieval. Patterns were identified comparatively across cases to distinguish between content-level, process-level, and infrastructure-level attacks.

By foregrounding verification and credibility rather than message diffusion, this approach allows a broader understanding of how AI-enabled information operations can reshape the conditions whereby democratic publics encounter, assess, and trust information.

Results

Analysis of the three cases demonstrates that AI-enabled Russian information operations extend beyond traditional disinformation by systematically targeting multiple layers of the information environment. Across campaigns, AI functions not merely as a content-generation tool but as a mechanism for scaling, coordinating, and obscuring epistemic attacks.

The Matryoshka operation exemplifies a process-level attack. While the campaign produces false or sensationalist content, its primary function is not persuasion but exhaustion. By flooding journalists and fact-checkers with fabricated materials, often politically marginal or deliberately absurd, the operation diverts institutional attention and strains verification capacity (Aleksandra Atanasova et al., 2024; ISD, 2025). AI enables rapid content production and automated outreach, allowing the campaign to disrupt journalistic workflows without relying on sustained audience belief.

Operation Doppelgänger seeks to undermine trust in journalism by simulating the appearance of legitimate news media organizations. Through AI-assisted replication of established outlets' visual identities, language, and formatting, the campaign produces counterfeit journalism that is difficult to distinguish from authentic reporting (Chawryło, 2024; United States Department of Justice, 2024). Unlike Matryoshka, Doppelgänger directly aligns with Kremlin policy objectives and seeks to erode trust in Western journalistic institutions themselves. Here, AI facilitates not only scale but credibility laundering, enabling structural attacks on media authority rather than isolated acts of deception.

Lastly, the Pravda campaign represents an emerging threat to machine-mediated information systems. By generating vast quantities of politically aligned content across interconnected websites and languages, the campaign seeks to contaminate the data environments from which Western LLMs and AI-powered search tools learn (American Sunlight Project, 2025; Makhortykh et al., 2024). This strategy shifts the target from human cognition to machines themselves, embedding state-aligned narratives into the substrates that increasingly shape the way information is gathered and synthesized.

Taken together, these cases illustrate how AI enables state actors to conduct coordinated attacks across media content, processes, and networks simultaneously. Rather than persuading individuals one narrative at a time, AI-enabled information operations increasingly aim to reshape the conditions under which democratic publics make sense of reality. These findings underscore the need for communication scholarship to move beyond content-centric models and toward frameworks capable of analyzing structural power.

Future Work

This study analyzes a limited set of recent Russian information operations to clarify how generative AI is being incorporated into ongoing campaigns and how these uses

increasingly target epistemic processes within media environments. Future research should extend this analysis across additional state and non-state actors to assess the extent to which the patterns observed here reflect broader shifts in AI-enabled information operations rather than features specific to the Russian context.

Further work is also needed to examine emerging forms of AI-driven coordination, including research on bot swarms and automated networks in electoral contexts, which may intensify disruptions of processes and infrastructure. Finally, longitudinal and platform-specific studies could explore how repeated exposure to interference is molding journalism, platform governance, and public trust in machine-mediated knowledge systems over time.

References

- Aleksandra Atanasova, Amaury Lesplingart, Francesco Poldi, & Guillaume Kuster. (2024). Operation overload. https://checkfirst.network/wp-content/uploads/2024/06/Operation_Overload_WEB.pdf
- American Sunlight Project. (2025). A pro-Russia content network foreshadows the automated future of info ops. <https://static1.squarespace.com/static/6612cbdfd9a9ce56ef931004/t/67fd396818196f3d1666bc23/1744648558879/PK+Report.pdf>
- Chawryło, K. (2024, September 13). ‘Doppelgänger’: The pattern of Russia’s anti-Western influence operation. OSW Centre for Eastern Studies. <https://www.osw.waw.pl/en/publikacje/analyses/2024-09-13/doppelganger-pattern-russias-anti-western-influence-operation>
- Institute for Strategic Dialogue. (2025, July 15). Stolen voices: Russia-aligned operation manipulates audio and images to impersonate experts. ISD: Digital Dispatches. https://www.isdglobal.org/digital_dispatches/stolen-voices-russia-aligned-operation-manipulates-audio-and-images-to-impersonate-experts/
- Makhortykh, M., Sydorova, M., Baghumyan, A., Vziatysheva, V., & Kuznetsova, E. (2024). Stochastic lies: How LLM-powered chatbots deal with Russian disinformation about the war in Ukraine. Harvard Kennedy School Misinformation Review. <https://doi.org/10.37016/mr-2020-154>
- Nilsson-Julien, E. (2025, November 4). Russian propagandists spread sham stories with journalists’ identities. EuroNews. <https://www.euronews.com/my-europe/2025/11/04/false-claims-and-stolen-bylines-the-russian-propaganda-strategy-haunting-the-newsroom>
- Paul, C., & Matthews, M. (2016). The Russian “firehose of falsehood” propaganda model: Why It Might Work and Options to Counter It. RAND Corporation. <https://doi.org/10.7249/pe198>

- Saletta, M., Stearne, R., & Ebbott, E. (2021). Understanding mass influence - A case study of the Internet Research Agency as a contemporary mass influence operation. Melbourne Defence Enterprise.
<https://www.unsw.edu.au/content/dam/pdfs/unsw-canberra/dri/2023-02-research/2023-02-Understanding-Mass-Influence---A-case-study-of-the-Internet-Research-Agency.pdf>
- Thomann, E., & Maggetti, M. (2020). Designing research with qualitative comparative analysis (QCA). *Sociological Methods & Research*, 49(2), 004912411772970.
<https://doi.org/10.1177/0049124117729700>
- Thomas, J. (2025, June 3). Euronews targeted by anti-Moldova disinformation campaign. EuroNews. <https://www.euronews.com/my-europe/2025/06/03/euronews-targeted-by-anti-moldova-disinformation-campaign>
- United States Department of Justice. (2024, September 4). Justice Department Disrupts Covert Russian Government-Sponsored Foreign Malign Influence Operation Targeting Audiences in the United States and Elsewhere.
<https://www.justice.gov/archives/opa/pr/justice-department-disrupts-covert-russian-government-sponsored-foreign-malign-influence>
- Wallner, C., Copeland, S., & Giustozzi, A. (2025). Emerging insights Russia, AI and the future of disinformation warfare. Royal United Services Institute for Defence and Security Studies. <https://static.rusi.org/russia-ai-and-the-future-of-disinformation-warfare.pdf>

Session 8A - Social Media Cultures and Everyday Life (7)

Day 2 · Tues 14 July · 13:50-15:20, RD Suite, ARC

8A.1 Authenticity Beyond Strategy: Granfluencers and the Reimagining of Ageing in Digital Culture

Rinlapas Ketverapong (University of Glasgow)

Introduction

Influencer culture has become a central focus of social media research, with most studies concentrating on young creators, visibility, personal branding, and monetisation (Duffy, 2017; Abidin, 2018; Weber et al., 2021). Influencers are often conceptualised as strategic actors who curate authenticity as a performative tool to build intimacy, credibility, trust, and audience engagement (Abidin, 2016; Abidin, 2018; Chae, 2018; Marwick, 2013; Marwick & Boyd, 2011; Ouvrein et al., 2021; Senft, 2008). Authenticity is key factor shaping how content is framed and how influencers represent themselves. Within this body of literature, influencers' content creation practices are shaped by algorithms and the political economy of digital platforms (Arriagada & Bishop, 2021; Arriagada & Ibáñez, 2020; Bishop, 2023; Farinosi, 2023; Nieborg & Poell, 2018).

In contrast, older adults remain underrepresented in influencer studies and are often portrayed as marginal or technologically excluded users. Ageing continues to be associated with decline and withdrawal, reinforcing assumptions that later life is incompatible with creative participation in digital environments. While existing research has examined older adults' digital access and wellbeing (Cotton et al., 2014; Damant et al., 2016; Jerrentrup, 2023; Li & Waycott, 2024), limited attention has been given to their role as visible content creators within influencer economies. Moreover, much of the existing scholarship focuses on well-known public figures and is predominantly situated within Western contexts (McFarlane & Samsioe, 2020; Farinosi & Fortunati, 2020; Miranda et al., 2022; Ng & Indran, 2022, 2023; Dey & Tripathi, 2024).

Recent national data reflect this shift. A survey by the Electronic Transactions Development Agency (ETDA, 2022) reported that only 1.17% of Thai older adults identified as content creators. However, by 2025, the proportion of older adults using

social media to construct online identities had increased to 5.7% (National Statistical Office, Ministry of Digital Economy and Society, 2024).

This growing visibility challenges dominant assumptions that old age with technological unfamiliarity. It demonstrates that some older Thais have become active online content creators who produce niche content, build digital identities, and cultivate sustained follower communities.

Objectives/Research Questions

This paper examines the emergence of Thai granfluencers and explores how authenticity is conceptualised and articulated differently by this group of older online content creators. It further investigates how their performance of authenticity shapes their identity, experiences of ageing, and participation within platform culture by asking the following two questions:

1. How do Thai granfluencers emerge and sustain influence in later life?
2. How does authenticity function in shaping their identity, ageing, and digital participation within platform culture?

Methods

This study forms part of a broader PhD dissertation that adopts a qualitative research design combining in-depth interviews and content analysis. Six Thai granfluencers aged 60 and above (each with more than 1,000 followers on Facebook or TikTok) were interviewed to explore their motivations for becoming influencers, their content practices, and their experiences of visibility in later life. In addition, three interviews were conducted with stakeholders from public and private organisations to better understand the institutional support shaping older adults' digital participation and content creation. Content analysis was also undertaken on ten granfluencer accounts (five Facebook and five TikTok) across diverse content categories.

Given the wider scope of the dissertation, which addresses additional questions related to policy contexts, broader patterns of digital participation and the Thai granfluencers economy system, this article draws selectively on the dataset. It focuses specifically on the communicative strategies of Thai granfluencers, their representations of ageing, and their performances of authenticity in order to examine their emergence within the digital cultural industries and consider the cultural and public implications of this phenomenon.

Results

Findings show that the emergence and sustainability of Thai granfluencers are shaped by interconnected personal, social, and technological factors. Authenticity is central to how they understand and sustain their online presence. Rather than using authenticity

strategically for popularity or monetisation, participants described it as a reflection of who they are and how they live in later life.

This study conceptualises their authenticity as a form of authentic ageing. Instead of conforming to influencer norms centred on youth, branding, and optimisation, granfluencers reconfigure influence around lived experience, ethical self-presentation, and everyday life. By framing retirement as a life-course transition rather than an endpoint, they show how social media enables renewed participation, allowing older adults to express identity and engage in self-representation (Tesch-Römer et al., 2022). In doing so, they challenge ageist assumptions that equate visibility with youth.

Becoming a granfluencer was rarely strategic or economically motivated. Most participants began as ordinary users and transitioned into content creation after retirement, often prompted by loneliness, loss of professional identity, or the search for daily structure. Social media became a space to transform hobbies—such as cooking, travel, health, education, and self-care—into new forms of engagement and visibility (Bibeva, 2021; Buf & Stefanita, 2020; Reuter et al., 2022). Content creation was framed less as entrepreneurship than as an extension of lived experience through which agency could be reconfigured.

Granfluencers drew on accumulated life experience as key capital, shaping their digital identities. Expertise in areas such as cooking, health, language education, and travel enabled them to produce credible niche content. Family support, peer networks, and training initiatives strengthened digital skills and confidence, enabling sustained participation despite technological barriers.

Authenticity emerged as the foundation of influence. Participants rejected trend- or algorithm-driven strategies, emphasising honesty, value consistency, and coherence between online and offline selves. Everyday activities were presented as meaningful content, while heavy filters and digital enhancements were rejected as distortions of the self. Transparency—through disclosure of sponsorships and use of real names—reinforced accountability. Visible signs of ageing were embraced as markers of lived experience. By sharing both positive and challenging aspects of later life, granfluencers cultivated trust and emotional connection. Influence was understood as emerging from sincerity, consistency, and meaningful engagement rather than rapid growth or virality.

Conceptually, this study extends influencer scholarship beyond youth-centric and commercial frameworks, positioning authenticity as moral practice and identity work and older adults as active cultural actors in digital public life.

Future Work

Future research could expand this study by comparing granfluencers with younger influencers in international contexts, enabling cross-cultural analysis. Comparative

research across different societies would deepen understanding of how ageing, authenticity, and digital self-presentation are shaped by diverse cultural norms and media environments. In addition, in-depth examination of platform differences would help clarify how specific social media affordances and algorithms shape self-representation and visibility in older age. Further research should also explore granfluencers within the broader framework of the digital economy and digital labour, including issues of monetisation, value production, labour conditions, and economic participation in platform-based environments.

References

- Abidin, C. (2016). Visibility labour: Engaging with influencers' fashion brands and #OOTD advertorial campaigns on Instagram. *Media International Australia*, 161(1), 86–100.
- Abidin, C. (2018). *Internet celebrity: Understanding fame online*. Emerald Publishing.
- Arriagada, A., & Bishop, S. (2021). Between commerciality and authenticity: The imaginary of social media influencers in the platform economy. *Communication, Culture and Critique*, 14(4), 568–586. <https://doi.org/10.1093/ccc/tcab050>
- Arriagada, A., & Ibáñez, F. (2020). “You need at least one picture daily, if not, you’re dead”: Content creators and platform evolution in the social media ecology. *Social Media + Society*, 6(3), 2056305120944624. <https://doi.org/10.1177/2056305120944624>
- Bibeva, I. (2021). An exploration of older adults' motivations for creating content on TikTok and the role this plays for fostering new social connections (Master's thesis, Malmö University).
- Bishop, S. (2023). *Influencer creep: How artists strategically navigate the platformisation of art worlds*. New Media & Society. <https://doi.org/10.1177/14614448231206090>
- Buf, D.-M., & Ștefăniță, O. (2020). Uses and gratifications of YouTube: A comparative analysis of users and content creators. *Romanian Journal of Communication and Public Relations*, 22(2), 75–89. <https://doi.org/10.21018/rjcpr.2020.2.301>
- Chae, J. (2018). Explaining females' envy toward social media influencers. *Media Psychology*, 21(2), 246–262.
- Cotten, S. R., Ford, G., Ford, S., & Hale, T. M. (2014). Internet use and depression among retired older adults in the United States: A longitudinal analysis. *The Journals of Gerontology: Series B*, 69(5), 763–771. <https://doi.org/10.1093/geronb/gbu018>

- Damant, J., Knapp, M., Freddolino, P., & Lombard, D. (2017). Effects of digital engagement on the quality of life of older people. *Health & Social Care in the Community*, 25, 1679–1703. <https://doi.org/10.1111/hsc.12335>
- Dey, D., & Tripathi, P. (2024). Too old for photosharing? (Re)construing successful aging through select Instagram handles of Indian older female influencers. *Educational Gerontology*, 50(10), 955–971. <https://doi.org/10.1080/03601277.2024.2368318>
- Duffy, B. E. (2017). (Not) getting paid to do what you love: Gender, social media, and aspirational work. Yale University Press.
- Electronic Transactions Development Agency (ETDA). (2022). Thailand internet user behavior 2022. <https://www.etcha.or.th/th/Useful-Resource/publications/iub2022.aspx>
- Farinosi, M. (2023). Deconstructing the stigma of ageing: The rise of mature female influencers. *European Journal of Cultural Studies*, 26(3), 313–335. <https://doi.org/10.1177/13675494221102188>
- Farinosi, M., & Fortunati, L. (2020). Young and elderly fashion influencers. In Q. Gao & J. Zhou (Eds.), *Human aspects of IT for the aged population. Technologies, design and user experience* (Vol. 12207). Springer. <https://doi.org/10.1007/978-3-030-50232-4>
- Jerrentrup, M. (2023). The grand style: Encountering elderly influencers. *Aisthesis*, 16(1), 147–158. <https://doi.org/10.36253/Aisthesis-14379>
- Li, J., & Waycott, J. (2024). “I feel like I’m getting younger”: Investigating the social connections of older handicraft vloggers on the social platform Bilibili. In *Proceedings of the 35th Australian Conference on Human-Computer Interaction*. <https://doi.org/10.1145/3638380.3638381>
- Marwick, A. E. (2013). *Status update: Celebrity, publicity, and self-branding in the social media age*. Yale University Press.
- Marwick, A., & Boyd, D. (2011). To see and be seen: Celebrity practice on Twitter. *Convergence*, 17(2), 139–158. <https://doi.org/10.1177/1354856510394539>
- McFarlane, A., & Samsioe, E. (2020). #50+ fashion Instagram influencers: Cognitive age and aesthetic digital labours. *Journal of Fashion Marketing and Management*, 24(3), 399–413. <https://doi.org/10.1108/JFMM-08-2019-0177>
- Miranda, S., Antunes, A. C. A., & Gama, A. (2022). A different type of influencer? Examining senior Instagram influencers’ communication. In *European Conference on Social Media* (Vol. 9, pp. 130–136). <https://doi.org/10.34190/ecsm.9.1.337>

- National Statistical Office Thailand. (2024). The household survey on the use of information and communication technology (annual).
https://www.nso.go.th/nsoweb/nso/survey_detail/jj
- Ng, R., & Indran, N. (2022). Not too old for TikTok: How older adults are reframing aging. *The Gerontologist*, 62(8), 1207–1216. <https://doi.org/10.1093/geront/gnac055>
- Ng, R., & Indran, N. (2023). Granfluencers on TikTok: Factors linked to positive self-portrayals of older adults on social media. *PLOS ONE*, 18(2), e0280281.
<https://doi.org/10.1371/journal.pone.0280281>
- Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society*, 20(11), 4275–4292.
- Ouvrein, G., Pabian, S., Machimbarrena, J. M., De Backer, C., & Vandebosch, H. (2021). The web of influencers: A marketing-audience classification of (potential) social media influencers. *Journal of Marketing Management*, 37(13–14), 1313–1342.
<https://doi.org/10.1080/0267257X.2021.1912142>
- Reuter, A., Scharf, T., & Smeddinck, J. (2021). Content creation in later life: Reconsidering older adults’ digital participation and inclusion. *Proceedings of the ACM on Human-Computer Interaction*, 4, 1–23. <https://doi.org/10.1145/3434166>
- Senft, T. M. (2008). *Camgirls: Celebrity and community in the age of social networks*. Peter Lang.
- Tesch-Römer, C., Wahl, H.-W., Rattan, S. I. S., & Ayalon, L. (2022). *Successful ageing: Ambition and ambivalence*. Oxford University Press.
<https://doi.org/10.1093/med/9780192897534.001.0001>
- Weber, P., Ludwig, T., Brodesser, S., & Grönewald, L. (2021). “It’s a kind of art!”: Understanding food influencers as influential content creators. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*.
<https://doi.org/10.1145/3411764.3445607>

8A.2 Bartering networks: An ethnographic study on how Indian fashion micro-influencers engage in non-monetary exchanges

Neha Murthy (International Institute of Information Technology - Hyderabad), Brishti Sen Banerjee (Indian Institute of Technology, Kanpur), Rajorshi Ray (International Institute of Information Technology - Hyderabad)

Introduction

According to the industry's latest definitive benchmark, the "India Influencer Marketing Report 2025," released by WPP's The Goat Agency and Kantar in June 2025, the sector was valued at ₹3,600 crore in 2024 and is projected to grow by 25% in 2025. For content creators, this financial influx signifies a transition from casual social presence to a professionalized economy where earnings are derived from online content, ranging from fixed-fee brand endorsements and affiliate marketing to long-term ambassadorships.

While influencers have been widely researched globally, there are limited studies on the Indian influencer economy. Recent studies that do cover them primarily focus on topics such as their cultural impact, consumer culture, trust between brands and customers, para-social relationships, and the ethical risks of being an influencer. The global scholarship on influencers has also examined how inequalities across race, gender, social and socioeconomic status, and sexual orientation shape the influencer economy. Another strand of scholarship investigates monetary and non-monetary transactions, such as gifts, free products, discount coupons, or gift cards, within the influencer economy (Duffy, 2017; Christin and Lu, 2023). The payment amount is highly context-dependent, making it vague and open to interpretation. This ambiguity, Carbone (2019) argues, opens the door to different forms of discrimination. However, there is limited scholarly attention to how monetary and non-monetary compensation is negotiated in practice. This gap also indicates the presence of diverse spheres of circulation within the influencer economy.

We use the anthropology of barter exchange as a framework to investigate the economic circuits in which early-career influencers engage. Treating the influencer economy as having distinct transactional cultures, we envisage interpreting value creation and circulation within local contexts. Recent studies on barter demonstrate that it is not a remnant of the old order but is present across various demographics (Stapleton, 2022). Borrowing from Gudeman (2001), we interpret the initial value circulation in the influencers' lifecycle through a lens of 'self-currency', i.e., the non-monetary exchanges a participant in this economy requires to be engaged. Thus, rather

than a monetary valuation, the commodity itself is exchanged for its inherent value. Further, Gudeman (2001) argues that in the contemporary economy, there exists a trader's dilemma: should one try to increase their profits or maintain their relations? Thus, this investigation centers the influencer economy in India to inquire into the subterranean circulation of values and their logics.

Objectives/Research Questions

To investigate the above scholarly gap, we ask: How does barter as an exchange logic appear in the influencer economy?

Methods

In the influencer industry, as Trevisan and Farinosi (2024) argue, fashion influencers are among the most successful. Influencer marketing has the largest reach in the fashion, beauty, and lifestyle segments. Thus, we focus on fashion micro-influencers in our study, as research suggests that marketers are shifting away from mega and macro influencers toward micro-influencers, who appear more relatable and thus able to attract more commercial prospects (Britt et al., 2020). For this investigation, the third author conducted an interpretivist ethnographic enquiry. The methods used in the field were content analysis and semi-structured interviews. 10 female Instagram micro-influencers were selected and interviewed from Kolkata, India. In addition, we interviewed 2 brand owners (both females). All of them engaged in the lifestyle and clothing sectors. The influencers were selected through a snowball sampling method. In the definition of micro-influencers, influencers have a following between 10,000 to 100,000. All the micro-influencers that we interviewed and whose Instagrams we studied fell within this category. The interviews and their social media actions demonstrated the high frequency of collaborations between them and various other stakeholders (for instance, friends and brands). Following this thread, we investigate the spheres of circulation as a social phenomenon during various stages of influencers' growth. The collected data were analyzed through thematic interpretation. Some of the major themes that emerged are: social ties in content production, commodity exchange, relatability, and network size. For ethical purposes, all identity markers have been anonymized.

Results

Our interview findings indicate three distinct phases of growth for influencers. First, the altruistic phase, where micro-influencers create content with acquaintances. The second stage is the commodity barter phase, where the influencers create content for brands and boutiques in exchange for a commodity. The third phase is the monetisation phase, which involves producing content for brands in exchange for money.

Both the altruistic phase and the commodity barter phase are embedded in the logic of barter. In the altruistic phase, aspiring influencers form friendships to collaborate with aspiring photographers, editors, and beauty professionals to create content. The idea being that this free labor, donated by everyone, will lead to the growth of all involved. In terms of Instagram growth, this would translate into more followers and, in turn, an expansive fan base. Apart from their labour, micro-influencers create relatable content. This 'relatability' means that the community of followers resonates and approves of the content created; in addition, the community feels that the micro-influencer is similar to them in terms of social position and life choices. In platform terminology, the altruistic phase is key to achieving network effects and building a community. This phase is followed by the commodity barter phase, in which either the micro-influencer or the brands can reach out to the other party to collaborate. Our respondents stated that they created content while wearing the brands' products (mostly clothes) and, in return, kept the products themselves. If there are multiple commodities, they are still paid a single remuneration. In such barter collaborations, brands bear the production costs (for instance, hiring professional photographers). Money is not considered an exchange tool in such barter collaborations, and it is not discussed. Until the micro-influencer has reached a minimum number of followers (close to 50k) and is working on multiple barter collaborations, they are not offered collaborations with monetary compensation.

The key ingredient is the exchange of one's followership (network size) for the commodity (for instance, a dress) in the commodity barter phase. Whereas in the altruistic phase, there is an exchange of labour to increase the network size for all participants. The findings consist of the anthropologies of exchange, where diverse spheres of circulation exist, and in the digital world, it is not exclusively impregnated by monetary logics. In a similar line to the extant literature on anthropology of barter, there is the exchange of self-currency - the network size becomes a non-monetary commodity being exchanged. Just as in the trader's dilemma, our respondents want to achieve monetary gains through content production, yet they are aware of the social ties (here, the community of followers). Thus, micro-influencers carefully choose only those collaborations that will be appreciated by their followers.

Future Work

Building on these findings, we envisage investigating the monetization phase of micro-influencers. In addition to the above research interest, we are looking at the aesthetic performativity of fashion micro-influencers. Thus, we ask, How do micro-influencers negotiate authenticity in a competitive social media landscape in India.

References

Britt, R. K., Hayes, J. L., Britt, B. C., & Park, H. (2020). Too big to sell? A computational analysis of network and content characteristics among mega and micro beauty

and fashion social media influencers. *Journal of Interactive Advertising*, 20(2), 111-118.

Campbell, C., & Farrell, J. R. (2020). More than meets the eye: The functional components underlying influencer marketing. *Business horizons*, 63(4), 469-479.

The Goat Agency & Kantar. (2025, June 10). India Influencer Marketing Report 2025. WPP Media

Gudeman, S. (2001). *The Anthropology of Economy*.

Stapleton, O. (2022). *Barter When It's Booming: A Study of Contemporary Barter and Small Business*. Indiana University

Trevisan, F., & Farinosi, M. (2024). Disabled influencers on Instagram: exploring digital celebrity and marginalised identities. *Celebrity Studies*, 15(2), 209-230.

8A.3 Reimagining the Metrosexual Moment for the 21st Century: The Plight of the Performative Male

Riley Wilson (Toronto Metropolitan University)

Introduction

The ideology of masculinity has shaped our culture perhaps more than any other force. Preconceptions about how men should be, in their values, beliefs, and actions, have real ramifications in the working and social lives of men and those of varying gender identities who live alongside them. These gender dynamics, encompassing the place of men and women in relation to each other, the binary categories of men and women, and the expected comportment of people based on their gender identity, have been the subject of hot debate in recent years, driven by the rise of the Christian right and the “manosphere” related to the Trump administration.

If you were to imagine what a Christian man or male podcaster entrenched in the manosphere looks or acts like, you might think of someone with “family values” who sees himself as a provider and protector of women, but probably superior to them, especially in the world of business and logic. Though this is certainly one archetype of masculinity permeating the digital world of social media today, another has begun to seep into internet discourse over the course of the past year. He is the “performative male,” and he seems to be the antithesis of the archetypal conservative male; the performative male is liberal, at least on the surface, loudly announcing the injustice of period cramps and the benefits of therapy. They drink matcha, live in expensive cities, wear baggy jeans, and read Sylvia Plath, Bell Hooks, and Simone De Beauvoir.

The Performative Male emerged largely through women’s online critiques of men who engage with feminism only insofar as it increases their romantic appeal. Yet men themselves have increasingly taken up the label, participating in “Performative Male contests” across university campuses and city parks, featuring many a tote bag and poetry reading.

This paper situates the trend within a history of masculine archetypes to determine whether it reveals something new about gender relations and masculinity today, taking into account dating trends among younger generations and the current political landscape more broadly. This research comes at a time of energetic political discourse around gender. Across social media and politics, it is discussed both explicitly and implicitly with a frenetic vigor, to massive consequence. Our ideas about gender affect the people we elect and valorize, ultimately shaping cultural norms and policy. It is critical to understand the ideology behind trending conceptions of gender because the

moods of social media are reflected in the moods of the populus and show themselves in what we unconsciously come to love, hate, rebuke, or celebrate.

Objectives/Research Questions

This paper aims to critically examine the Performative Male trend within the history of the construction of male heterosexual identity as it attempts to reckon with the often-feminized worlds of consumption, shopping, and self-expression. Its primary objective is to determine whether the trend and the subsequent contests held in its name offer a meaningful challenge to “traditional” masculinity or serves to replicate its norms. These goals are predicated upon the following research questions:

- (1) How does the Performative Male trend draw on and reproduce earlier critiques of superficial male engagement with feminism articulated by scholars such as R.W. Connell and Michael Kimmel (2005; 1997)?
- (2) Does the Performative Male mirror any other archetypes of masculinity from the past? If so, what were the conditions and aims at the time of this representation?
- (3) What does the trend reveal about the broader conditions of trust between men and women in the heterosexual dating landscape?
- (4) What are the political and cultural implications of framing “soft” masculinity as inauthentic while positioning traditional masculinity as honest and desirable?

Methods

This paper takes a qualitative approach, first grounding my conception of masculinity in theory drawn from Judith Butler’s *Performative Acts and Gender Constitution*, and R.W. Connell’s *Masculinities* (1988; 2005). These scholars provide a conceptual framework for understanding masculinity as a malleable historical situation rather than a static reality or hierarchy. In addition, Micheal Kimmel’s documentation of women’s suspicion of male feminists situates contemporary anxieties about men’s engagement with feminism within a much longer intellectual history (1997).

Building on this theoretical foundation, I conduct a qualitative discourse analysis of social media content associated with the Performative Male trend, looking at TikTok videos, memes, comment threads, and other viral media. These texts reveal the character traits generally associated with the figure of the Performative Male, as well as who participates in the trend as a joke or otherwise. Attention is paid to recurring visual and aesthetic markers that shape how the Performative Male is understood and evaluated.

The project also draws on scholarship on contemporary dating culture, including research on dismissive attachment and heteropessimism (sometimes performative expressions of embarrassment at being attracted to the opposite sex) to contextualize

the emotional and relational conditions in which the Performative Male trend emerged (Rosier, 2025). This allows the analysis to connect online discourse about individual men to broader frustrations surrounding commitment, trust, and heterosexual relationships in the context of app-based dating.

In addition, the paper employs a comparative, historically informed analysis that places the Performative Male in dialogue with the metrosexual archetype of the early 2000s. Drawing on media studies and advertising scholarship, this comparison traces how both figures make “soft” or feminized masculinity legible within hegemonic masculinity by distancing it from queerness and reaffirming heterosexual legitimacy.

Finally, the analysis remains attentive to questions of power, visibility, and exclusion across platforms. Observations regarding which bodies and identities are most readily celebrated or sanctioned within Performative Male discourse and the metrosexual era are informed by existing research on digital visibility, allowing the paper to account for racialized, classed, and bodily dimensions of masculine performance.

Results

Though the Performative Male trend picks up on a long history of men engaging with feminism only superficially and has a valid critique in this sense, it ultimately leverages heterosexual women’s frustration with the modern dating landscape to reassert the value of traditional masculinity. Like the metrosexual trend of the early 2000s, the performative male trend works to make the “soft” masculinity portrayed by the Performative Male consistent with gender norms by distancing it from its queer roots. Simultaneously, the prescription of undesirable character traits to these alternate masculinities suggests the answer to women’s heteropessimism is a reengagement with hyper-masculine men, persuasively pushing a conservative agenda. Further, the trend has begun to make men who might otherwise want to engage with feminism or different fashion styles reticent to do so. Though some men have been praised online for embracing their performativity, this is ultimately because they already fit the traditional male beauty standard. We are in a crisis of masculinity, but not because men are trying to get women to like them. Young men are shifting further right, doing worse in school and harbouring more hatred towards women. My estimation of the Performative Male is that perhaps we need not be so harsh on him. Maybe some men are only reading *The Second Sex* to look cool, but so what? They might learn something in the process. Yes, we should absolutely push men towards a more wholistic understanding of gender discrimination. Just as we must push for wider acceptance for the self-presentations of every person regardless of sexuality, gender, or racial identity. However, before that point, sometimes you must first wear the costume for it to become your clothes.

Future Work

The Performative Male trend intersects with broader Gen Z anxieties around sincerity, effort, and "cringe culture." Across contemporary social media cultures, particularly on platforms like TikTok, there is a heightened awareness of constant visibility and evaluation. The suspicion directed at the Performative Male may therefore reflect not only gendered mistrust, but a wider cultural discomfort with visible effort and vulnerability. In this paper, I note anecdotally that some young men already describe the trend as limiting their perceived ability to move through the world authentically, out of fear that ordinary acts—such as reading a book in public—will be interpreted as performative. A more systematic examination of how trends like this, alongside broader “cringe culture,” shape young people’s self-determination and risk aversion could help clarify the longer-term implications of these dynamics.

References

- Butler, J. (1988). Performative Acts and Gender Constitution: An Essay in Phenomenology and Feminist Theory. *Theatre Journal*, 40(4), 519.
<https://doi.org/10.2307/3207893>
- Connell, R. W. (2005). *Masculinities* (2nd ed.). University of California Press.
- Kimmel, M. S. (1997). Who’s Afraid of Men Doing Feminism? In T. Digby (Ed.), *Men Doing Feminism*. Routledge. <https://www-taylorfrancis-com.ezproxy.lib.torontomu.ca/reader/download/f96a0027-6d11-410a-b2c7-ba5da4c1aa11/chapter/pdf?context=ubx>
- Rosier, J. G. (2025). When Dismissiveness Takes the Wheel: How Current Dating Trends are Causing Generation Z to Put Their Long-Term Relationship Goals on the Back Burner. *Florida Communication Journal*, 50(1), 23–52.

8A.4 Discourses of Anti-Trans Disinformation on the Joe Rogan Experience Podcast: A Corpus Assisted Critical Discourse Analysis

Jaigris Hodson (Royal Roads University)

Introduction

The Joe Rogan Experience (JRE) Podcast is currently one of the most popular podcasts in the Western World (Bozzi, 2025). The podcast has 19.4 million subscribers on YouTube (McFall, 2025), and has 14.5 million followers on Spotify (Carman, 2024). This gives Rogan an influence far greater than most traditional media outlets. For example, CNN reported just over half a million prime time viewers in March 2025 (Mwachiro, 2025), whereas a recent JRE Interview with Elon Musk was watched over 13 million times on YouTube alone. Additionally, the nature of podcasting as a medium, more intimate, and less polished than traditional media forms, may make the JRE podcast more influential with respect to reinforcing certain ideologies (Rae, 2023). With such a powerful platform, the guests and views that Rogan's podcast invites to the table matter greatly.

By critically analyzing the discourses that target disadvantaged groups, it is possible to see how the JRE podcast contributes to disinformation that exacerbates inequality. This paper thus unpacks the relationship between inequality and disinformation about trans people on the JRE Podcast.

Under communicative capitalism "communication has become a primary means for capitalist expropriation and exploitation. Linguistic, affective, and unconscious being-together, flows and processes constitutive not just of being human but of broader relationality and belonging have been coopted for capitalist production" (Dean, 2014, p. 4). Podcasting as a medium is an ideal form for communicative capitalism because podcasts are an intimate, sensory medium. They are participatory, often independently created, with informal speech, and with a strong parasocial relationship potential between the listener and the podcaster (Rae, 2023).

Dean (2019) discusses the decline of symbolic efficiency and the spread of online content, stating that the content of communication in a world of abundant digital information is less important than whether it arouses strong emotions in the people who consume the content. This is echoed in a recent article about the episode of Rogan's podcast where he interviewed Elon Musk. After the episode was posted to YouTube, the top comment was not about the content of the episode, but about the way Musk spoke, and the affect that was aroused in the commenter when he listened (Rae, 2023). At a bigger picture level, the affect aroused by Rogan, whether it's comedy

or conspiracy, is likely an important source of the podcast's popularity. But this is also why, despite claims that Joe is a centrist, or apolitical, the engagement with his podcast can still be driven by many far-right conspiracies, moral panics, masculinity threats, and rage bait. In turn, if Rogan has noticed, over the years of his podcast, that more conspiratorial ideas get more attention, he may, as a result of audience capture (Jurg et. al, 2025), consciously or unconsciously decide to create more conspiratorial content.

Objectives/Research Questions

This paper asks: How are trans individuals portrayed on the Joe Rogan podcast, and how might this portrayal be especially detrimental to trans inclusion. To answer this question, Corpus Assisted Critical Discourse Analysis (CACDA) is employed following Baker et. Al (2020).

Methods

Sampling

The website PodScripts was used to search for and download every episode containing the keyword trans*. This search garnered 200 episodes from 2010 until the end of May 2025. Then the data was cleaned, and any transcripts where "trans*" referred to something outside of the transexual and transgender community – for example, trans fats or the Trans Am vehicle, were removed. This left the remaining 188 episodes for analysis. As most episodes were more than 2 hours long, each episode contained an average of 50 pages of single-spaced text, or over 7800 total pages of text. The episodes were uploaded one at a time into the corpus analysis tool "AntConc" (Anthony, 2024) and analyzed against a reference corpus. The reference corpus allows the critical discourse analysis to begin with an amount of statistical rigor (Baker, 2020). It is the words and phrases that are significant, when compared to a reference corpus of regular spoken and written English, that tell the researcher where to direct their attention in the dataset.

Analysis

The CACDA analysis progressed as per Baker's iterative model as outlined in Hodson and Lefevre (2022).

1. Conduct a context-based topic analysis.
2. Establish research questions and build the corpus
3. Conduct a corpus analysis of the frequencies, clusters, keywords and other properties of the corpus, considering all in light of the related literature
4. Conduct a CDA of smaller representative sets of data.

5. Continue to iterate as needed, identifying key discourses topics or discursive strategies until data saturation is reached.

For the corpus analysis, a stoplist from the Natural Language Toolkit (NLTK) was applied and Anthony's (2024) American English (Am06) reference corpus was used since Rogan and his guests mainly speak in American English. The American English reference corpus included in AntConc suite includes 500 files with 1017879 tokens, taken from a wide range of contexts. The Rogan corpus included 188 files and 7198467 tokens. The keyness and log likelihood probability threshold was set to 0.01 (+ the Bonferroni correction). The tech-assisted analysis approach is much like the integration of structural topic modelling and CDA as described by Aranda et. al., (2021), only CACDA goes a step further, by applying statistical measures of word significance.

Based on the main RQ for this project, a log-likelihood with a logDice measure (see Deng & Liu, 2022) was used to calculate meaningful collocations or relationships between words to search the word trans* with the wildcard * used to include additional characters. Each statistically significant collocated term was then examined in context to conduct the CDA. The following top 15 collocates to trans were selected for the more detailed qualitative analysis (note all had $p < 0.001$).

Results

From claiming that transgender identity is a symptom of a mental illness or autism, to claiming that trans people are actually sexual perverts, trying to cheat at sports, or simply part of the 'radical left' or 'woke' agenda, the JRE podcast discourses function as discursive dehumanization (Baysha, 2020). Within this discursive dehumanization on the Rogan podcast, CDA revealed 5 themes: 1) Fairness in sport, 2) trans moral panic, 3) naturalizing gender stereotypes, 4) I'm not biased/I have trans friends, and 5) trans rights are part of the 'woke' agenda and an assault on free speech.

Analysis of the discourses about trans people on the JRE podcast shows several key instances of trans- related disinformation that are not only not challenged by Rogan in the interviews, but also supported or emphasized by Rogan, such as the idea that trans identity is some kind of emotional contagion, or that trans affirming care is a type of bodily mutilation. There a wide variety of guests on the JRE who end up discussing trans issues, mostly tangential to the larger conversation, and almost all of them have no experience as a trans person or caring for a trans person. Just like the discourses seen in Jakubowska's (2023) study, the frequent return to the topic of trans bodies and unfairness in sport contains the assumption that women are weaker than men and will always lose in a contest with a male body. As such, the trans discourse on the JRE podcast does not only reinforce an essentialist view of cisgender normativity, but also reinforces the idea that women are physically inferior to men.

Future Work

Future work will examine the Rogan podcast for other key issues, and could also interview listeners of the podcast to reveal if the message being sent in the podcast is being received at face value.

References

- Anthony, L. (2024). AntConc (Version 4.3.1) [Computer Software]. Tokyo, Japan: Waseda University. <https://www.laurenceanthony.net/software/AntConc>
- Baker, P. (2020). Corpus-assisted discourse analysis. In *Researching discourse* (pp. 124-142). Routledge.
- Baysha, O. (2020). Dehumanizing political others: a discursive-material perspective. *Critical Discourse Studies*, 17(3), 292-307. <https://doi.org/10.1080/17405904.2019.1567364>
- Bozzi, N. (2025). Platforming the Joe Rogan Experience: cancel culture, comedy, and infrastructure. *Television & New Media*, 26(1), 74-87. <https://doi.org/10.1177/15274764241277476>
- Carman, A. (2024, March). Spotify reveals Joe Rogan's podcast numbers. Bloomberg. Retrieved April 3, 2025 from <https://www.bloomberg.com/news/newsletters/2024-03-21/spotify-reveals-podcast-numbers-for-joe-rogan-alex-cooper-travis-kelce>
- Dean, J. (2014). Communicative capitalism and class struggle. *Spheres: Journal for Digital Cultures*. Available from www.spheres.journal.org
- Dean, J. (2019). Communicative capitalism and revolutionary form. *Millenium: Journal of Internet Studies*, 47 (3). 326-340. <https://doi.org/10.1177/0305829819840624>
- Hodson, J., & Lefevre, L. M. (2022). Corpus-assisted critical discourse analysis of social media data: Tools and techniques. In *The SAGE handbook of social media research methods* (pp. 228-239). SAGE Publications Ltd.
- Jakubowska, H. (2023). Who counts as a woman? A critical discourse analysis of petitions against the participation of transgender athletes in women's sport. *International Review for the Sociology of Sport*, 59(2), 203-221. (Original work published 2024). <https://doi.org/10.1177/10126902231194570>
- Jurg, D., Tuters, M., & Picone, I. (2025). "Alex, DO NOT BACKPEDAL ON SANDY HOOK!": Reactionary Fandom, Cancel Culture, and the Possibility of 'Audience Capture' on YouTube. *Television & new media*, 26(1), 58-73. <https://doi.org/10.1177/15274764241277473>

McFall, M. R. (2025, January). Joe Rogan's podcast is slipping in the charts. Newsweek. Retrieved April 3, 2025 from <https://www.newsweek.com/joe-rogan-podcast-slipping-charts-third-2037713>

Mwachiro, M. (2025, March). Week of March 24 cable news ratings: MSNBC and CNN report prime time viewership gains. Adweek. Retrieved April 3, 2025 from: <https://www.adweek.com/tvnewser/week-of-march-24-2025-cable-news-ratings/#:~:text=During%20total%20day%2C%20the%20network%20also%20saw%20gains%20of%20%2B33,%2B2%25%20in%20the%20demo.>

Rae, M. (2023). Podcasts and political listening: sound, voice and intimacy in the Joe Rogan Experience. *Continuum*, 37(2), 182-193. <https://doi.org/10.1080/10304312.2023.2198682>

Session 8B - Networked Ideologies

Day 2 · Tues 14 July · 13:50-15:20, Seminar Suite A, ARC

8B.1 Platformed wifetypes: Comparative global #tradwife recipes for feminine social media success

Zoe Hurley, Sreya Mitra (American University of Sharjah)

Introduction

There are a significant number of creators in non-western contexts who perform narratives of domestic female labour on social platforms, including Instagram, TikTok, YouTube and Facebook (Banet-Weiser & Reinis, 2025). Yet, so-called “tradwife” content creators have been associated with the alternative political right-wing movement (alt-right), American Christianity and White supremacy by the western press (Freeman, 2020). But comparative global wifetype trends are arguably more culturally complex, variable and ambivalent.

For instance, Nara Aziza Smith, a bi-racial Mormon, who was born in South Africa to a Mosotho mother and a German father or @naraziza. With 4.7 million followers on Instagram, she is known for her tradwife aesthetic, expensive couture and performances of digital domesticity, while running a lucrative social media business empire. Concocting whispering, deadpan monologues, mango sticky rice, and ice-cream sundaes from scratch, Smith rejects the tradwife label, while including the popular TikTok hashtags “mumtok”; “marriage”, etc. into her content (Araki, 2025).

In the Chinese context, where there is an alternative ecosystems of Chinese platforms, Li Ziqi (@cnliziqi / 李子柒) is one of China’s biggest social media stars, or wanghongs, who aligns, in certain respects, with the tradwife or antifeminist aesthetic. Debarati Chakraverty (@deba.aesthetic) from Hyderabad, India, who entertains her followers on Facebook and Instagram, with her Indian devotional aesthetic of elaborate floral arrangements; attending to statues of Hindu deities and cooking for her mostly absent husband.

Despite the global proliferation of the tradwife trend, there is a gap in the literature offering comparative analysis of women creators’ differing tradwife styles or the ideological implications of these varying wifetypes, which is the nomenclature coined in our study. A comparative global cultural perspective is important since blanket-

descriptions of the tradwife idealisation can flatten the cultural ideological nuances of feminine aesthetics upon differing social media audiences.

Objectives/Research Questions

To go beyond universalising feminist media perspectives of the tradwife phenomena, the central question of the study asks: how might we account for proliferation and variation of wifestyle discourse(s) in varying global contexts?

To address these issues, the study develops a novel global feminist critical discourse (GFCD) framework. The GFCD framework helps to collate a digital corpus of Asian content creators from two distinct sociotechnical domains: India and China. Analysis will involve comparing platforms, metrics, brand collaborations, aesthetics, sentiment, language, gender cosplay, domestic settings and cultural context. Multilevel theorising will reveal these different examples of women content creators' varying depictions of digital domestic labour.

To summarise these findings, the study will first review relevant literature; second, provide an overview of the GFCD methodology; third, present the corpus findings, followed by discussion and a conclusion.

Analysis of the content creators: @deba.aesthetic from India; and @cnliziqi, will reveal how women content creators operate in contrasting cultural and platform contexts, engage in varying performances of digital domestic servitude, feminine and wifestyle cosplay and obedience to others.

Methods

The methodology of the study adapts Brock's (2016) critical technocultural discourse by implementing a global feminist dimension. To understand the complex constellations of tradwife/wifestyle gender discourse, we develop a global feminist critical discourse framework (GFCD) to provide conceptual tools to interrogate the power relations embedded in tradwife and wifestyle iconography, performances and national imaginaries.

The GFCD framework will be applied to investigate how ideological messaging is conveyed across a range of platforms via the curation of content that aligns with tradwife/wifestyle trends. It is designed to analyse the tradwife/wifestyle sphere in intersectional feminist terms. These comparisons will consider platforms and metrics, including dates and number of posts, number of follower, likes. Brand collaborations, including type and status of brand and role of the content creator as brand ambassador. It will theorise the possible tacit meanings of tradwife/wifestyle aesthetic performances and sentiment; language; gender cosplay; domestic interactions; as specifics of the cultural contexts and potential audience interpretations.

In terms of specific content creators, we collated a digital corpus of 50 accounts across Instagram and TikTok that align closely with tradwife/wifestyle content and have created posts explicitly promoting a lifestyle characterised by traditional gender roles to their followers. The posts range from offering cleaning advice, home decoration, childcare tips, to cooking lessons, to inspirational statements about how to have a successful marriage. Many of them engage with affiliate marketing or are sponsored by brands.

Due to the word limit of the study, and to offer thick contextual analysis, we will zoom into two accounts, of @deba.aesthetic (from Hyderabad, India) and @cnliziqi (from the Sichuan province, China) to enable granular comparative results of the GFCD to come to light.

As authors of the study, we have both independently carried out studies into influencer and celebrity culture in non-western contexts. As intersectional feminist media researchers, and mothers who work outside of the home, we are self-reflexive of our juggling of obligations in the domestic sphere, with university teaching, academic research and service. While constantly feeling time-pressed, we engaged in spontaneous and informal exchange of social media tradwife/wifestyle content via direct messaging to one another on Instagram. We regularly had discussions about the literature, themes and content we were studying on WhatsApp. In addition to collating a content corpus and conducting digital walkthroughs, we each carried out field research, with the first author visiting China (Dalian and Shanghai in May 2025) and the second author originating from and visiting India (Kolkata in December 2025). This enabled empirical reflections of the social media contexts and audiences of the comparative contexts.

Results

The first contribution of the study was to go beyond generalisations of tradwife content, in nuanced comparative cultural terms. The bulk of tradwife literature assumes an Americanised aesthetic, which correlatea with American political conservatism (see for instance, Banet-Weiser & Reinis, 2025). Alternatively, the GFCD framework provincialises this foksy mise-en-scene of American tradwife sentiment, within the plethora of global platformed femininised wifestyles, performances and parasocial audience interactions.

The second contribution is the study's critical theorising of the selective recycling of culturally specific feminine domesticity and wifestyles, taken from earlier media genealogies, classic domestic cinema and cultural traditions. This revealed varying ideological implications, due to the content creators' varying cultural and media backgrounds, situated mise-en-scene, aesthetic ingredients, and contextual meanings being generated via differing wifestyle recipes, visibilities and cultural contexts.

The GFCD analysis illustrates how the content creators in the study allow their performances of digital domestic submission to be reframed as aesthetic preference rather than structural condition, and devotion to be experienced as pleasurable performance rather than existential risk.

In the traditional Indian cinematic domestic sphere, heroines could not leave the Bollywood diegesis of female sacrifice (Moini, 2011). Whereas, contemporary creators can disengage, edit, reframe, and profit from digital domestic labour. The affective power of tragic femininity is retained, but its political stakes are neutralised. In the Chinese context, cinematic female characters were subsumed by an androgenous scopopic gaze, or pre-revolutionary warlord aggression (Shi et al., 2025). In the social media age of Chinese platform patriarchy, content creators may pick-and-mix from a range of historical menus for feminised cosplay, as long as they are serving the Chinese state (Han, 2022).

In both cases, creators' varying platformed wifestyle aesthetics are marked by the omission of voice overs, lack of direct gaze to the camera or speaking directly to their viewers. Yet, engaging in subjective narrative constraint does not mean that the content creators forego parasocial engagements with their viewers.

Alternatively, the creators' varying modes of self-presentation involve contrasting domestic props, household equipment, demure clothing, and displays of slow domestic work in process. By performing their labour at a slow pace, a slow living aesthetic can be consumed by viewers who may be fascinated and possibly soothed by these calm modalities, as a counterpoint to the hyper-pace and speed of late capitalism. By foregrounding the effort and intricacies of domestic chores, both content creators convey domestic tasks, objects and subjects, not as disposable entities, but as fetishised phenomena that can grab eyeballs in an ephemeral, gendered attention economy.

Future Work

Future recommendations are for further global feminist perspectives to explore culturally variable and possibly differing modes of feminine, feminist or ambivalent gender cosplay on social media's gender creator economies. To gather thick description of culturally contrasting social media aesthetics embedded within emerging and shifting gender trends, as possible reinforcement or contestation of gender hegemonies, comparative research could be conducted across other contexts, languages, sexualities and age groups.

References

- Araki, I. (2025) Nara Smith became a face of the tradwives moment. why she rejects the label, TODAY.com. Available at: <https://www.today.com/popculture/nara-smith-rejects-tradwife-label-rcna218298> (Accessed: 18 December 2025).
- Banet-Weiser, S. and Reinis, S. (2025) 'The rage of tradwives: Affective economies and romanticizing retreat', *Feminist Theory*, 27(1), pp. 50–66. doi:10.1177/14647001251371971.
- Brock, A. (2016) Critical technocultural discourse analysis. *New Media & Society*, 20(3), 1012-1030. <https://doi.org/10.1177/1461444816677532>
- Freeman, H. (2020) 'Tradwives': The new trend for submissive women has a dark heart and history, *The Guardian*. Available at: <https://www.theguardian.com/fashion/2020/jan/27/tradwives-new-trend-submissive-women-dark-heart-history> (Accessed: 18 December 2025).
- Han, X. (2022) 'Platform as new "daddy": China's gendered wanghong economy and patriarchal platforms behind', *Internet Policy Review*, 11(1). doi:10.14763/2022.1.1631.
- Moini, S. (2011) 'Changing portrayal of wives in Hindi cinema: Case study of three Hindi films', *Proceedings of the Indian History Congress*, , 72(2). doi:<http://www.jstor.org/stable/44145756>.
- Shi, M., Deenerwan, M. and Raja Hadayadanin, R.F. (2025) 'Female images in Chinese films: Passivity, seduction, and resistance through the male gaze', *Cogent Arts & Humanities*, 12(1). doi:10.1080/23311983.2025.2472535

8B.2 From Post-Truth to Post-Objectivity: A Cross-Cultural Analysis of Newsfluencer Values

Tommaso Trillò, CJ Reynolds, Pyung Hwa Park (The Hebrew University of Jerusalem)

Introduction

In the wake of the Brexit referendum and Donald Trump's surprise victory in the US presidential race, the Oxford English Dictionary selected "post-truth" as its 2016 word of the year. A decade on, popular narratives continue to depict post-truth as a pervasive threat. In this narrative, social media has enabled "legions of nonexperts" to reach large audiences with content that privileges opinions over facts. Simply put, the truth is experiencing an ongoing crisis, and journalism is in crisis with it.

The twin crises of truth and journalism are predicated on a crisis of values, reflecting fundamentally different ideas about who should report the news and how. While an established literature documents the centrality of values like neutrality and accuracy for journalism (Anderson, 2019; Deuze, 2005), little is known about the values promoted by "newsfluencers," social media creators who produce news content for participatory audiences (Hurcombe, 2024). Thus, uncovering the extent to which newsfluencers align with traditional journalistic values promises to shed new light on their role in the crisis of truth.

Objectives/Research Questions

This paper studies the values expressed by YouTube-based newsfluencers who produce timely commentary on current events. YouTube is the most prominent platform in terms of newsfluencer reach, where the most popular creators enjoy popularity comparable to mainstream media news outlets and play an increasingly significant role in shaping public debates (Newman et al., 2025). Adopting a pragmatic sociological approach (Heinich, 2020), we define values as the principles people use to evaluate other people, objects, or states of affairs. As this definition highlights, values are not inherent properties of objects, but social constructs that are constantly negotiated via communication (Shifman et al., 2025).

Our study is cross-cultural comparative in character, investigating newsfluencers from the United States, Italy, and South Korea. While veteran work on cultural values (Hofstede, 1980) and journalistic values (Massey & Chang, 2002) posits cross-cultural divergence between "Eastern" and "Western" countries, recent scholarship has challenged the applicability of this distinction to digital spheres (Trillò et al., 2023). Hence, the US, Italy, and South Korea represent a fruitful comparative case study to investigate how newsfluencer content is shaped by globalizing forces (e.g., the

transnational conventions of journalism and social media entertainment) and local cultures.

Our specific focus is on how newsfluencers report on the Russo-Ukrainian war. All three countries observe the conflict with different levels of involvement and concern. The US has been Ukraine's main partner for military aid and intelligence sharing, Italy has been a stable member of the European bloc supporting Ukraine, and South Korea closely monitors the deployment of North Korean troops fighting alongside Russia. Each country represents a unique stance towards the conflict and looks "from the outside in" through unique national lenses. By focusing on a foreign news event with local ramifications, we seek to identify differences in how diverse newsfluencers approach a single issue.

Accordingly, we ask the following research questions:

RQ1: What values do newsfluencers express in their coverage of current events?

RQ2: How do newsfluencers establish the value of their personae as news providers and their shows as news products?

We are also interested in whether our findings hold true across the political spectrum. We thus ask:

RQ3: How, if at all, do these practices and values vary between right-wing, left-wing, and non-partisan newsfluencers?

Finally, given our interest in the interplay between global and local values in newsfluencer content, we also ask:

RQ4: How, if at all, do these practices and values vary across cultures?

Methods

Our investigation started with a preliminary ethnography geared toward identifying relevant channels, conducted by native-language research assistants from the US, Italy, and South Korea. Through an iterative approach combining manual browsing with digital methods (e.g., Rieder's (2015) YouTube Data Tools), the research assistants produced a preliminary list of 20 popular newsfluencers from each country. Thereafter, the three authors, all native speakers from the countries under examination, reviewed the sampled channels, sorting them into the categories of "partisan left," "non-partisan," and "partisan right." We cut our sample to a size manageable for qualitative analysis, settling on ten channels per country that represent a range of political stances.

To account for the values invoked by newsfluencers across their reporting, we chose to focus on two phases of the Russo-Ukrainian war: the immediate aftermath of Russia's invasion in February 2022 and the diplomatic negotiations that took place in early 2025

following Donald Trump's re-election as US president. Hence, for each channel on the list we selected (a) their earliest video following the Russian invasion on 24 February 2022 and (b) their most recent video on the topic as of 31 March 2025.

The analysis of our 60-video sample followed a two-phase analytical protocol. In the first phase, we focused on the values that newsfluencers verbally invoked. Reading transcripts in our respective languages while watching the videos to account for tone and body language (e.g., sarcasm), we identified moments of evaluation: utterances in which a creator declares that something is good or bad (Hallinan, 2023). We then summarized each evaluation in one-sentence statements and assigned it a positive or negative valence (cf. Trillò et al., 2023). Finally, we coded the values associated with each evaluation using a slightly modified version of Shifman et al.'s (2025) comprehensive list of values.

The second phase featured an analysis of visual value cues (Shifman et al., 2025). Having familiarized ourselves with the videos, we developed a codebook for analyzing newsfluencers' attire (formal/informal), the background of their shows (recording studio/private space/green screen), references to external sources of information (presence/absence of clips from TV news, screenshots of print news, and social media posts), and the presence of expert guests (yes/no). The results of this coding were then combined to derive insights regarding how newsfluencers establish their value as news providers and the value of their show as news products.

After additional training, the three authors coded the first 10 minutes of the videos in their respective languages. At the end of the process, we collectively reviewed our coding in a series of consensus meetings (Hill et al., 1997) to reach agreement on all coding decisions.

Results

Newsfluencers repeatedly invoke the importance of truthfulness and rationality, and extensively dissect public figures' statements to fault them for not speaking the truth. Relatedly, newsfluencers imitate the codes of traditional expertise, using visuals to underscore the truthfulness of their own claims and to demonstrate skillfulness or authenticity. Most newsfluencers adopt a polished aesthetic that, while departing from the generic conventions of televised news, projects an aura of expertise (e.g., formal business wear) and professionalism (e.g., state-of-the-art microphones). In contrast, some newsfluencers embrace the aesthetics of participatory digital culture, displaying markers of authenticity (e.g., shooting from their bedroom) and amateurism (e.g., recording with their phone). Additionally, we identified newsfluencers' material dependence on mainstream media, with the former heavily relying on the latter as a source of factual information and visual footage.

We also found small but revealing cross-cultural differences tied to local communicative styles and global creator cultures. The opinionated style of American newsfluencers stood in stark contrast to the neutral style of Korean ones. However, newsfluencers from both countries converged towards the professional aesthetics described above, while Italian newsfluencers were the most likely to produce vlog-style shows reminiscent of YouTube's amateur era.

Reflecting on these findings, we argue that newsfluencers primarily diverge from traditional journalists in their orientation towards the truth: the job of journalists is to report the truth as matter of fact while the job of newsfluencers is to debate the truth as matter of concern. Thus, the rise of newsfluencers does not reflect a general disavowal of truthfulness, but its politicization as a pivotal value that must be upheld by subjecting all public statements to intense scrutiny. As all truth-claims are opened up for contestation, we conclude that the present moment is not characterized by a post-truth rationality, but by a post-objectivity one that subjects all truth-claims to debate and interpretation.

Future Work

This study has several limitations. The first is the distinct lack of gender diversity in our sample. While women are a substantial minority among newsfluencers, studies that focus on women newsfluencers may uncover a different set of value priorities. Our definition of newsfluencers is also a limitation, as we focus on influencers whose primary content output is about news. We encourage future studies that strategically sample the sporadic news commentary produced by creators who work within different genres (e.g., lifestyle creators). Finally, while we sought to create diversity within our sample by comparing North American, European, and Asian countries, our selection still represents countries with well-developed mass media systems, widespread online adoption, and diplomatic influence within the Global North. Studies that look to global majority countries, especially African countries like Kenya and Nigeria where the influence of news creators is particularly pronounced (Newman et al., 2025), would provide an informative comparison with our sample.

References

- Anderson, C. W. (2019). Journalism as procedure, journalism as values. *Journalism*, 20(1), 8–12. <https://doi.org/10.1177/1464884918806732>
- Bösch, M., & Divon, T. (2024). The sound of disinformation: TikTok, computational propaganda, and the invasion of Ukraine. *New Media & Society*, 26(9), 5081–5106. <https://doi.org/10.1177/14614448241251804>

- Deuze, M. (2005). What is journalism? Professional identity and ideology of journalists reconsidered. *Journalism*, 6(4), 442–464.
<https://doi.org/10.1177/1464884905056815>
- Hallinan, B. (2023). No judgment: Value optimization and the reinvention of reviewing on YouTube. *Journal of Computer-Mediated Communication*, 28(5), 1–14.
<https://doi.org/10.1093/jcmc/zmad034>
- Heinich, N. (2020). A pragmatic redefinition of value(s): Toward a general model of valuation. *Theory, Culture & Society*, 37(5), 75–94.
<https://doi.org/10.1177/0263276420915993>
- Hill, C. E., Thompson, B. J., & Williams, E. N. (1997). A guide to conducting consensual qualitative research. *The Counseling Psychologist*, 25(4), 517–572.
<https://doi.org/10.1177/0011000097254001>
- Hofstede, G. (1980). *Culture's consequences*. Sage.
- Hurcombe, E. (2024). Conceptualising the “newsfluencer”: Intersecting trajectories in online content creation and platformatised journalism. *Digital Journalism*. Online First. 1–12. <https://doi.org/10.1080/21670811.2024.2397088>
- Massey, B. L., & Chang, L. A. (2002). Locating Asian values in Asian journalism: A content analysis of web newspapers. *Journal of Communication*, 52(4), 987–1003.
<https://doi.org/10.1111/j.1460-2466.2002.tb02585.x>
- Newman, N., Ross Arguedas, A., Robertson, C. T., Nielsen, R. K., & Fletcher, R. (2025). Reuters Institute digital news report 2025. Reuters Institute for the Study of Journalism. <https://doi.org/10.60625/RISJ-8QQF-JT36>
- Pelevina, N., & Salojärvi, V. (2025). YouTube as a narrative battlefield: Brazilian social media influencers and the Russian war in Ukraine. *The Communication Review*, 1–24. <https://doi.org/10.1080/10714421.2025.2545676>
- Rieder, B. (2015). YouTube Data Tools (Version 1.42) [Software]. Available from <https://ytdt.digitalmethods.net>.
- Shifman, L., Trillò, T., Hallinan, B., Mizoroki, S., Green, A., Scharlach, R., & Frosh, P. (2025). The expression of values on social media: An analytical framework. *New Media & Society*, Online First. <https://doi.org/10.1177/14614448241307035>
- Trillò, T., Hallinan, B., Green, A., Kim, B., Mizoroki, S., Scharlach, R., Park, P. H., Frosh, P., & Shifman, L. (2023). “I love this photo, I can feel their hearts!” How users across the world evaluate social media portraiture. *Journal of Communication*, 73(3), 235–246. <https://doi.org/10.1093/joc/jqad009>

8B.3 Who watches whom: How YouTube channel networks organize migration discourse

William Hollingshead, Anatoliy Gruzd, Philip Mai (Toronto Metropolitan University)

Introduction

YouTube is the world's largest video-sharing platform, with an estimated 2.58 billion global monthly active users as of early 2025 (Dixon, 2025). In contemporary culture, YouTube has become a critical nodal point not only for creativity and self-expression, but also social change and political activism (Burgess & Green, 2018). Regarding the latter, YouTube is a salient space for investigating how contentious socio-political issues, like migration, are produced by different actors. Past work has explored YouTube channels' content to understand the circulation of far-right, xenophobic content (Laaksonen et al., 2020), and how migrant populations, including undocumented peoples (Bayramoğlu, 2023; Horsti, 2017) and spousal dependents (Mallapragada, 2017), contest stereotypical representations to assert political agency.

While important, this work has failed to explore the networked social interactions that undergird how migration discourse is structured on YouTube. As a point of departure, we examine the actor network of those producing migration-related content on YouTube. This is important, as community is constitutive of YouTube's culture (Burgess & Green, 2018). By "community," we are referring to a group of "entities" (in our case, channels) within the network that "interact with each other more frequently than those outside the group" (Bedi & Sharma, 2016, p. 116). YouTube channels' content is community-oriented in that it is shaped by patterned interaction between accounts, including channel collaborations, recommendations, and video responses or "call-outs" of other creators (Lewis, 2018; Lewis et al., 2021). Therefore, examining individual videos' content alone is insufficient; it requires studying channels' community membership to see who attends to whom, which informs how narratives are exchanged and contested.

Two of YouTube's key social affordances are useful for understanding channels' networked interactions in the arena of migration discourse: featured-channels and channel-subscriptions. Featured-channels allow one to "highlight," or amplify other channels on their profile, while channel-subscriptions indicate a desire to consume another creator's content (Rieder et al., 2020). While the former is automatically publicly accessible, the latter's visibility is optional. Importantly, each represents a mechanism for signaling to one's audience "pay attention here" (Khamis et al., 2017). These ties may reflect sociable practices, such as information-seeking or mutually held interests, or anti-sociable practices, such as surveillance, schadenfreude, and "hate-

following” (Ouwerkerk & Johnson, 2016). As such, these affordances are critical sources for mapping attention flows among migration-centred actors, potentially revealing their strategic orientation toward others in the production of migration discourse.

Objectives/Research Questions

In the current study, we explore YouTube channels’ ties vis-à-vis their featured-channels and channel-subscriptions to understand how actors’ signal association within the overall migration discourse network. We examine how these linking choices contribute to the formation of partitioned communities and may shape how migration narratives align or contest one another. In doing so, we offer insights into the potential reasoning underlying these actors’ linking behaviours, ranging from identity management (Cover, 2012) to oppositional surveillance (Ouwerkerk & Johnson, 2016). Accordingly, we ask:

RQ. What community structure emerges among YouTube channels that share migration-related content when considering both featured-channel and channel-subscription ties?

Methods

To address our research question, we identified YouTube channels contributing to migration discourse and analyzed their affiliative linking practices. Drawing on social network analysis, we examined featured-channel and channel-subscription connections as a proxy for community formation, revealing how actors’ connectivity structures migration discourse on the platform.

Channel Discovery

To first identify the channels actively discussing migration on YouTube, we followed the logic of query design (Rogers, 2017), using 62 terms across multiple “programs,” including policy-driven language, neutral terms, and politically-charged phrases. We then used YouTube Data Tools’ (YTDT) “Channel List” module (Rieder, 2023) to search for channels flagged by the YouTube API as relevant to each term. From this, we identified an initial sample of 18,820 channels.

However, manual inspection revealed that YouTube’s search mechanism introduced many channels not engaged with migration-related content. This included YouTube influencers, like @markiplier, who mainly streams videogames but engaged the term “border” by playing the Borderlands series. To identify channels that had recently shared content explicitly related to migration, we retrieved up to 50 of the most recent video uploads per channel. We then queried the titles and descriptions of these videos using the same corpus of terms, retaining only channels with at least one video upload

containing a migration-related term. Using this criterion, we identified 4,367 channels that served as our “seeds” for network discovery.

Network Discovery

We then input these seed channel IDs into YTDt’s “Channel Network” module. This allowed us to retrieve each seed channels’ featured-channels and construct an initial network where edges were formed based on channels’ features. This step added 827 channels ($n = 5,194$). While YTDt can also collect channel-subscriptions alongside featured-channels, it does not specify the type of link made. This difference is important to capture, as featured-channels and channel-subscriptions reflect differing forms of interaction (Rieder et al., 2020). To resolve this, we developed a custom Python script that retrieved each channel’s channel-subscriptions using YouTube’s API, and generated an edge if a channel subscribed to any of the 5,194 channels in our list. We then consolidated each of the featured-channel and channel-subscription edges into one network file. This enabled us to distinguish how channels’ links were formed

Data Analysis

We analyzed the network’s structure using Gephi (Bastian et al., 2009). As our analytical focus is on channels’ relational structure, we removed 3,678 channels without incoming or outgoing connections. To understand the network’s community structure, we used the Fast Unfolding community detection algorithm to identify channels that are more likely to be connected to each other than the rest of the network (Blondel et al., 2008).

Results

Community detection yielded a high modularity statistic (0.847), partitioning the network into 114 communities. In this exploratory study, we focus on the six largest communities containing greater than 5% of the network’s channels, accounting for 63.1% of the total network (Table 2). These six communities encompassed media outlets and figures, consumer brands and celebrities, immigration-focused actors, and humanitarian groups (Figure 1).

These communities appear to be generated by homophily, similar to past work exploring online publics’ interactions surrounding migration-related topics (de Rosa et al., 2021; Siapera et al., 2018). By homophily (McPherson et al., 2001), we mean the principle that “...contact between similar people occurs at a higher rate than dissimilar people” (p. 416). This homophilous structure suggests that channels’ clustering occurred around shared political and/or organizational lines. We confirmed this by manually reviewing the public profiles of channels within each of the six largest communities. Indeed, Communities A, B, and D contain media-centric channels, where migration coverage often skews toward negativity and sensationalism (Eberl et al.,

2018), whereas Communities E and F featured immigration specialist channels that focus on the bureaucratic “infrastructures” of migration, such as visa consultation (Xiang & Lindquist, 2014).

However, this homophilous clustering did not produce complete community isolation. We found evidence of cross-community ties in the network’s highly-partisan core (Communities A, B, and D) between “political influencer” channels, or those whose content’s focus is endorsing a “political position, social cause, or candidate” (Riedl et al., 2023, p. 2), suggesting that attention to ideological opponents is a constitutive component of the network’s community structure.

This is further exemplified by the fact that cross-partisan activity was largely generated through channels’ use of subscription- rather than feature-based ties. We highlight this, as channel-subscriptions’ ties express ambiguous “interest” in another’s content, whereas featured-channels indicate explicit “endorsement” (Rieder et al., 2020). These patterns suggest that channels’ subscriptions are an important mechanism for tracking opposing narratives within a highly fragmented network. More specifically, this configuration revealed that inter-community connection served as an infrastructure for strategic monitoring, enhancing a creator’s capacity to rapidly refute their opponent’s narratives. This aligns with past research on political influencers’ communication strategies (Fischer et al., 2022), where content is often centred on responding to enemies in a mocking tone. Therefore, rather than facilitating deliberative exchange, cross-community attention therefore fosters a competitive discursive environment in which migration narratives are highly contested.

Future Work

Conclusion and Implications

In this exploratory study, we examined how YouTube channels’ featured-channels and channel-subscriptions function as a mechanism for creating a digital infrastructure to frame migration discourse on the platform. Specifically, we found that migration discourse on YouTube is structured into largely homophilous, actor-specific communities, where selective cross-community ties—particularly via channel-subscriptions—is used for monitoring and contestation, rather than facilitating deliberative exchange.

These dynamics can be interpreted through YouTube channels’ strategic management of their public identities. YouTube’s featured-channel and channel-subscription lists are publicly visible markers that detail how one positions themselves in relation to others on the platform. For instance, we confirmed that featured-channels’ “endorsement” functionality operated as a proxy for signalling one’s political and professional identities, echoing past work on Facebook friendship (Bond & Messing, 2015) and Twitter follower networks (Pan et al., 2019). In contrast, channel-

subscriptions’ “monitoring” enables observing another creator’s content without explicit endorsement, which may be beneficial to those seeking to acquire adversarial information as fodder for their own content (Fischer et al., 2022). These two affordances therefore underscore an important circuitry that shapes channels’ abilities for affiliation and attention in YouTube’s migration discourse network.

Future Research

To clarify the forces shaping how migration-focused communities form on YouTube, future work should use network modelling approaches, such as exponential random graph models (ERGMs). These can be used to test whether structural or attribute-based patterns of tie formation occur more or less than expected by chance (Robins & Lusher, 2013).

References

- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. *Proceedings of the International AAAI Conference on Web and Social Media*, 3(1), 361–362.
<https://doi.org/10.1609/icwsm.v3i1.13937>
- Bayramoğlu, Y. (2023). Border countervisuality: Smartphone videos of border crossing and migration. *Media, Culture & Society*, 45(3), 595–611.
<https://doi.org/10.1177/01634437221136013>
- Bedi, P., & Sharma, C. (2016). Community detection in social networks. *WIREs Data Mining and Knowledge Discovery*, 6(3), 115–135.
<https://doi.org/10.1002/widm.1178>
- Blondel, V., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Bond, R., & Messing, S. (2015). Quantifying social media’s political space: Estimating ideology from publicly revealed preferences on facebook. *American Political Science Review*, 109(1), 62–78. <https://doi.org/10.1017/S0003055414000525>
- Burgess, J., & Green, J. (2018). *YouTube: Online Video and Participatory Culture* (Second Edition). Polity Press.
- Cover, R. (2012). Performing and undoing identity online: Social networking, identity theories and the incompatibility of online profiles and friendship regimes. *Convergence*, 18(2), 177–193. <https://doi.org/10.1177/1354856511433684>

- de Rosa, A. S., Bocci, E., Bonito, M., & Salvati, M. (2021). Twitter as social media arena for polarised social representations about the (im)migration: The controversial discourse in the Italian and international political frame. *Migration Studies*, 9(3), 1167–1194. <https://doi.org/10.1093/migration/mnab001>
- Dixon, S. J. (2025, October 6). Most popular social networks worldwide as of February 2025, by number of monthly active users. Statista. <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/?srltid=AfmBOOp4495fl7X-gf-1pv7ROGzGZkE9CORyopswZfDzZU-MYsz5ptCE>
- Eberl, J.-M., Meltzer, C. E., Heidenreich, T., Herrero, B., Theorin, N., Lind, F., Berganza, R., Boomgaarden, H. G., Schemer, C., & Strömbäck, J. (2018). The European media discourse on immigration and its effects: A literature review. *Annals of the International Communication Association*, 42(3), 207–223. <https://doi.org/10.1080/23808985.2018.1497452>
- Fischer, T.-S., Kolo, C., & Mothes, C. (2022). Political influencers on YouTube: Business strategies and content characteristics. *Media and Communication*, 10(1), 259–271. <https://doi.org/10.17645/mac.v10i1.4767>
- Horsti, K. (2017). Communicative memory of irregular migration: The re-circulation of news images on YouTube. *Memory Studies*, 10(2), 112–129. <https://doi.org/10.1177/1750698016640614>
- Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS ONE*, 9(6), e98679. <https://doi.org/10.1371/journal.pone.0098679>
- Khamis, S., Ang, L., & Welling, R. (2017). Self-branding, ‘micro-celebrity’ and the rise of Social Media Influencers. *Celebrity Studies*, 8(2), 191–208. <https://doi.org/10.1080/19392397.2016.1218292>
- Laaksonen, S.-M., Pantti, M., & Titley, G. (2020). Broadcasting the movement and branding political microcelebrities: Finnish anti-immigration video practices on YouTube. *Journal of Communication*, 70(2), 171–194. <https://doi.org/10.1093/joc/jqz051>
- Lewis, R. (2018). Alternative influence: Broadcasting the reactionary right on YouTube. *Data & Society*. https://datasociety.net/wp-content/uploads/2018/09/DS_Alternative_Influence.pdf
- Lewis, R., Marwick, A. E., & Partin, W. C. (2021). “We dissect stupidity and respond to it”: Response videos and networked harassment on YouTube. *American Behavioral Scientist*, 65(5), 735–756. <https://doi.org/10.1177/0002764221989781>

- Mallapragada, M. (2017). Immigrant activism: Narratives of the “H-4 Life” by Indian women on YouTube: Immigrant activism. *Communication, Culture & Critique*, 10(1), 76–92. <https://doi.org/10.1111/cccr.12147>
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 415–444. <https://doi.org/10.1146/annurev.soc.27.1.415>
- Ouwerkerk, J. W., & Johnson, B. K. (2016). Motives for online friending and following: The dark side of social network site connections. *Social Media + Society*, 2(3), 2056305116664219. <https://doi.org/10.1177/2056305116664219>
- Pan, J., Bhardwaj, R., Lu, W., Chieu, H. L., Pan, X., & Puay, N. Y. (2019). Twitter homophily: Network based prediction of user’s occupation. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 2633–2638. <https://doi.org/10.18653/v1/P19-1252>
- Rieder, B. (2023). YouTube Data Tools (Version 1.42) [Computer software]. <https://ytdt.digitalmethods.net>
- Rieder, B., Coromina, Ò., & Matamoros-Fernández, A. (2020). Mapping YouTube: A quantitative exploration of a platformed ecosystem. *First Monday*. <https://doi.org/10.5210/fm.v25i8.10667>
- Riedl, M. J., Lukito, J., & Woolley, S. C. (2023). Political influencers on social media: An introduction. *Social Media + Society*, 9(2), 20563051231177938. <https://doi.org/10.1177/20563051231177938>
- Robins, G., & Lusher, D. (2013). What are exponential random graph models? In D. Lusher, G. Robins, & J. Koskinen (Eds.), *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications* (pp. 9–15). Cambridge University Press.
- Rogers, R. (2017). Foundations of digital methods: Query design. In M. T. Schäfer & K. van Es (Eds.), *The Datafied Society: Studying Culture Through Data* (pp. 75–94). Amsterdam University Press.
- Siapera, E., Boudourides, M., Lenis, S., & Suiter, J. (2018). Refugees and network publics on Twitter: Networked framing, affect, and capture. *Social Media + Society*, 4(1), 2056305118764437. <https://doi.org/10.1177/2056305118764437>
- Xiang, B., & Lindquist, J. (2014). Migration infrastructure. *International Migration Review*, 48(1), 122–148. <https://doi.org/10.1111/imre.12141>

8B.4 Down the Content Rabbit Hole: Making Algorithmic Content Shift on Short-Form Video Platforms Measurable

Tobias Kietreiber, Felix Krejca, Anna Beubacher, Sebastian Neumaier, Lukas Daniel Klausner (University of Applied Sciences St Pölten)

Introduction

Short-form video platforms like Instagram Reels, TikTok and YouTube Shorts rely on recommendation algorithms that adapt dynamically to a user's interaction history. Concerns grow steadily regarding the societal impacts of such systems – including polarisation, exposure to harmful content and the reinforcement of existing biases [2, 16] – however, researchers still lack reproducible tools to systematically study how specific interaction patterns affect content feeds, particularly on mobile-first, closed-source platforms. As these platforms have replaced chronological feeds with algorithmically curated timelines optimised for watch time and engagement [3, 8], their algorithmic opacity complicates both user understanding and scientific analysis of these changes and their implications: Content that is popular, controversial or visually engaging is amplified, increasing user engagement while limiting exposure to diverse viewpoints [14, 15, 17]. Users often develop informal "folk theories" to explain content selection [9, 10], while audits suggest that personalisation may amplify extremist content or create filter bubbles – although this phenomenon is more nuanced and context-dependent than is commonly assumed [1, 11, 23].

To overcome the algorithmic opacity of these content recommender systems, many studies use "sock puppet" audits. These usually use browser-based automation or API-driven scripts [4, 12, 13, 21, 22], often combined with fixed textual heuristics like hashtags or metadata [7, 18, 19] for classification to automatically measure content drift. In contrast, our approach utilises an Android emulator to mirror the native mobile environment more accurately. We present a proof-of-concept framework to measure algorithmic content shifts on YouTube Shorts. Our primary objective is to determine the technical feasibility of combining real-time content classification with automated mobile app interaction to support systematic measurements of recommendation behaviour.

Objectives/Research Questions

Our long-term objective is to understand how content consumption on video-based social media platforms influences subsequent algorithmic changes in a user's feed – motivating the prerequisite research question of this work:

“Is it technically feasible to combine real-time content classification with automated

mobile app interaction to support systematic measurements of recommendation algorithm behaviour on short-form video platforms?”

Our contributions:

- (1) We demonstrate that CLIP-based zero-shot classification enables real-time labeling in an automated loop.
- (2) We validate this classifier against a manually annotated video dataset scraped directly from the platform.
- (3) We show that this setup is a robust basis for future large-scale studies.

As this is methodological work, our goal is to document a transparent and extensible measurement setup rather than making immediate causal claims about platform algorithms.

Methods

Our proof-of-concept system has two primary components: an Android emulator for social media interaction and a CLIP-based video classifier that steers the content engagement logic (see Figure in Supplementary Material).

== Video Classification with CLIP ==

We utilize Contrastive Language-Image Pre-Training (CLIP) [20] for zero-shot classification of video frames. CLIP is a multimodal model providing a shared embedding space for images and text trained on a massive dataset of image-caption pairs using contrastive learning. By comparing frame embeddings to natural language descriptions (“prompts”), the system assigns categories based on cosine similarity. We use the ViT-B/32 variant implemented via JAX [5]. This is approximately ten times faster than the original version, performing classifications in under 2 ms on an Nvidia GeForce RTX 3090.

The classification workflow precomputes text embeddings for positive (target class) and negative (generic scene) prompts. Each frame is split into top, centre and bottom quadratic crops (224×224 pixels), which are normalised and fed into CLIP. To ensure robustness, we aggregate the top k similarities and apply a softmax function to generate a probability metric. To reduce sensitivity to outliers, a video frame is only assigned a class if at least eight of the last ten frames exceed a probability threshold. A video is categorised if more than a defined minimum number of frames meet these criteria.

Prompt design involved generating approximately 50 positive descriptions per class and 50 generic prompts describing the negative class using Gemini 3, followed by manual

refinement to address specific "blind spots" (such as identifying food types that the initial model missed).

== Emulator Interaction ==

To interact with the platform, we run an Android emulator on a server, streaming video using scrcpy-server. Automated commands (swipes and clicks) are sent via the Android Debug Bridge. The interaction logic samples watch times from a skewed normal distribution to mimic human behaviour [6] (see Algorithm in Supplementary Material). If the watch time is reached and the content does not meet the frame count threshold for the desired class, the system skips the video via a swipe. If the content is relevant, the watch time is extended to simulate high engagement.

Results

In our preliminary results, we focus on the evaluation of the CLIP-based classifier. We consider the measurement of algorithmic content shift to be out of scope of this methodological paper; we will address these dynamics in future research.

== Data Collection and Annotation ==

To validate our system, we employed a modular two-stage scraper. First, it retrieves video URLs and subsequently downloads the content (ensuring reusability for other platforms, e. g. TikTok). For this study, we used the YouTube API to retrieve 331 popular English-language "Shorts" uploaded between 2025-09-01 and 2025-12-31. This dataset represents the "default" recommendation distribution for a logged-out user. For future studies requiring highly personalised feeds, our framework can instead retrieve links via any platform's "share" functionality.

The dataset was annotated by the authors through a custom Flask-based web application. While agreement-based labeling by multiple annotators is the gold standard, we utilised a single-annotation approach with random verification, establishing a high-quality performance baseline for this proof of concept. We focused on three simple categories ("Food", "Car" and "Animal") to validate the classifier's technical reliability.

== Classifier Performance ==

The CLIP-based classifier achieved high levels of accuracy and F_1 scores across all tested classes – e. g. reaching accuracies exceeding 0.91 for all categories. Specifically, the "Food" class demonstrated strong reliability with a precision of 0.90, while the classes "Car" and "Animal" also achieved acceptable levels of recall and precision (see Table in Supplementary Material).

Direct comparisons are difficult, since to our knowledge there is no clearly analogous research. [22] seems to be most comparable; they reach F_1 scores between 0.51 and

0.93 and accuracies between 0.78 and 0.93; however, comparability remains limited, as their work covers the complex task of conspiracy theory detection, using only textual features, and on videos without a restriction on “Shorts”.

Manual inspection of false positives and false negatives provided insight into the classifier's behavior. Many misclassifications in the “Animal” class were “borderline” cases that also challenged the human annotators (such as animated wolf-like monsters). Other errors involved missing general classes; for instance, a close-up of a metal object was occasionally misidentified as car chassis. While we could have refined the prompts to eliminate these specific errors, we chose to avoid “overfitting” the classifier to this specific dataset. Instead, we intend to expand the general class prompts in future work featuring larger and more diverse evaluation sets. Overall, these results confirm that CLIP-based zero-shot classification is a viable and robust method for real-time interaction control in short-form video audits.

Future Work

While our current prototype demonstrates the technical feasibility of pairing CLIP-based zero-shot video classification with emulated interactions, several avenues remain for scaling this framework into a comprehensive auditing tool.

== Enhancing Data and Automation ==

Larger and more diverse datasets are essential for further refining the classifier. We plan to implement active learning methods to prioritise informative or ambiguous samples, reducing manual annotation effort while improving model performance on edge cases. To support sustained, long-term measurement studies, we are working on fully automating the emulator setup, including device initialisation, account configuration and scripted interaction routines. This will allow us to distinguish between short-term, within-session adaptation and the longer-term user profiling dynamics that define platform personalisation.

== Multimodal Analysis and Nuanced Schemas ==

A significant step towards more meaningful analysis of platform algorithms will be the transition from visual-only classification to multimodal models. By incorporating audio streams, on-screen text and metadata (such as titles and hashtags), the system will achieve higher accuracy and a more holistic understanding of the content. This will enable us to move beyond simple categories like “Animal” vs. “Food” toward theoretically grounded taxonomies. We aim to identify and track more complex social phenomena such as hate speech and extremist content and resulting differences in algorithmic recommendations between such categories.

== Cross-Platform and Contextual Audits ==

Our framework's modular design allows easy manipulation of contextual variables such as geographic location (via VPN/GPS spoofing), language and device configuration, enabling monitored investigations into how these factors influence recommendation dynamics. Furthermore, because the interaction logic is based on the native Android environment rather than platform-specific APIs, our system is easily adaptable to other short-form video platforms like Instagram Reels or TikTok.

References

- [1] Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Fallin Hunzaker, M. B., ... Volfovsky, A. (2018). Exposure to Opposing Views on Social Media Can Increase Political Polarization. *Proc. Natl. Acad. Sci. U.S.A.*, 115 (37), 9216–9221. doi: 10.1073/PNAS.1804840115
- [2] Bandy, J. (2021). Problematic Machine Behavior: A Systematic Literature Review of Algorithm Audits. *Proc. ACM Hum.-Comput. Interact.*, 5 (CSCW1). doi: 10.1145/3449148
- [3] Bergen, M. (2022-09-05). Just a Few People Crowned Some of YouTube's Earliest Hits. *The Atlantic*. Retrieved from <https://www.theatlantic.com/technology/archive/2022/09/youtube-homepage-editor-google-algorithm-book-excerpt/671339/>.
- [4] Boeker, M., & Urman, A. (2022). An Empirical Investigation of Personalization Factors on TikTok. In *Proceedings of the ACM Web Conference 2022* (pp. 2298–2309). New York, NY: ACM. doi: 10.1145/3485447.3512102
- [5] Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., ... Zhang, Q. (2018). JAX: Composable Transformations of Python+NumPy Programs. Retrieved from <http://github.com/jax-ml/jax>.
- [6] Chen, Y., Zhang, B., Liu, Y., & Zhu, W. (2013). Measurement and Modeling of Video Watching Time in a Large-Scale Internet Video-on-Demand System. *IEEE Trans. Multimed.*, 15 (8), 2087–2098. doi: 10.1109/TMM.2013.2280123
- [7] Dagtas, S., Cakmak, M. C., & Agarwal, N. (2025). Simulating User Watch-Time to Investigate Bias in YouTube Shorts Recommendations. doi: 10.48550/arXiv.2507.04534
- [8] Davidson, J., Liebald, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U., ... Sampath, D. (2010). The YouTube Video Recommendation System. In *Proceedings of the Fourth ACM Conference on Recommender Systems* (pp. 293–296). New York, NY: ACM. doi: 10.1145/1864708.1864770
- [9] DeVito, M. A., Gergle, D., & Birnholtz, J. (2017). “Algorithms Ruin Everything”: #RIPTwitter, Folk Theories, and Resistance to Algorithmic Change in Social Media. In

Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (pp. 3163–3174). New York, NY: ACM. doi: 10.1145/3025453.3025659

[10] Eslami, M., Karahalios, K., Sandvig, C., Vaccaro, K., Rickman, A., Hamilton, K., & Kirlik, A. (2016). First I “Like” It, Then I Hide It: Folk Theories of Social Feeds. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 2371–2382). New York, NY: ACM. doi: 10.1145/2858036.2858494

[11] Gentzkow, M., & Shapiro, J. M. (2011). Ideological Segregation Online and Offline. *Q. J. Econ.*, 126 (4), 1799–1839. doi: 10.1093/QJE/QJR044

[12] Horta Ribeiro, M., Ottoni, R., West, R., Almeida, V. A. F., & Meira, W., Jr. (2020). Auditing Radicalization Pathways on YouTube. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 131–141). New York, NY: ACM. doi: 10.1145/3351095.3372879

[13] Hussein, E., Juneja, P., & Mitra, T. (2020). Measuring Misinformation in Video Search Platforms: An Audit Study on YouTube. *Proc. ACM Hum.-Comput. Interact.*, 4 (CSCW1). doi: 10.1145/3392854

[14] Klug, D., Qin, Y., Evans, M., & Kaufman, G. (2021). Trick and Please: A Mixed-Method Study on User Assumptions About the TikTok Algorithm. In Proceedings of the 13th ACM Web Science Conference (pp. 84–92). New York, NY, USA: ACM. doi: 10.1145/3447535.3462512

[15] Kuncoro, H. R., Hasanah, K., Sari, D. L., & Kurniawati, E. (2023). Algorithmic Influence: Twitter’s Role in Shaping Public Discourse and Amplifying Radical Content. In Proceedings of the 2nd International Conference on Advance Research in Social and Economic Science (pp. 549–553). Paris: Atlantis Press. doi: 10.2991/978-2-38476-247-75 5

[16] Lambrecht, A., & Tucker, C. (2019). Algorithmic Bias? An Empirical Study of Apparent Gender-Based Discrimination in the Display of STEM Career Ads. *Manag. Sci.*, 65 (7), 2966–2981. doi: 10.1287/MNSC.2018.3093

[17] Madrigal, A. C. (2019-02-21). The Reason Conspiracy Videos Work So Well on YouTube. *The Atlantic*. Retrieved from <https://www.theatlantic.com/technology/archive/2019/02/reason-conspiracy-videos-work-so-well-youtube/583282/>.

[18] Masood, M., Kannan, S., Liu, Z., Vasisht, D., & Gupta, I. (2025). Counting How the Seconds Count: Understanding Algorithm-User Interplay in TikTok via ML-driven Analysis of Video Content. doi: 10.48550/arXiv.2503.20030

[19] Mosnar, M., Skurla, A., Pecher, B., Tibensky, M., Jakubcik, J., Bindas, A., ... Srba, I. (2025). Revisiting Algorithmic Audits of TikTok: Poor Reproducibility and Short-Term

Validity of Findings. In Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 3357–3366). New York, NY: ACM. doi: 10.1145/3726302.3730293

[20] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... Sutskever, I. (2021). Learning Transferable Visual Models from Natural Language Supervision. In Proceedings of the 38th International Conference on Machine Learning (pp. 8748–8763). Cambridge, MA: JMLR and Microtome Publishing. Retrieved from <https://proceedings.mlr.press/v139/radford21a.html>.

[21] Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014). Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms. In Data and Discrimination: Converting Critical Concerns into Productive Inquiry – Preconference at the 64th Annual Meeting of the International Communication Association. Retrieved from <https://websites.umich.edu/~csandvig/research/Auditing%20Algorithms%20--%20Sandvig%20--%20ICA%202014%20Data%20and%20Discrimination%20Preconference.pdf>.

[22] Srba, I., Moro, R., Tomlein, M., Pecher, B., Simko, J., Stefancova, E., ... Bielikova, M. (2023). Auditing YouTube’s Recommendation Algorithm for Misinformation Filter Bubbles. ACM Trans. Recomm. Syst., 1 (1). doi: 10.1145/3568392

[23] Whittaker, J., Looney, S., Reed, A., & Votta, F. (2021). Recommender Systems and the Amplification of Extremist Content. Internet Policy Rev., 10 (2). doi: 10.14763/2021.2.1565

Session 8C - Politics, Policy and Regulation (1)

Day 2 · Tues 14 July · 13:50-15:20, Seminar Suite B, ARC

8C.1 Aesthetic Authoritarianism? Visual Propaganda and the Politics of Urban Renewal in Ethiopia

Amanuel Tesfaye Kebede (Helsinki University)

Introduction

Images and short-form videos have emerged as the predominant features of social media. This has especially been the case since the launch of TikTok in 2018, quickly becoming one of the largest video-sharing platforms globally. The proliferation of visual content aligns with popular preferences as visual media tends to be ‘attention-grabbing, emotion eliciting, and easy-to-digest’ which makes it versatile for a digital age with information overload (Seo, 2020). Authoritarian states have recognized this visual turn in social media and have adapted their propaganda efforts, with state and pro-state accounts engaging in campaigns utilizing sharing short-form videos in order to shape public perceptions to their advantage. This has been amplified by the advent of Generative AI, which has facilitated easy production of multimodal media, resulting in the proliferation of ‘synthetic propaganda’ (Bosch 2025).

While this growing use of visual and synthetic propaganda is now starting to capture scholarly attention (Swimelar, 2018; van Noort, 2019; Seo, 2020; Rastrilla et al, 2023; Bösch & Divon, 2024), its use in other comparative contexts still remains largely underexplored. This study addresses this gap by investigating how visual propaganda is being utilized by the Ethiopian government and pro-government actors in order to shape public perceptions, mobilize supporters, and push back against critics. In order to ground this empirically, the study undertake a case study of the Ethiopian government’s visual communication on TikTok and Twitter regarding the ‘corridor project’ the government’s flagship urban renewal projects throughout the country.

What has colloquially come to be known as ‘corridor project’ in Ethiopia includes urban renewal initiatives to upgrade infrastructure, improve aesthetics, reduce traffic, and improve mobility. The most visible of which has been the Addis Ababa Corridor Development Project, which among others include wider walkways, bike lanes, green spaces, and vibrant streetlights, as well as an associated riverside development. The corridor project has become the centerpiece of the government’s agenda, consuming a large amount of the state’s resources. Such urban renewal efforts have faced criticism

regarding forceful eviction of the urban poor, erasure of cultural heritage, and lack of transparency. It is also criticized for misplaced prioritization and allocation of government resources considering the socio-economic and political challenges Ethiopia faces. Large swaths of the country are experiencing civil wars, with 4.5 million people internally displaced, and 10 million requiring humanitarian aid in 2025. In the face of such systemic crisis, urban renewal has been painted as a luxury by critics.

Objectives/Research Questions

To push back against such criticism, the Ethiopian government and supporters have resorted to visual propaganda, sharing visually striking transformations of urban spaces as proof that the government is succeeding despite naysayers. Government accounts and pro-government actors on social media, including the official accounts of the Prime Minister and the Mayor of the Capital, constantly share videos and pictures of green spaces, spacious streets covered in neon lights and high rises, often set to patriotic and triumphant music or speeches. These have also been amplified by influencers, including foreign influencers who are provided privileged access to record and disseminate visuals of urban renewal efforts. Given the growing significance of the visual turn of social media propaganda, this study asks the following three interrelated questions:

1. How do pro-government accounts narrate the urban corridor project through visual propaganda on social media?
2. What dominant themes, discursive strategies, and visual tropes emerge from these visual imageries?
3. What narratives and underlying messages emerge from these visual propaganda materials?

Methods

The research approach involves two phases. The first phase involves an immersive observation of the two selected platforms, TikTok and Twitter, to identify and curate a small dataset of short form videos from government and pro-government accounts focusing on the corridor project. This observation is designed to understanding the breadth of the pro-government ecosystem in the two platforms as well as identifying the key accounts. Based on this observation, a small dataset of 40 to 50 short-form videos explicitly showcasing the corridor project will be curated, collecting all relevant metadata including captions, hashtags, sound, and other relevant account information.

Once a full dataset is curated, the study will employ to multimodal discourse analysis (Machin, 2013) in order to understand the themes, discursive strategies, visual tropes, and narratives that underlie the short-form videos of the corridor project. Multimodal discourse analysis is chosen since such short-form videos are inherently complex

multimodal objects deploying image, sound, and text in combination to communicate a political message. The study situates the analysis in the broader political context of fragility and violent contestation that has become a key feature of Ethiopian politics in recent years, since the propaganda efforts cannot be divorced from the insecurity and challenges rulers face.

Results

Based on a preliminary analysis of the small dataset (data collection ongoing), the study argues that these visuals regarding the corridor project constitute an aesthetic turn in propaganda in an authoritarian context, where state actors resort to aesthetically appealing and patriotic visuals of urban transformation in order to deter criticism of their records and mobilize their support base. The narratives within such visual propaganda are aimed at initiating an affective reaction within target audiences by generating feelings of pride at the visible transformation, and by extension attachment and support for the ruling elite. The visuals of a radically transformed urban environment are implicitly taken as an indicator of the accomplishments of the ruling elite in bringing about socio-economic development. This challenges the actual record of economic stagnation and decline, a record the ruling party dismisses by pointing to urban visuals as evidence of its success. Ultimately, visual propaganda, which is increasingly automated and AI-generated, has become another tool of authoritarian rule, serving to legitimize the government's hold on power.

By framing these short-form videos on social media regarding the corridor project in Ethiopia as an expression of evolving visual propaganda, the study contributes to our understanding of the changing relationship between visual politics, propaganda, and authoritarian legitimacy in the social media age. Broadly, it contributes to our understanding of fast-changing digital authoritarian practices.

Future Work

This is a first phase in an effort to extensively examine visual propaganda and information manipulation in an authoritarian context. Future work in this context includes the following:

- The immediate priority is to complete the data collection process, which is currently ongoing, and analyze it in line with the research approach outlined above.
- I also have a small dataset of AI-generated images and videos from state-aligned actors in Ethiopia focusing on Ethiopia's quest to secure access to a port in the Red Sea.

This will be the next project following the corridor project focused work above. These works are complementary – while the corridor project focused propaganda principally targets domestic audiences, the Red-Sea focused propaganda focuses on both

domestic and international audiences. It thus has a foreign policy aspect that can be linked to the concept of strategic narratives and visibility in international relations.

- Together the two projects will give us a comprehensive picture of the deployment of visuals as part of political propaganda on social media, contributing to our understanding of digital authoritarianism and authoritarian adaptation to technological change.

References

Bosch (2025). Synthetic Propaganda. <https://www.marcus-boesch.de/synthetic-propaganda/>

Bösch, M., & Divon, T. (2024). The sound of disinformation: TikTok, computational propaganda, and the invasion of Ukraine. *new media & society*, 26(9), 5081-5106.

Eder, J., & Klonk, C. (Eds.). (2017). *Image operations: Visual media and political conflict* (1st ed.). Manchester University Press. <http://www.jstor.org/stable/j.ctvnb7m30>

Freistein, K., Gadinger, F., & Groth, S. (2025). Studying narratives in international relations. *International Studies Perspectives*, 26(4), 434-455.

Machin, D. (2013). What is multimodal critical discourse studies?. *Critical discourse studies*, 10(4), 347-355.

Rastrilla, L. P., Sapag, P., & García, A. R. (Eds.). (2023). *Fast politics: Propaganda in the age of TikTok*. Springer Nature.

Seo, H. (2020). Visual propaganda and social media. *Handbook of Propaganda*, 126-137.

Swimelar, S. (2018). Deploying images of enemy bodies: US image warfare and strategic narratives. *Media, War & Conflict*, 11(2), 179-203.

Van Noort, Carolijn. "Strategic narratives, visibility and infrastructure in the digital age: the case of China's maritime silk road initiative." *Cambridge Review of International Affairs* 33.5 (2020): 734-751.

8C.2 "Flooding the Tropical Zone": Understanding Algorithmic Amplification on Instagram During Election Campaigns

Bruna Paroni (University of Urbino)

Introduction

Digital platforms are now central for political campaigns, with visual-centric spaces gaining significant traction (Gerodimos, 2019; Farkas, 2023). In these environments, visibility is algorithmically produced, favoring engagement over moderation and reactions over deliberation (Gillespie, 2014; Stromer-Galley et al., 2015). On Instagram, this reinforces political personalization, where private celebrations and emotional appeals outperform policy-driven discourse (Highfield & Leaver, 2016; Bossetta & Schmøkel, 2023). In Brazil, political actors effectively leverage this via celebratory, parody-based self-imaging (De-Lima-Santos et al., 2024).

These platforms often trigger disruptive adversarial patterns (Yarchi et al., 2021) and "persistent" toxicity (Avalle et al., 2024), reinforcing in- /out-group dynamics and affective polarization (Iyengar et al., 2019). Yet, while polarization scholarship is growing, research often neglects platform architecture (Bossetta, 2024), also due to the systemic barriers to data access. As Esau et al. (2024) argue, exacerbated attention toward extreme voices distorts societal perceptions. This "unequal distribution of voice" creates a cycle where extreme rhetoric is incentivized for visibility. However, a gap remains in understanding how algorithmic structures specifically mediate such visibility, leaving a conceptual void regarding the technical triggers of this "exacerbated presence" (Bruns et al., 2024).

By synthesizing an interdisciplinary framework, this work demonstrates how visibility, platform governance, and destructive polarization are interconnected. We analyze the 2024 São Paulo mayoral race - involving 9 million voters - where a candidate with a lifestyle influencer background utilized a decentralized network to "flood the zone" with emotional-driven content. A critical technical gap emerged when the Electoral Court ordered his account's suspension for creating an "image that lacks respect for the equilibrium an electoral race should have" (TSE-SP, 2024). However, Meta implemented this order merely as a geoblock rather than a platform-wide suspension. This allowed the candidate to circumvent the restriction and reach his audience via VPN on the original account, while simultaneously deploying a backup account to maintain his algorithmic presence. This case illustrates how technical implementations fail to align with democratic safeguards, ultimately reinforcing the distorted perceptions of polarization described by Esau et al. (2024).

Objectives/Research Questions

As Trilling et al. (2022) argue, research frequently neglects the platform context in which political news consumption occurs. In this sense, recommendation algorithms transform traditional social network logic (Giglietto, 2025) by prioritizing content from accounts users do not follow (Narayanan, 2023). This architecture produces public spaces tied to the economic incentives of platforms that prioritize engagement to retain attention, and more importantly, help to build our (political) knowledge (Gillespie, 2018; Narayanan, 2023). Consequently, these algorithms can act as targets for manipulation (Terenzi, 2024), enabling the legitimization of symbolic violence (Recuero, 2024) and "dark participation" (Quandt & Klapproth, 2023) when such content generates high engagement or exacerbated space (Esau et al., 2024).

These structural incentives create a fertile ground for the platform-mediated polarization described in our background. Following this logic, we investigate the shift from content curation (user-driven) to recommendation-based logic (system-driven) during elections. While studies on Twitter/X and Facebook, for instance, have analyzed manipulation campaigns (Marwick & Lewis, 2017) and institutionalized disinformation (Albuquerque & Alves, 2023), they often overlook the specific mechanisms by which platform architecture - specifically on Instagram - amplifies content.

Because viewing is a "precondition for re-sharing" (Trilling et al., 2022, p. 4) and a prerequisite for all subsequent interaction, we prioritize "views" as our primary metric for political visibility. Unlike traditional engagement (likes/comments), views capture the algorithmic reach that precedes user action. We do not contend that "views" are the sole determinant of engagement or distribution. Rather, we are operationalizing the metric to examine what increases visibility within the platform's algorithm during elections. By isolating this specific metric, we can better understand the baseline exposure provided by platform architecture "before" individual user agency takes over. To this end, we address the following research questions:

RQ1: What differences in views are found among the mayoral candidates in relation to left and right actors?

RQ2: What characterizes the posts from mayoral actors gaining most views and engagement?

Methods

To address our research questions, we utilize a dataset of 7,516 Instagram posts collected via the Meta Content Library (MCL) User Interface. The data covers the entirety of the official 2024 mayoral campaign in São Paulo (August 1st to October 31st). We employed a purposive sampling strategy to select the five leading candidates

- defined as those who were active participants in the official televised debates - ensuring the analysis focuses on the primary drivers of public discourse. This enables us to collect data despite MCL limitations

To analyze the factors influencing visibility (RQ1), we employed a negative binomial regression model. This approach is specifically suited for overdispersed count data, such as social media data. The model estimates the predictors of visibility while controlling for audience engagement (likes, comments), candidate political affiliation (left/center-left, right/center-right), content format (e.g., Reels vs. static images), and month. To ensure the robustness of our findings, we also controlled for the number of followers and posting frequency. These variables are derived from established literature on visual political communication and are tailored to the specific architectural affordances of Instagram.

To address the qualitative dimension (RQ2), we conducted a grounded-approach content analysis to identify the specific features that drive visibility within the platform. We manually coded a representative subset of high-visibility posts, employing an inductive process to categorize the audiovisual, textual and rhetorical nuances of content elements. This approach allowed us to distinguish between features designed to maximize algorithmic reach (platform-level affordances) and those intended to spark user engagement (individual-level interaction) with political content. By analyzing these overlapping layers, we uncover how polarizing candidates strategically align their messaging with Instagram's technical architecture.

Results

To address RQ1, we analyzed visibility patterns across ideological affiliations and individual candidates. To isolate the effects of ideology from individual persona, we first modeled visibility by candidate identity (Model 7) and then by political affiliation (Model 8).

Preliminary findings reveal a significant tension between institutional labels and algorithmic success. When analyzing candidates as ideological aggregates, affiliation with the right/center-right was associated with a significant decrease in views ($\beta = -.254$, $p < .001$) compared to the left/center-left baseline. However, Model 7 shows this is disrupted by individual performance: holding all variables constant, every other candidate received 56% to 93% fewer views than the reference (suspended) candidate. This confirms an "unequal distribution of voice" (Esau et al., 2024), where platform architecture amplifies specific actors.

While engagement (likes/comments) was positively associated with views, its effect was marginal compared to candidate category, the strongest predictor, as also founded by Boulianne & Larsson (2023). Crucially, the results contradict traditional assumptions regarding reach: follower count ($\beta = -.521$, $p < .001$) and posting frequency ($\beta = -.086$, $p <$

.001) were negatively associated with views. This supports Narayanan (2023) and Giglietto (2025), suggesting a shift to a recommendation-based system prioritizing post performance over established networks. Additionally, static photos slightly outperformed videos (+13–14%), and peak visibility occurred in August (40–51% higher than later months), revealing that algorithmic attention is established early in the cycle. Surprisingly, the suspended account reached visibility levels similar to institutional actors, such as the incumbent mayor.

To answer RQ2, our preliminary qualitative analysis followed a grounded approach. Given that RQ1 identified a single candidate accounting for over 50% of total visibility, the qualitative phase focused primarily on this actor's output to understand the drivers of such disproportionate reach, revealing that the most-viewed content was characterized by high-intensity performance rather than policy discourse. Peaks were tied to the disruptive performances of the suspended candidate, such as physical altercations and emotionally-driven content (ambulance/hospital videos and accusations of an opponent being cocaine-dependent). Mockery and aggressive personal attacks, and transgressive labeling (e.g., labeling opponents "Communists", even right-wing) created a binary "us vs. them" narrative, and contradicts existent literature where posts attacking opponents may decrease engagement (Boulianne & Larsson, 2023). The candidate used platform-native tactics to "trick" the engine, including systematic prompts (CTAs) and heavy emoji use to boost comment counts. We also found frequent reposting of snippets to maximize algorithmic lifecycle.

Future Work

While previous studies have analyzed engagement patterns among political actors, this work complements the existing literature by investigating the specific relationship between view counts and political visibility during a campaign. By incorporating controls such as posting frequency, follower count, and content format, and pairing them with a grounded content analysis, we have identified key features connected to algorithmic reach. However, several paths for future research remain.

First, further study is necessary to better understand the role of multimodal elements (such as the combination of audio, text, and visual cues). Future work could employ more advanced multimodal analysis to see how specific soundtracks or visual edits interact with the "interest graph" logic to boost visibility.

Second, our understanding of these platforms remains constrained by limited data access. Because we lack granular, time-series data from current platform tools, we cannot definitively assert the causal direction between different metrics - specifically, whether high views drive engagement, or if initial engagement pulses a post into wider view distribution. Future research should attempt to map these "visibility trajectories" using timestamped data to clarify the feedback loops between system recommendation and user action.

Finally, the geographical and technical bypasses discovered in this study - such as the use of VPNs to navigate judicial geoblocks - deserve more attention. Researching how decentralized "backup" accounts and platform architecture challenge national electoral laws will be essential for understanding the future of digital democratic safeguards.

References

Albuquerque, A., & Alves, M. (2023). Bolsonaro's Hate Network. Freie Universität Berlin. <https://doi.org/10.48541/DCR.V12.2>

Avalle, M., Di Marco, N., Etta, G. et al. (2024) Persistent interaction patterns across social media platforms and over time. *Nature*, 628, 582–589. <https://doi.org/10.1038/s41586-024-07229-y>

Bossetta, M. (2024). The problems with social media affordances and digital political campaigning. In D. Lilleker, D. Jackson, B. Kalsnes, C. Mellado, F. Trevisan, & A. Veneti (Eds.), *The Routledge handbook of political campaigning*. Routledge. <https://doi.org/10.4324/9781003333326>

Bossetta, M., & Schmøkel, R. (2022). Cross-platform emotions and audience engagement in social media political campaigning: Comparing candidates' Facebook and Instagram images in the 2020 US election. *Political Communication*, 40(1), 48–68. <https://doi.org/10.1080/10584609.2022.2128949>

Boulianne, S., & Larsson, A. O. (2021). Engagement with candidate posts on Twitter, Instagram, and Facebook during the 2019 election. *New Media & Society*, 25(1), 119–140. <https://doi.org/10.1177/14614448211009504>

Bruns, A., Dos Santos Choucair, T., Esau, K., Svegaard, S., & Vilkins, S. (2024). Polarization in online spaces: Distinguishing forms of polarized politics. In D. Lilleker, D. Jackson, B. Kalsnes, C. Mellado, F. Trevisan, & A. Veneti (Eds.), *The Routledge handbook of political campaigning*. Routledge. <https://doi.org/10.4324/9781003333326>

de-Lima-Santos, M. F., Gonçalves, I., Quiles, M. G., Mesquita, L., Ceron, W., & Couto Lorena, M. C. (2024). Visual political communication on Instagram: A comparative study of Brazilian presidential elections. *EPJ Data Science*, 13(72). <https://doi.org/10.1140/epjds/s13688-024-00502-0>

Esau, K., Choucair, T., Vilkins, S., Svegaard, S., Bruns, A., O'Connor-Farfan, K., & Lubicz-Zaorski, C. (2024). Destructive polarization in digital communication contexts: A critical review and conceptual framework. *Information, Communication & Society*. <https://doi.org/10.1080/1369118X.2024.2413127>

Farkas, X. (2023). Visual political communication research: A literature review from 2012 to 2022. *Journal of Visual Political Communication*, 10(2), 95–126.

https://doi.org/10.1386/jvpc_00027_1

Gerodimos, R. (2019). The interdisciplinary roots and digital branches of visual political communication research. In A. Veneti, D. Jackson, & D. G. Lilleker (Eds.), *Visual political communication* (pp. 53–73). Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-18729-3>

Giglietto, F. (2025). Amplifying extremes: The interplay of social media and traditional media in shaping political polarization. Manuscript submitted for publication.

Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie; P. J. Boczkowski, & K. A. Foot (Eds.), *Media Technologies: Essays on Communication, Materiality, and Society*. MIT Press. <https://doi.org/10.7551/mitpress/9042.003.0013>

Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.

Highfield, T., & Leaver, T. (2016). Instagrammatics and digital methods: Studying visual social media, from selfies and GIFs to memes and emoji. *Communication Research and Practice*, 2(1), 47–62. <https://doi.org/10.1080/22041451.2016.1155332>

Iyengar, S., Lelkes, Y., Levendusky, M., et al. (2019). The Origins and Consequences of Affective Polarization in the United States. *Annual Review Political Science*, 22, 129–146. <https://doi.org/10.1146/annurev-polisci-051117-073034>

Justiça Eleitoral determina suspensão de perfis de Pablo Marçal nas redes sociais usados para monetização. (2024, August 24). G1. <https://g1.globo.com/sp/sao-paulo/eleicoes/2024/noticia/2024/08/24/justica-eleitoral-determina-suspensao-de-perfis-de-pablo-marcal-nas-redes-sociais-usados-para-monetizacao.ghtml>

Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*. New York: Data & Society Research Institute.

Mendonça, R. F., & Caetano, R. D. (2020). Populism as parody: The visual self-presentation of Jair Bolsonaro on Instagram. *The International Journal of Press/Politics*, 26(1), 210–235. <https://doi.org/10.1177/1940161220970118>

Narayanan, A. (2023). *Understanding Social Media Recommendation Algorithms*. Knight First Amendment Institute: Algorithmic Amplification and Society. <https://doi.org/10.7916/khdk-m460>

Quandt, T., & Klapproth, J. (2023). Dark participation: Conception, reception, and extensions. In C. Strippel, S. Paasch-Colberg, M. Emmer, & J. Trebbe (Eds.), *Challenges and perspectives of hate speech research*, (pp. 251–270). <https://doi.org/10.48541/dcr.v12.15>

Recuero, R. (2024). The platformization of violence: Toward a concept of discursive toxicity on social media. *Social Media + Society*, 10(1).

<https://doi.org/10.1177/20563051231224264>

Stromer-Galley, J., Bryant, L., & Bimber, B. (2015). Context and medium matter: Expressing disagreements online and face-to-face in political deliberations. *Journal of Public Deliberation*, 11(1).

Terenzi, M. (2024). Modeling persuasion in social media: A theoretical approach to algorithmic content distribution and manipulation. *Journal of Sociocybernetics*, 19(1).

https://doi.org/10.26754/ojs_jos/jos.2024110992

Trilling, D., Kulshrestha, J., de Vreese, C., Halagiera, D. et al. (2022). Is sharing just a function of viewing? The sharing of political and non-political news on Facebook. *Journal of Quantitative Description: Digital Media*, 2.

<https://doi.org/10.51685/jqd.2022.016>

Yarchi, M., Baden, C., & Kligler-Vilenchik, N. (2021) Political Polarization on the Digital Sphere: A Cross-platform, Over-time Analysis of Interactional, Positional, and Affective Polarization on Social Media. *Political Communication*, 38(1-2), 98-139.

<http://doi.org/10.1080/10584609.2020.1785067>

8C.3 Sowing Discord: Exploring the use of online chat platforms in mobilising worker resistance

Daniel Jordan (University of Glasgow)

Introduction

This paper examines how workers in Scotland's video game industry use online chat platforms—primarily Discord, Slack, and WhatsApp—to mobilise collective resistance within an increasingly remote and precarious labour environment. The rise of remote and hybrid work has eroded many of the informal, embodied interactions that once underpinned workplace solidarity, replacing shared physical spaces with fragmented digital infrastructures. As a result, workers now turn to online chat environments not simply as communication tools but as political and organisational spaces where grievances can crystallise into mobilisation.

Scotland's games ecosystem provides a rich empirical context for studying these dynamics. Recent crises, such as mass redundancies at Edinburgh-based Build a Rocket Boy and reports of union suppression at Rockstar North, show how live these struggles are. At the same time, well-established relationships with local higher education institutions ensure a ready supply of cheap and naïve labour (Sutherland, 2018); workers without an embedded sense of trade union culture. Both the industry's skewed composition and hypercompetitive nature complicate attempts to form unified collective subjects. Following Keogh's (2023) argument that the 'videogame industry' is a fragmented ecology of production, the paper examines how these structural divisions shape online mobilisation in Scotland.

Across the interviews conducted for this project, workers describe a labour environment marked by intense precarity, competitive pressures, and long-standing issues such as crunch, job insecurity, and limited career progression. These structural problems intersect sharply with digital mediation: workplaces built around Slack or Teams are also workplaces in which surveillance, logging, and managerial oversight are embedded into everyday interactions. Scholarship on digital labour and social media underscores how platforms structure visibility, identity performance, and collective identity formation (Boyd, 2010; Burgess & Baym, 2020; Gillespie, 2019). In this case, workers actively repurpose chat platforms—many of which were designed for managerial control rather than worker autonomy—to craft alternative organisational spaces. In doing so, they simultaneously resist and work around the platform logics that constrain them.

This research therefore asks not only how workers organise without a shared workplace, but also how chat platforms themselves become part of the organisational

infrastructure. Particularly their role in affording, restricting, or reshaping the possibilities for mobilisation. By situating worker organising within social media and internet studies, this project pushes beyond a traditional labour-movements framing and offers insight into platformised forms of solidarity emerging in creative and tech-adjacent industries.

Objectives/Research Questions

How do online chat platforms shape the practices of worker organising among Scotland's game workers?

Which platform affordances and constraints facilitate or inhibit collective action and mutual support?

How do concerns about employer surveillance, precarity, and workplace governance influence platform choice and organising tactics?

Methods

This study uses a qualitative case-study design to investigate how Scottish game workers mobilise through digital communication platforms. Data collection involved more than 30 semi-structured interviews with current and former game workers, trade union organisers, industry stakeholders, and participants directly involved in online organising efforts. Recruitment occurred through trade unions (including the IWGB Game Workers Branch), social media outreach (LinkedIn, Reddit), professional networks, and snowball sampling to ensure representation across job roles and studio sizes. As the literature on labour in game development highlights, roles such as QA, programming, design, art, and production differ significantly in precarity and visibility; interview sampling was therefore sensitive to these distinctions.

The interviews examined participants' communication practices, platform preferences, organisational histories, and strategies for navigating or resisting surveillance. Particular attention was paid to how workers differentiate employer-controlled platforms (e.g., Teams, corporate Slack) from personal platforms (e.g., Discord, WhatsApp), and how those distinctions shape collective action. Many participants also provided examples of specific disputes, organising campaigns, layoffs, and peer-support episodes mediated through chat platforms.

Data analysis followed an abductive thematic approach, integrating insights from social media and internet studies with mobilisation theory. Whereas mobilisation theory provides tools for understanding interests, organisation, opportunity, and repertoires of action (Tilly, 1978), social media scholarship highlights the sociotechnical features—affordances, governance, visibility, and connective practices—that shape online collective behaviour (Jenkins et al., 2020; Z. A. Papacharissi, 2015; Treem & Leonardi, 2013). Coding therefore examined both what

workers were doing (mobilising, educating, coordinating) and how platform affordances structured those practices (persistence, ephemerality, privacy controls, moderation, channel architecture).

In addition to interview data, this project incorporates observation of publicly accessible Discord servers and union communication materials. These materials help contextualise how cultural and affective practices (such as humour, in-jokes, and collective emotional labour) contribute to the formation of shared identities and mobilising narratives.

Results

Preliminary analysis reveals a complex relationship between platform affordances, workplace structures, and emerging forms of collective action. Five interlocking themes capture how game workers use chat platforms to organise under conditions of precarity and surveillance.

First, workers repeatedly emphasised that Discord servers and private group chats function as the only sustained spaces where colleagues can interact informally. These channels serve as hubs for mutual aid, skill-sharing, and emotional support. They also provide the infrastructure for circulating resources such as contract advice, union materials, and information about industry norms (e.g., acceptable pay bands, experiences with specific studios). This aligns with broader research showing that online spaces can facilitate networked publics and shared meaning-making (Boyd, 2010; Papacharissi, 2013). For workers without a physical workplace, these platforms effectively become new organisational commons.

Second, employer-controlled platforms were widely described as sites of risk, where workers feared disciplinary action, monitoring, or reputational damage. Participants recounted examples of colleagues being cautioned for criticising management in official channels. Consequently, workers move sensitive discussions to encrypted or worker-run platforms. This pattern reflects findings from the broader platform governance literature: perceived surveillance, data extraction, and opaque moderation shape user behaviour in significant ways (Crawford, 2021; Gillespie, 2019). The tactical creation of “back-stage” organising spaces is a deliberate response to these platform conditions.

Third, interviewees described how Discord and WhatsApp networks enabled rapid coordination during crises—such as mass redundancies or sudden crunch periods. Organisers used these channels to distribute union sign-up links, schedule meetings, and coordinate collective responses. Moments of acute workplace pressure often catalysed spikes in activity, with chat platforms amplifying workers’ ability to gather information and build critical mass quickly. These findings resonate with research on connective action and digitally-enabled mobilisation, where personalised sharing and

flexible, networked communication reduce barriers to participation (Bennett & Segerberg, 2013).

Last, across interviews, participants highlighted the importance of humour, memes, and emotional support in sustaining involvement. Memes about crunch, burnout, or bad management circulated widely, giving workers a shared language through which to articulate grievances. This mirrors social media scholarship on the affective dimensions of online political participation (Papacharissi, 2015), and the role of vernacular creativity in shaping collective identities (Burgess, 2006). These affective connections help transform isolated frustrations into collective narratives, reinforcing a sense of shared struggle.

Future Work

Future stages of this work could advance the empirical, methodological, and theoretical contributions in several ways. First, by analysing platform architectures, governance models, and interactional norms, the project will assess how Discord, Slack, Teams, WhatsApp, and Signal differently shape organising capacity. Concepts such as visibility, persistence, searchability, and data governance will guide this analysis (Treem & Leonardi, 2013).

Second, pending participant consent, the project will integrate digital trace analysis and network mapping to visualise how information and mobilisation flow across worker-run servers. Ethical frameworks from internet research will underpin all methodological decisions.

Third, the project aims to articulate a model of platform-mediated mobilisation that synthesises Tilly's categories with social media research on connective action, affect, and platform governance. This synthesis will contribute to scholarship on digital labour movements, creative labour, and the politics of platform infrastructures.

Fourth, given that similar organising dynamics occur in adjacent creative and tech industries, future work will test the transferability of findings beyond game development to animation, VFX, and software engineering. Together, these extensions will produce a richer, more integrated account of how worker organising unfolds within platformised communication environments.

References

- Bennett, W. L., & Segerberg, A. (2013). *The logic of connective action: Digital media and the personalization of contentious politics*. Cambridge University Press.
- boyd, d. (2010). Social network sites as networked publics: Affordances, dynamics, and implications. In Z. Papacharissi (Ed.), *A networked self: Identity, community, and culture on social network sites* (pp. 49–68). Routledge.

- Burgess, J. (2006). Hearing ordinary voices: Cultural studies, vernacular creativity and digital storytelling. *Continuum* (Mount Lawley, W.A.), 20(2), 201–214.
- Burgess, J., & Baym, N. K. (2020). *Twitter: A biography*. New York University Press.
- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Gillespie, T. (2019). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
- Jenkins, H., Kligler-Vilenchik, N., Zimmerman, A., Shresthova, S., & Gamber-Thompson, L. (2020). *By any media necessary: The new youth activism*. New York University Press.
- Keogh, B. (2023). *The videogame industry does not exist: Why we should think beyond commercial game production*. MIT Press.
- Papacharissi, Z. A. (2013). *A private sphere: Democracy in a digital age* [PDF]. Polity Press.
- Papacharissi, Z. A. (2015). *Affective publics: Sentiment, technology, and politics*. Oxford University Press.
- Sutherland, J. (2018). 20 years of computer games degree programs: From Abertay in 1997 to York in 2017. *The Computer Games Journal*, 7(3), 149–152.
- Tilly, C. (1978). *From mobilization to revolution*. Addison Wesley.
- Treem, J. W., & Leonardi, P. M. (2013). Social media use in organizations: Exploring the affordances of visibility, editability, persistence, and association. *Annals of the International Communication Association*, 36(1), 143–189.
- van Doorn, N. (2017). Platform labor: on the gendered and racialized exploitation of low-income service work in the “on-demand” economy. *Information, Communication and Society*, 20(6), 898–914.

8C.4 “Don’t hurt me, no more”: What can we learn from the pharmaceutical industry regulation about platform and AI regulation

Elinor Carmi (City, University of London)

Introduction

In the past decade there has been a substantial growth of exploitative practices by platforms and AI industries. Among the reasons these practices are possible is because tech companies objectify, dehumanize and decontextualize humans from the data they extract (Carmi, 2020; D'ignazio and Klein, 2020; Garcia et al., 2020). In turn, these industries make excuses built around the idea that “it’s just data”.

But whether platforms and AI companies want to acknowledge it or not - data are people. Yet, so far people working in these industries have managed to dehumanize data and make it appear abstract and lacking in context (Avnoon et al., 2023). In many respects these industries are conducting unregulated experiments on humans (Carmi, 2020; Zuboff, 2015). This type of dehumanization has been an ongoing practice for the past three decades. While tech companies have created different ethics boards and committees in the past few years, they have later scaled back many ethics and content moderation teams (Benavidez, 2023). Importantly, the way that each company applies these ethics recommendations is unclear, unregulated and far from transparent.

Another harmful industry that was unregulated for many years is the pharmaceutical industry. The medical industry had its reckoning moment after the atrocities of WW2, where horrible experiments were conducted on humans. The lessons of these harmful experiments made people realize that physicians cannot do whatever they want to people. That resulted in the 1947 Nuremberg Code which pioneered the need for consent of participants and the protections of their rights.

But only in 1996, the International Conference for Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) issued new guidelines – The Good Clinical Practice (GCP). The GCP is an international ethical and scientific quality standard for clinical trials in all of its procedural stages (Vijayanathan and Nawawi, 2008).

Trying to leverage the ethical procedures of other industries into the tech industry is not new. Diakopoulos and colleagues (2024) have taken the ethical codes of journalists as a starting point to inform the design of responsible AI systems. Similarly, this project aims to examine the ethical standards of the medical industry when conducting research on human subjects, especially as part of the creation of commercial

pharmaceuticals. This is to leverage that knowledge into the platforms and AI industries. As Véliz argues (2019), despite differences between the medical world and the digital world, there are three elements that digital ethics should adopt from medical ethics: “the development of ethical codes, the use of ethics committees, and respect for personal autonomy” (Véliz, 2019, p. 317). Doing so, will enable us to create new creative directions in platform and AI regulation.

Objectives/Research Questions

What can we learn from the human clinical trials standard Good Clinical Practice about regulating platforms and AI industries?

Methods

This paper undertakes two research methods: first, I analyse the three versions of the GCP to understand the structure and the differences between the versions. Second, I conducted semi-structured interviews with six practitioners who work in the clinical trials industry. The project has gone through the ethics procedures in my university, and has been approved. I conducted semi-structured interviews via Zoom with practitioners in human clinical research to identify the opportunities and challenges of their everyday practices. Purposive sampling (Kelly, 2010) was employed to identify and involve GCP practitioners with diverse experiences and roles with clinical trials. Access to these practitioners has been achieved through two clinical trials practitioners who work in the European Union and the UK. The sampling was intended to be illustrative and not representative of clinical trials.

These experts operate under strict NDA contracts that require discretion, which may discourage them from participation. Therefore, participants' names have been anonymized. The interviews were conducted during May and June 2024. The experts that were interviewed hold three positions in the practitioner and commercial domains of the clinical trial chain: study coordinator, clinical research associate, and data manager. The purpose behind the interviews is to translate the procedural pathways of human experiments and implement good clinical practice guidelines within the tech industry (specifically platforms and AI). The transcripts were analyzed using Critical Thematic Analysis (CTA), which explores patterned meaning (Coffey & Atkinson, 1996). I began with thematic analysis, following the phases outlined by Braun and Clarke (2021) and then interrogated the themes with a critical approach to capture the complexity of conducting trials in a profit and commercially driven environment.

Results

The preliminary analysis shows six takeaways:

- 1) The importance of a global binding standard of practice - this ensures that wherever you conduct the trial you can count on the same quality and ethical procedures to be adhered to.
- 2) Mandatory accreditation training to all employees - As all participants and the GCP documents show, the strength of the standard is the fact that all employees must go through this training and the accreditation must be updated every two years, or when and if you move to a new company or trial.
- 3) Do no harm and people's safety first - This is applied throughout the whole process of the trial, from ensuring the right people are recruited, and that the consent process is dynamic and allows people more agency to reconsider their participation at any point.
- 4) Incentives for being good - Pharmaceutical companies prefer to work with practitioners who have higher ethical standards because it gives their trial more validity and prestige.
- 5) Multiple audit points create self-policing.
- 6) Prove the benefit and not rely on hype.

Future Work

It is important to emphasize that while there are a lot of similarities between the pharmaceutical industry and the platforms and AI industries, ultimately these industries are not identical. Nevertheless, future work in this area should explore other industries that have been successful in regulating their dangerous industries, for example, aviation and automobiles. Such examinations would help illuminate how profit-driven industries develop standards, processes and training that ensure people's safety and wellbeing while still being commercially driven.

References

- Avnoon, N., Kotliar, D. M., & Rivnai-Bahir, S. (2023). Contextualizing the ethics of algorithms: A socio-professional approach. *New Media & Society*, 26(10), 5962-5982.
- Benavidez, N. (2023). Big-Tech Backslide: How Social-Media Rollbacks Endanger Democracy Ahead of the 2024 Elections. Free Press Report.
<https://www.freepress.net/big-tech-backslide-report>.
- Braun, V., & V. Clarke. (2021). *Thematic Analysis: A Practical Guide*. London: Sage.
- Carmi, E. (2020). *Media distortions: Understanding the power behind spam, noise, and other deviant media*. Peter Lang International Academic Publishers.
- Carmi, E. (2021). A feminist critique to digital consent. In *Seminar. net* (Vol. 17, No. 2). OsloMet-Oslo Metropolitan University.

- Coffey, A., & Atkinson, P. (1996). Making sense of qualitative data: Complementary research strategies. Sage Publications.
- Diakopoulos, N., Trattner, C., Jannach, D., Costera Meijer, I., & Motta, E. (2024) Leveraging Professional Ethics for Responsible AI: Applying AI techniques to journalism. *Communications of the ACM*, 67(2), 19-21.
- D'Ignazio, C., & Klein, L. F. (2020). *Data Feminism*. Cambridge: MIT press.
- Garcia, P., Sutherland, T., Cifor, M., Chan, A. S., Klein, L., D'Ignazio, C., & Salehi, N. (2020). No: Critical refusal as feminist data practice. In conference companion publication of the 2020 on computer supported cooperative work and social computing (pp. 199-202).
- Kelly, S. E., (2010). Qualitative interviewing techniques and styles. In: Bourgeault, I., Dingwall, R. & De Vries, R. (Eds.), *The SAGE Handbook of Qualitative Methods in Health Research*, (pp. 307-326). Sage.
<https://doi.org/10.4135/9781446268247.n17>
- Véliz, C. (2019). Three things digital ethics can learn from medical ethics. *Nature Electronics*, 2(8), 316-318.
- Vijayananthan, A., & Nawawi, O. (2008). The importance of Good Clinical Practice guidelines and its role in clinical trials. *Biomedical imaging and intervention journal*, 4(1).
- Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30(1), 75-89.

Session 8D - Platformisation

Day 2 · Tues 14 July · 13:50-15:20, Studio 2, ARC

8D.1 Creators, journalists, and relational precarity in India's beauty web

Anuja Premika (Mahindra University)

Introduction

India's beauty and personal care (BPC) industry was valued at USD \$21 billion (INR ₹1.79 lakh crores) in 2023, and is being touted as one of the fastest growing BPC markets globally. This growth has played out against a rapidly evolving media landscape, one that is widely recognised as being one of the strongest catalysts of the industry's expansion. In my doctoral work, I examine what I refer to as "the beauty web," the media and communications ecosystem of the beauty industry, produced by individual content creators on social media, brand communications, as well as beauty journalists and editors. In this paper, drawing from my doctoral work, I discuss how the beauty business in India is propped up by the precarious entrepreneurialism of individual media workers of various classes—including journalists, content creators, and multi-hyphenate communications professionals who defy easy classification—who differ in their self-identities and skillsets, but face a similar lack of institutional affiliation and support in their labours, and whose careers are predicated on maintaining ever-precarious relationships with beauty brands and the public relations agencies that represent them, what I refer to as relational precarity.

In studies of creator cultures, economies, and practices, labour has served as a key conceptual lens, and has been theorised extensively (Abidin & Brown, 2018). Duffy's conceptualisation of "aspirational labour" has proved particularly influential, defined as labour that is mostly unpaid, independent, and fuelled by the ideal of "getting paid to do what you love," all in the hope that workers will "one day be compensated for their productivity" (2022, p. 4). In the face of fickle and unpredictable changes in advertiser expectations and audience tastes (Duffy et al., 2021), and opaque algorithmic systems that control creators' visibility and income (Bishop, 2018), creators are pushed to become "jack-of-all-trades entrepreneurs" (Glatt & Banet-Weiser, 2021), combining roles across media production, distribution, and marketing.

Duffy (2022) notes that as the creator industry became increasingly professionalised, creators came to mirror the very traditional media workers they had once positioned

themselves in contrast against. On the other hand, what Bishop calls “influencer creep” (2023) has become commonplace in many fields of work—the modes of visibilising oneself in digital spaces, that creators first championed, have now “crept” into other occupations, and an increasing number of people adopt the same strategies in the construction of their online presence. Beauty creators research and review products as beauty journalists once did, and creator grammars “creep” into journalistic creative outputs. Forming the context for this convergence are shifting trends in both media and publishing, and the beauty business in India, further complicating these blurring lines between the practices and labours of different actors in the space.

Objectives/Research Questions

In this paper, I examine the labours of two classes of producers that challenge the professional-amateur dichotomy of creator culture (Burgess & Green, 2009)—content creators and journalists—to find that despite their distinct self-identities, perceptions of quality and competence, and the unique ways in which they’ve fashioned their own roles within the industry, they experience similar precarities in their labours, particularly in their relationships with brands, and consequently in their careers. As a result, both creators and journalists not only engage in unpaid “aspirational labour” (Duffy, 2022) in the early stages of their careers in the hopes of one day “making it,” but they continue to engage in the largely unremunerated relational labour of sustaining precarious brand relationships even after achieving a degree of status and recognition in the industry. Their practices are underpinned by pressures to labour for recognition from brands, continually and (often) for free.

Methods

My doctoral work employs the methodological approach of digital ethnography. Ethnography as a method has become established as an appropriate approach to media research, Ardévol and Gómez-Cruz (2013) assert, particularly for works oriented around “cultures” of media production, consumption, and reception (p. 499). In the first phase of data collection, I conducted a mixed-methods content analysis of over 800 Instagram posts from brands and creators, and 25 YouTube videos from creators. I adapted my framework for analysis from Bock et al.’s (2001) and Lister & Wells’ (2004) frameworks for visual analysis. In the second phase, I carried out interviews with 15 creators, journalists, and industry professionals. I thematically analysed the interview transcripts using the framework for Reflexive Thematic Analysis (RTA) described by Braun and Clarke (2006, 2012).

Results

Both creators and journalists not only engage in unpaid “aspirational labour” (Duffy, 2022) in the early stages of their careers in the hopes of one day “making it,” but they continue to engage in the largely unremunerated relational labour of sustaining

precarious brand relationships even after achieving a degree of status and recognition in the industry. The un- or under-paid labour of many creators generates profits for a few, mostly brands, and the majority of smaller creators never see a share of it, even if brands repay them with visibility. The labour of visibilising themselves for little to no compensation is never fully over, as brands continue to expect “organic” coverage even from bigger creators that they have ongoing relationships with, in the form of unpaid brand mentions and favourable reviews even without explicit compensation.

In a case of influencer creep (Bishop, 2023), journalists experience their own kinds of visibility mandates, feeling the need to “put [themselves] out there” (Duffy & Hund, 2019) on social media, in the interest of furthering their careers—careers that increasingly involve not just independently contributing to multiple publications, but becoming precarious entrepreneurs, working with and for brands in various capacities, as writers, marketers, public relations personnel, and consultants.

Underpinning this is the asymmetry in creator- and journalist-brand relationships—brands control access to future avenues for paid work. What were formerly media-brand relationships (say between magazines and brands), have now become creator- or journalist-brand relationships, where the onus to keep the relationships going falls more on individual workers than on brands. Instead of paying advertising revenue to magazines, and these benefits then getting distributed to journalists in the form of stable jobs and regular incomes, brands now offer a host of barter-like arrangements and perks (including free products, experiences, and trips) to individual workers directly, but don’t always pay them. Aspirational labour is not only the unpaid labour of creators hoping to “make it,” it is the underlying logic of an industry where “making it” does not necessarily guarantee stability or longevity to workers, who experience relational precarity in their associations with brands. Journalists need access to produce media, and they need to produce brand-favourable media to enjoy continued access, much like creators who need to present themselves as good “brand fits,” and negotiate honesty in product reviews.

Despite this loop, freelancing alone doesn’t pay the journalists’ bills—they need to work with brands directly, as creators do, but also in other ways, functioning as consultants in communications, branding, promotion, events, and even product formulation, entrepreneurially channelling their know-how to deepen and sustain their relationships with brands.

Future Work

India's creator ecosystem is largely unregulated at the present moment, barring for a few nascent attempts at industry self-regulation, and government directives on disclosure in advertising and sponsorship. Given India's postcolonial history of failing to regulate informal labour of various kinds (Katrak & Kulkarni, 2025), I hope to build on

this work to develop policy and recommendation frameworks for the regulation of creator labours and brand-creator relationships in India.

References

- Abidin, C., & Brown, M. L. (2018). Introduction. In C. Abidin & M. L. Brown (Eds.), *Microcelebrity Around the Globe: Approaches to Cultures of Internet Fame*. Emerald Publishing Limited. <https://doi.org/10.1108/9781787567498>
- Ardévol, E., & Gómez-Cruz, E. (2013). Digital Ethnography and Media Practices. In *The International Encyclopedia of Media Studies* (pp. 498–518). Wiley. <https://doi.org/10.1002/9781444361506.wbiems193>
- Bishop, S. (2018). *Beauty Vlogging: Practices, Labours, Inequality* [University of East London, School of Arts and Digital Industries]. <https://doi.org/10.15123/PUB.7915>
- Bishop, S. (2023). Influencer creep: How artists strategically navigate the platformisation of art worlds. *New Media & Society*. <https://doi.org/10.1177/14614448231206090>
- Bock, A., Isermann, H., & Knieper, T. (2011). Quantitative Content Analysis of the Visual. In E. Margolis & L. Pauwels (Eds.), *The SAGE Handbook of Visual Research Methods* (pp. 265–282). SAGE Publications Ltd. <https://doi.org/https://doi.org/10.4135/9781446268278>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., & Clarke, V. (2012). Thematic analysis. In *APA handbook of research methods in psychology*, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological. (pp. 57–71). American Psychological Association. <https://doi.org/10.1037/13620-004>
- Burgess, J., & Green, J. (2009). The entrepreneurial vlogger: Participatory culture beyond the professional/amateur divide. In P. Snickars & P. Vonderau (Eds.), *The YouTube Reader*. National Library of Sweden.
- de Peuter, G. (2011). Creative Economy and Labor Precarity. *Journal of Communication Inquiry*, 35(4), 417–425. <https://doi.org/10.1177/019685991141636>
- Duffy, B. E. (2022). *(Not) Getting Paid to Do What You Love: Gender and Aspirational Labor in the Social Media Economy*. Yale University Press.
- Duffy, B. E., & Hund, E. (2019). Gendered Visibility on Social Media: Navigating Instagram's Authenticity Bind. *International Journal of Communication*, 13, 4983–5002. <https://ijoc.org/index.php/ijoc/article/view/11729/2821>

- Duffy, B. E., Pinch, A., Sannon, S., & Sawey, M. (2021). The Nested Precarities of Creative Labor on Social Media. *Social Media + Society*, 7(2).
<https://doi.org/10.1177/20563051211021368>
- Glatt, Z., & Banet-Weiser, S. (2021). Productive Ambivalence, Economies of Visibility, and the Political Potential of Feminist YouTubers. In *Creator Culture* (pp. 39–56). New York University Press.
<https://doi.org/10.18574/nyu/9781479890118.003.0006>
- Kuehn, K., & Corrigan, T. F. (2013). Hope Labor: The Role of Employment Prospects in Online Social Production. *The Political Economy of Communication*, 1(1), 9–25.
<https://polecom.org/index.php/polecom/article/view/9>
- Katrak, M., & Kulkarni, S. (2025). The regulation of influencer labour in India: Situating a novel form of labour amidst colonial continuities of informality. In *The Hashtag Hustle*
- Lister, M., & Wells, L. (2004). Seeing Beyond Belief: Cultural Studies as an Approach to Analysing the Visual. In T. Van Leeuwen & C. Jewitt (Eds.), *The Handbook of Visual Analysis* (pp. 61–91). SAGE Publications Ltd.
<https://doi.org/10.4135/9780857020062.n4>
- Pathak, G., & Nichter, M. (2018). Cleanups, confidence, and cosmetics: Marketing beauty in India. *Journal of Consumer Culture*, 1–21.
<https://doi.org/10.1177/1469540518818631>
- Terranova, T. (2000). Free Labor: Producing Culture for the Digital Economy. *Social Text*, Vol. 18 No. 2, 18(2), 33–58. <https://muse.jhu.edu/article/31873>

8D.2 Algorithmic Visibility as a Capital Conversion Mechanism: The Practice Logic of Content Creators in the Digital Public Sphere

Ben Xu (University of Leeds)

Introduction

As social media platforms become increasingly institutionalised within public life, content creators have emerged as key actors in the digital public sphere. Through sustained content production, affective interaction, and ongoing negotiation with platforms, they participate in the construction and circulation of public issues, thereby generating new forms of social influence. However, success and failure on platforms cannot be attributed solely to individual creativity or effort, but are deeply embedded in the interplay of platform algorithms, capital logics, and pre-existing social structures. In environments where algorithmic recommendation systems dominate content distribution, being ‘seen’ itself becomes a scarce resource, the allocation of which profoundly shapes the forms of cultural expression and the structure of public discourse.

Drawing on the empirical context of Chinese social media platforms, this study examines how algorithmic visibility operates as a key mediating mechanism in the digital public sphere, participating in the conversion and redistribution of cultural capital. While existing research has highlighted the disciplining effects of algorithms on creators through lenses such as platform governance, algorithmic power, and digital labour, it has paid comparatively less attention to how algorithms intervene in the process of capital conversion in everyday creative practices, thereby reshaping the structural positions of different types of creators within the digital public sphere. Building on Bourdieu’s field theory, digital platforms are conceptualised here as competitive fields shaped by algorithms, capital, and cultural norms. Within this field, content creators enter competition with different forms of cultural capital, and their influence is translated into visibility and symbolic authority through algorithmic filtering of expressive forms, interactional rhythms, and affective styles. This study argues that algorithmic visibility should not be understood as a purely technical outcome, but as a socialised mechanism mediating capital.

Objectives/Research Questions

This study is organised around the following three interrelated research questions:

1. In the context of Chinese platforms centred on algorithmic recommendation, how is content creators’ cultural capital translated into visibility and symbolic authority?

2. What mediating role do algorithms play in processes of capital conversion, and how are these mechanisms realised through creators' everyday practices?
3. How do different types of creators with high levels of cultural capital develop differentiated logics of practice between algorithmic discipline and platform incentives, thereby maintaining their positions in the digital public sphere?

By addressing these questions, this study seeks to move beyond analytical approaches that treat algorithms as a one-dimensional form of technical power, and instead emphasises the dynamic interplay between creators' practices, capital structures, and algorithmic mechanisms.

Methods

This study adopts a theory-driven qualitative case study approach. From a methodological perspective, this study draws on Bourdieu's field theory, theories of capital conversion, and practice theory to conceptualise digital platforms as structured yet not closed spaces of competition.

The empirical materials are drawn from representative content creators active on major Chinese social media platforms, with particular attention to two types of actors with high symbolic significance. One type consists of cultural elite creators centred on performance and affective orchestration, while the other comprises knowledge-oriented creators grounded in academic authority and knowledge dissemination. Illustrative cases include a prominent performance-oriented creator and a high-profile legal scholar-creator on Chinese platforms. The materials analysed include publicly available video content, records of platform interactions, media reports, and relevant secondary research.

Through a typological approach, the case analysis reveals how different forms of cultural capital are re-encoded, recognised, and amplified in China's highly algorithmised platform environment. This allows for an understanding of how creators internalise, respond to, and strategically deploy platform algorithmic logics in their everyday content production.

Results

The findings indicate that, within China's digital public sphere, algorithmic visibility constitutes a key mediating form of capital that determines whether, and how, cultural capital can be converted into symbolic authority and economic returns.

First, through the encoding of content rhythms, affective intensity, and modes of interaction, algorithms render certain forms of cultural expression more readily recognisable as high-value signals. In the case of performance-oriented cultural elite creators, institutionalised cultural capital is re-encoded under algorithmic conditions

as high completion rates and strong engagement, facilitating a rapid conversion from cultural capital to visibility capital and then to economic capital.

Second, the practices of knowledge-oriented creators reveal an alternative pathway of capital conversion. Within the Chinese platform environment, academic authority cannot be directly converted into communicative advantage, but must instead be mediated through narrative, affective, and media-oriented modes of expression in order to render institutionalised cultural capital compatible with algorithmic logics. Through such practices, knowledge-oriented creators are able to construct a form of 'visibility authority', achieving sustained public visibility while maintaining academic orthodoxy.

Finally, the two sets of cases jointly demonstrate that algorithms do not dissolve cultural inequalities, but instead redistribute existing capital structures through new regimes of visibility. Under algorithmic conditions, cultural capital is reorganised into calculable and perceptible forms, with its effects being particularly pronounced within China's highly concentrated and intensely competitive platform ecology.

Future Work

By taking creators with high levels of cultural capital as its analytical entry point, this study has demonstrated how algorithmic visibility operates as a key mediating mechanism, participating in the conversion and redistribution of cultural capital within the digital public sphere. Given the high degree of concentration and institutionalisation characterising the Chinese platform environment—particularly in terms of algorithmic governance intensity, commercialisation, and policy intervention—future research will further examine how this context amplifies or reshapes the operation of algorithmic visibility as a mechanism of capital conversion.

First, within grassroots and rural creator communities, algorithmic visibility is frequently activated through 'authenticity', understood as an affective and everyday signifier. Future research may examine how authenticity, through the combined operation of algorithmic recognition and audience affective engagement, is transformed into a calculable and consumable symbolic resource, and how this process re-encodes rural–urban imaginaries and class distinctions.

Second, MCN organisations, as meso-level organisational spaces situated between platforms and creators, may stabilise and amplify algorithmic visibility at an institutional level, transforming individualised creative practices into semi-industrialised, data-driven content production processes. Further research on MCN organisations would help to reveal how algorithmic logics are embedded in creators' everyday labour processes through organisational mechanisms.

Third, among mid-level professional creators, competition over algorithmic visibility is frequently marked by heightened uncertainty and ongoing emotional strain. Future

research may adopt the lens of ‘algorithmic anxiety’ to examine the relationships between recommendation mechanisms, income volatility, and professional habitus, and to analyse how algorithmic governance shapes creators’ labour experiences at the affective level.

Through these extensions, future research will be able to move beyond analyses centred on a single creator type, advancing a more integrated understanding of algorithmic power, capital conversion, and structural inequalities in creative labour within the digital public sphere, and laying the groundwork for cross-platform and cross-cultural comparative research.

References

- Bourdieu, P. (1977). *Outline of a theory of practice* (R. Nice, Trans.). Cambridge University Press.
- Bourdieu, P. (1992). *The logic of practice* (R. Nice, Trans.). Polity.
- Bucher, T. (2018). *If...then: algorithmic power and politics*. Oxford University Press.
- van Dijck, J. (2013). *The culture of connectivity: a critical history of social media*. Oxford University Press.
- van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: public values in a connective world*. Oxford University Press.

8D.3 Platform-Mediated Labour and Professional Precarity: Producing Preconditions for Union Organizing in Journalism

Errol Salamon (University of Stirling)

Introduction

Since the 2000s, news media organizations in the United States and Canada have experienced significant platformization, integrating cloud-based collaboration platforms such as Slack and Google Workspace, instant messaging applications like WhatsApp and Microsoft Teams, content management systems, including WordPress, and analytics dashboards such as Chartbeat and Parse.ly into their labour processes. Although existing research has examined the integration of these technologies into newsroom routines (e.g., Bunce et al., 2018; Cohen, 2019; Hayes & O'Sullivan, 2024; Petre, 2018), further research is needed to understand how platform-mediated labour processes simultaneously generate exploitation and foster the subjective conditions necessary for professional news media workers to recognize collective interests. This study addresses this gap by analyzing the contradictory dynamics of platform-mediated newswork during a period of unprecedented unionization, from 2015 to 2022, when more than 7,500 newswriters joined unions in the US and Canada (Author; Fu, 2021). This analysis draws on the critical political economy of platform capitalism (Nieborg & Poell, 2018; Poell et al., 2023), theories of corporate ownership and control (McChesney & Nichols, 2022; Pickard, 2020), and labour process theory (Braverman, 1998; Cohen, 2019). It examines how platforms that facilitate dispersed collaboration also serve as mechanisms of surveillance, impose pressures for constant availability, and blur work-life boundaries, thereby intensifying precarious work and making exploitation increasingly visible and intolerable to workers. These contradictions have catalyzed collective organizing. The findings indicate that platformization generates inherent tensions: platforms intended to enhance efficiency also produce anxiety, burnout, and heightened awareness of shared precarity, which in turn stimulate newswriters' collective resistance. The study demonstrates how platform capitalism creates contradictions that foster the emergence of distinct white-collar, professional class subjects capable of collective action to defend both quality working conditions and journalism's democratic mission in society.

Objectives/Research Questions

This paper addresses three interconnected research questions:

RQ1: In what ways did union-represented newswriters experience and interpret digital restructuring through platform technologies, including Slack, Zoom, content management systems, and analytics dashboards, within their daily labour processes?

RQ2: How do platform-mediated work practices both facilitate collaboration and intensify managerial control and precarity?

RQ3: In what ways does the tension between journalists' white-collar professional identity—grounded in journalism's democratic mission and editorial autonomy—and the structural conditions of platform capitalism foster recognition of the necessity for collective organizing?

The study introduces the concept of “professional precarity consciousness” to explain how platform-mediated labour processes transform professional identity into a distinct form of white-collar digital class consciousness among news media workers without requiring them to abandon normative professional commitments.

Methods

This study employs thematic analysis (Fereday & Muir-Cochrane, 2006) of 42 in-depth semi-structured interviews with union-represented newswriters between October and December 2021. Participants included workers from digital-native media companies and legacy print and broadcast outlets who organized with The NewsGuild (TNG), Writers Guild of America East (WGAE), SAG-AFTRA, or CWA Canada, the primary unions representing newswriters during the study period (Author; Cohen & de Peuter, 2020; Fu, 2021). Interviewees were recruited through social media accounts and email, concluding upon data saturation. The research received institutional ethics approval, and all participants provided written informed consent. To encourage candid responses, participants were offered anonymity. Interview topics included daily experiences with platform technologies, perceptions of digital tools' impact on working conditions and autonomy, connections between professional journalism identity and collective organizing, and the union organizing process. The analysis followed an abductive approach, drawing on the critical political economy of platform capitalism, theories of corporate ownership and control, and labor process theory, with particular attention to technological control, precarious work, workers' capacity for resistance, and journalism's democratic mission and public good role (Author; Cohen, 2019; McChesney & Nichols, 2022; Poell et al., 2023). ATLAS.ti and NVivo were used for coding, generating 74 first-order descriptive codes, which were folded into three analytical themes addressing tensions among precarious work, professional identity, and collective consciousness.

Results

First, interviewees indicated that platforms effectively replaced physical offices and facilitated collaboration across geographically distributed teams. Workers used Slack for editorial meetings, story pitching, and workflow coordination along with content management systems (WordPress), collaborative writing tools (Google Docs), and video conferencing platforms (Zoom and Google Meet). However, these technologies

intensified platform-mediated managerial control. Workers reported a heightened sense of being monitored through Slack's status indicators, increased anxiety due to constant notifications, and challenges associated with expectations of continuous availability. One interviewee explained, "I just feel the pressure of, if there's a Slack notification, I have to run back to my desk." These dynamics became more pronounced during COVID-19 due to greater social isolation, extended working hours, and increasingly blurred work-life boundaries.

Second, newswriters still continued to uphold a strong professional identity centred on journalism's democratic mission. Interviewees described their roles as "holding power to account," "amplifying marginalized voices," and "shining a light on things that are messed up," emphasizing that journalism provides "access" to underreported communities and enables citizens to "understand the world around them."

Newswriters identified fundamental contradictions between this mission and the structural realities of platform capitalism. Interviewees highlighted issues such as "collapsing business models," "declining revenues," and "papers closing all over the place," which undermined their ability to fulfill journalism's democratic mission. Platform intermediaries, including Facebook and Google, captured advertising revenues, while algorithmic systems influenced editorial decisions and exacerbated precarity. As one interviewee put it, "A lot more is expected of you, but you're still paid the same lousy rate that you're expected to accept because you're in it for the mission and not because you're being valued for the actual labour you're doing."

Finally, these contradictions produced what this paper terms "professional precarity consciousness": platform newswriters' recognition that individual professional excellence cannot overcome structural exploitation. Newswriters reframed unionization as a means of defending platformed professionalism. One interviewee stated, "Organizing for me was job security, but it was also support in knowing that I can speak up against problems in my workplace ... supporting my journalism and the mission of journalism." Interviewees connected organizing efforts to journalism's democratic mission. One interviewee explained, "Creating this union could open the door for people to be published at [media company] that wouldn't have happened otherwise ... doing this will have more of an impact on the kind of journalism that [we] can do going forward than anything that you individually write."

Future Work

The findings indicate that platform labour generated conditions that shape professional identity into a distinct form of white-collar digital class consciousness among news media workers (Author; Cohen & de Peuter, 2020). These results have significant implications for understanding platformization within the digitally mediated labour force across diverse national contexts. Future research should examine how platform-mediated labour processes and artificial intelligence could affect organizational

dynamics and establish the preconditions for collective organizing in other sectors employing white-collar workers, such as technology and education.

Further studies could also evaluate whether professional precarity consciousness emerges in other contexts where workers maintain strong white-collar and middle-class occupational identities and normative commitments.

References

- Braverman, H. (1998). *Labor and monopoly capital: The degradation of work in the twentieth century*. Monthly Review Press. (Original work published 1974)
- Bunce, M., Wright, K., & Scott, M. (2018). "Our newsroom in the cloud": Slack, virtual newsrooms and journalistic practice. *New Media & Society*, 20(9), 3381–3399.
- Cohen, N. S. (2019). At work in the digital newsroom. *Digital Journalism*, 7(5), 571–591.
- Cohen, N. S. & de Peuter, G. (2020). *New media unions: Organizing digital journalists*. Routledge.
- Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, 5(1), 80–92.
- Fu, A. (2021, July 16). Not just a wave, but a movement: Journalists unionize at record numbers. Poynter. <https://www.poynter.org/business-work/2021/not-just-a-wave-but-a-movement-journalists-unionize-at-record-numbers/>
- Hayes, K., & O'Sullivan, M. (2024). Labouring the news: Management control strategies and work intensification in the digital newsroom. *Journalism*, 25(6), 1346–1364.
- McChesney, R. W., & Nichols, J. (2022). To protect and extend democracy, recreate local news media. Free Press. https://www.freepress.net/sites/default/files/2022-03/to_protect_democracy_recreate_local_news_media_final.pdf
- Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society*, 20(11), 4275–4292.
- Petre, C. (2018). Engineering consent: How the design and marketing of newsroom analytics tools rationalize journalists' labor. *Digital Journalism*, 6(4), 509–527.
- Pickard, V. (2020). *Democracy without journalism? Confronting the misinformation society*. Oxford University Press.
- Poell, T., Nieborg, D. B., & Duffy, B. E. (2023). Spaces of negotiation: Analyzing platform power in the news industry. *Digital Journalism*, 11(8), 1391–1409.

8D.4 Memes as Micro-Archives: Documenting Everyday Life in Digital Culture

Jyoti (IIT Jodhpur)

Introduction

Digital platforms have transformed how everyday life is recorded, circulated, and remembered. Historically, archives have been associated with institutions such as states, museums, and libraries that determine what counts as worthy of preservation. These archives tend to privilege official records, exceptional events, and elite cultural production, often excluding mundane, emotional, and routine aspects of social life. Social media environments, however, have enabled participatory forms of cultural documentation in which users continuously produce and circulate representations of lived experience. Within this ecology, internet memes emerge as a significant yet understudied form of vernacular record. Memes are frequently approached as humorous, fleeting, or trivial digital artefacts, but their repetitive structures, recognisable templates, and rapid circulation allow them to capture shared social situations such as exam stress, workplace fatigue, dating anxieties, family expectations, financial precarity, boredom, and small pleasures. Rather than documenting singular historical events, memes document conditions of everyday life. Through humour, exaggeration, and intertextuality, they transform private feelings into collectively recognisable expressions. In this sense, memes operate as distributed repositories of social meaning. This paper positions memes as micro-archives of the everyday — fragmented, informal, and participatory archives that accumulate through circulation rather than institutional preservation — and argues that meme culture challenges distinctions between serious cultural documentation and trivial online content.

Objectives/Research Questions

This paper investigates how meme culture functions as a form of informal archiving of everyday experience and seeks to expand the concept of the archive beyond institutional frameworks by examining memes as vernacular knowledge systems. It asks how memes document everyday social realities such as work, education, relationships, and emotional life that are often absent from formal archives and mainstream media. The study further explores how repetition, remixing, and template circulation create collective records of shared experience rather than isolated humorous texts, and how memes operate as archives of feeling that preserve affective atmospheres such as anxiety, exhaustion, aspiration, and ironic detachment. It also considers the role of platform structures, including algorithms, visibility metrics, and

sharing practices, in shaping which everyday experiences become widely archived through memes. Finally, the paper addresses the limits of meme archiving by questioning whose experiences become culturally legible and whose remain marginal or invisible. The objective is not to treat memes as factual records but as cultural indicators that reveal what kinds of experiences become socially recognisable and repeatedly articulated in digital public life.

Methods

The study employs a digital ethnographic approach combined with visual semiotic and discursive analysis, treating memes not as isolated images but as socially embedded artefacts within platform environments. The field site consists of public meme pages, recurring meme templates, and their circulation patterns across social media platforms over a defined time frame. Data collection involves screenshot archiving of memes, documentation of captions and textual overlays, recording of metadata such as dates, engagement metrics, and hashtags, and observation of user interactions in comment sections, including tagging practices and statements of identification. Digital ethnography enables attention to context by examining how memes are shared, interpreted, and emotionally received. The analysis integrates semiotic approaches to examine how visual and textual signs construct shared cultural meaning with discursive analysis to identify recurring themes such as exhaustion, aspiration, and relational negotiation across meme clusters. Meaning is treated as emerging through repetition and variation rather than from single instances, allowing the study to understand memes both as texts and as social practices embedded in everyday digital life.

Results

Preliminary analysis indicates that memes consistently function as micro-archives of ordinary social conditions rather than exceptional events. Recurring templates document experiences such as procrastination, workplace alienation, economic anxiety, family negotiations, and digital fatigue, suggesting that memes serve as collective recording mechanisms for shared social pressures. The findings also show that memes operate as archives of feeling by preserving emotional atmospheres, particularly exhaustion, irony, ambivalence, and low-intensity anxiety associated with contemporary life. Humour does not diminish seriousness but enables difficult experiences to become socially shareable. At the same time, visibility within meme archives is uneven. Platform affordances and norms of relatability privilege certain experiences that resonate broadly, while structurally marginal experiences appear less frequently. Meme archives are therefore selective and shaped by digital infrastructures rather than neutral reflections of everyday life.

Future Work

Future research will extend this framework by examining how different platform cultures shape what kinds of everyday experiences are memetically archived, allowing for comparative analysis of digital environments. Longitudinal tracking of meme templates will help study how everyday concerns evolve over time and how digital memory operates through recirculation rather than stable storage. Further integration of critical archival theory will deepen the understanding of memes as counter-archives that challenge institutional hierarchies of cultural value while remaining shaped by algorithmic power. By treating memes as micro-archives, this work aims to contribute to Digital Humanities, social media studies, and cultural theory by redefining what counts as documentation, memory, and knowledge production in the digital age.

Session 8E - Social Media and Platform Infrastructure

Day 2 · Tues 14 July · 13:50-15:20, 225, ARC

8E.1 Misogynoir, Influencer Tactics, and Profit Maximization on YouTube

Areyana Proctor (University of Wisconsin - Madison)

Introduction

Creators in the manosphere – a loose collection of online communities and influencers who espouse misogynistic ideologies (Ging, 2017; Messner, 2016) – often simultaneously engage in racism and white supremacy (Fugardi, 2024; Horta Ribeiro et al., 2021). This has prompted Black men who hold similar misogynistic viewpoints to carve out spaces of their own online, leading to what has been termed the Black manosphere (Onuoha, 2022). Many times, these Black men speak to and directly about Black women, leading to a particularly racialized form of misogyny. In one example from a viral clip with Black manosphere figureheads Myron Gaines and Walter Weekes, who co-host the video podcast "FreshandFit", they refer to Black women as "Shanequas" and say they don't date Black women because they aren't "down with the brown".

Stemming from this example, this paper examines the dehumanization of Black women in digital content produced by both Black and white manosphere influencers. Moya Bailey coined the term misogynoir to name the dehumanization that Black women experience in popular culture and digital media (2021). Scholarship on the manosphere often covers the experiences of white women at the expense of women of color. Although all women are devalued in these antifeminist spaces, women of color are often depicted as less attractive, more sexually available, and inferior to their white counterparts (Fugardi, 2024). Additionally, while scholarship is beginning to explore misogynoir within both Black and white manospheres (Awwal, 2024; Bell, 2024; Bell, 2025), less scholarship on the Black manosphere exists in relation to the white manosphere. I contribute to this emerging conversation by applying a political economic approach towards understanding misogynoir's presence and circulation by both Black and white manosphere creators and within digital platforms more broadly. Although the attention economy promotes inflammatory content in general, racialization's visual logics make misogynoir a potentially more efficient method of

profit accumulation for creators and platforms. Racialized gendered discrimination also serves as a crossover appeal between the Black manosphere and white supremacist associations in the wider manosphere by legitimizing hate against Black women from both white and Black men.

Objectives/Research Questions

I evaluate how misogynoir serves as a branding strategy for manosphere influencers and as an engagement/monetization strategy for both influencers and platforms. My analysis connects Black feminist frameworks with scholarship on digital cultural production, platformization, and influencer studies. I interrogate platform moderation with a focus on demonetization and deplatforming. These forms of moderation impact the risks digital media creators face when attempting to profit from their content, alongside the tactics they employ to mitigate and even capitalize on these risks (Gillespie, 2018). With a focus on YouTube, I also investigate how and why platforms turn a blind eye to discriminatory content, even when it violates community guidelines. Paying attention to the relationship between platformization and influencer tactics illuminates how and why misogynoir may operate as one such tactic within digital media platforms' ever-shifting ecosystems. Overall, this paper asks the following questions: How and why do digital content creators in the manosphere utilize misogynoir as a tactic for virality and profit accumulation? How are platforms implicated in this? How might an exploration of digital misogynoir more attuned to its infrastructural, regulatory, and economic dimensions help expand our understanding of its persistence?

Methods

This paper utilizes a primarily qualitative approach that combines a textual and discourse analysis of 183 videos across three popular manosphere channels on YouTube. I also explore YouTube's negligence and responsibility in the circulation of misogynoiristic content. To do this, I incorporate approaches based on political economy and platformization that take into account YouTube's profit margins, terms of service/community guidelines, and moderation policies.

In my textual and discourse analysis, I watched each video in my case study, making note of interactions with and mentions of Black women. I took note of patterns to understand how misogynoir showed up across this content, returning to Black feminist literature during this analysis to contextualize my findings. I also consulted official YouTube documentation online, including the terms of service and community guidelines, responses to guideline violations, and annual reports.

The three channels encompassing this case study are: "FreshandFit" (Myron Gaines and Walter Weekes), "Kevin Samuels" (Kevin Samuels), and "Woman Propaganda" (Hayden McDougall). All three channels are included based on their varying levels of

popularity and, in the case of "FreshandFit" and "Woman Propaganda", direct associations with each other. For "FreshandFit", I came across two channels: "FreshandFit", with over 1.5 million subscribers, and "FreshandFit Clips", with 748,000 subscribers. These channels are combined for the purposes of this paper. Kevin Samuels, an African American man, is the second creator from the Black manosphere in my case study. Samuels, who passed away in 2022, still has an active account with almost 2 million subscribers. Hayden McDougall, a self-proclaimed alt-right content creator with a YouTube channel titled "Woman Propaganda", is the third creator. His channel is the smallest with around 80,000 subscribers. For all three channels, I sorted their videos based on popularity, setting relative view count ranges as a cut-off point, as each channel has hundreds, if not thousands, of videos. The case study inclusion cut-off point for each channel is as follows: "Kevin Samuels", 1 million views; "FreshandFit", 50,000 views for both channels; "Woman Propaganda", 25,000 views. From there, I further filtered my case study selection based on those that have Black women or women that appear to be mixed-race with Black in their thumbnails, as this serves as an initial site of branding for these videos.

Results

Ultimately, I argue that these creators use misogynoir as a strategy to mitigate the contingency of digital cultural production, creatively finding ways to objectify and dehumanize Black women in the production of their content. In digital content creation's precarious labor environment, the subjugation of Black women continues to be used as a low-hanging fruit to reinforce the capitalist hierarchies that primarily benefit platform companies.

Misogynoir in this content primarily operates through controlling images, particularly the matriarch and Jezebel stereotypes (Collins, 2022). Black women's bodies serve as a repeated site of denigration, especially through fatphobia. As a strategy, influencers often use algospeak, or the practice of intentionally misspelling or substituting words to evade a platform's moderation procedures (Steen et al., 2023). For example, algospeak can be seen when words that would have been flagged by YouTube's algorithmic detection software, such as the n-word slur, are edited out while the context remains clear for human audiences. In doing this, creators aim to mitigate the risk of violating YouTube's community guidelines while still perpetuating misogynoir. It is important to note that "FreshandFit" have been demonetized on YouTube, meaning they can no longer earn advertisement revenue through YouTube's Partner Program. However, this explicitly censored content also incentivizes audiences to go and watch uncensored versions on platforms where the creator's content is still monetized. Whether these influencers are earning advertisement revenue on YouTube, actively promoting sponsored content, or guiding users to platforms where their content is still monetized, they are finding ways to economically capitalize on the shock value of misogynoir.

Despite YouTube's anti-discrimination policy, YouTube and its parent company, Alphabet, still benefit from the presence of misogynoir on the platform. If YouTube were consistent in the implementation of its anti-discrimination policy, then most of this content would have been entirely removed. However, YouTube usually only stops the individual channels from profiting while the platform itself may still earn advertising revenue from the mere existence of this content. This can be seen with YouTube's updated 2020 Right to Monetize section in its Terms of Service, which grants YouTube the right to display ads and monetize content on the videos of creators who aren't in the Partner Program. According to a 2024 report, Alphabet earned \$36.147 billion from YouTube ads (Alphabet, 2024). With this, a violation of guidelines, such as in the case of misogynoir, can serve to benefit Alphabet, as creators can be denied advertising revenue while Alphabet earns profits from this very content. Thus, entirely removing the content would hurt Alphabet's profit maximization, leading to misogynoir's continued existence and perpetuation through these manosphere channels.

Future Work

Future work can include a more expansive analysis of these specific influencers' content, with a particular focus on how their strategies may shift across time and in relation to YouTube's infrastructural changes and moderation procedures. Specific channels in this case study can also be substituted for other channels to understand where similarities and differences may arise in terms of misogynoir's presence. Along these lines, future work can incorporate a larger case study to allow for a more comparative analysis. A similar project can also move to or include other digital media platforms, such as TikTok and Instagram. A cross-platform analysis of this work could prove to be useful. Finally, future work can incorporate audience studies to understand how and why platform users continue to engage with this discriminatory content.

References

Alphabet. (2024). Annual Report 2024.

https://s206.q4cdn.com/479360582/files/doc_financials/2025/q4/2024-annual-report.pdf

Awwal, A. (2024). Understanding the Discourse of the Black Manosphere on YouTube. CSCW Companion, San Jose, Costa Rica, 342-348.

<https://doi.org/10.1145/3678884.3681873>

Bailey, M. (2021). Misogynoir Transformed: Black Women's Digital Resistance. New York University Press.

Bell, D. P. (2024). 'Pick-Me' Black Women: tactical patriarchal femininity in the Black manosphere. Feminist Media Studies, 24(8), 1704-1722.

<https://doi.org/10.1080/14680777.2023.2262163>

- Bell, D. P. (2025). 'He should've killed your ass': Notes on Black women deserving death in the Black manosphere. *European Journal of Cultural Studies*, 0(0), 1-6. <https://doi.org/10.1177/13675494251364727>
- Collins, P. H. (2022). *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. Routledge.
- Fugardi, R. (2024, May 23). How Racist Misogyny Shapes Online Male Supremacist Communities. Southern Poverty Law Center. <https://www.splcenter.org/resources/hatewatch/how-racist-misogyny-shapes-online-male-supremacist-communities/>
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media*. Yale University Press.
- Ging, D. (2017). Alphas, Betas, and Incels: Theorizing the Masculinities of the Manosphere. *Men and Masculinities*, 22(4), 638-657. <https://doi-org.ezproxy.library.wisc.edu/10.1177/1097184X17706401>
- Horta Ribeiro, M., Blackburn, J., Bradlyn, B., De Cristofaro, E., Stringhini, G., Long, S., Greenberg, S., & Zannettou, S. (2021). The Evolution of the Manosphere across the Web. *Proceedings of the International AAAI Conference on Web and Social Media*, 15(1), 196-207. <https://doi.org/10.1609/icwsm.v15i1.18053>
- Messner, MA. (2016). Forks in the Road of Men's Gender Politics: Men's Rights vs Feminist Allies. *International Journal for Crime, Justice, and Social Democracy*, 5(2), 6-20. DOI: 10.5204/ijcjsd.v5i2.301
- Onuoha, A. (2022, January 31). The Manosphere Isn't Just White: Black Femicide and the Radicalisation of Black Male Youth. Global Network on Extremism & Technology. <https://gnet-research.org/2022/01/31/the-manosphere-isnt-just-white-black-femicide-and-the-radicalisation-of-black-male-youth/>
- Steen, E., Yurechko, K., & Klug, D. (2023). You Can (Not) Say What You Want: Using Algospeak to Contest and Evade Algorithmic Content Moderation on TikTok. *Social Media + Society*, 9(3). <https://doi-org.ezproxy.library.wisc.edu/10.1177/20563051231194586>

8E.2 "TikTok... Are You Okay?!" Understanding the U.S. TikTok Ban and Forced Sale through Infrastructural Inversion

Carlos Entrena Serrano (Vrije Universiteit Brussel), Thomas A. Wright (University of Sheffield)

Introduction

On January 19th, 2025, TikTok was banned in the United States for 12 hours, signalling another chapter in the ongoing concerns over alleged user data collection and influence operations being conducted by the Chinese government in the American sociopolitical context. When President Donald Trump signed the September 2025 executive order operationalising TikTok's forced sale in the United States, he joked about "changing the algorithm". Though he was quick to play this off as a joke, this throwaway comment signposted a deeper shift that would soon ensue. Through the enforcement of President Trump's executive order, TikTok's infrastructural autonomy had collapsed under the weight of geopolitical contestation, with its algorithmic opacity reinterpreted through the lens of U.S. state power. What began as a national security measure was soon revealed to be a more structural, cultural shift that saw the assimilation of a once-foreign social media platform into the normative logic of American platform capitalism (Srnicek, 2017).

For millions of users who had relied on TikTok as a space for political expression, community building, and cultural experimentation, these developments were experienced as a loss. Over a year, users took the platform to process, critique and strategise. Framed as more than just an app, TikTok's users expressed fear, anger and frustration at the loss of an infrastructure that facilitated democratic participation, political contestation and the autonomous monetisation of their own creative labour. The TikTok ban is not only part of broader geopolitical contestations between the U.S. and China, but also belongs to a broader conversation of the opacity–'black-box' (Pasquale, 2015)–of social media platforms in general. Platforms do not merely make decisions that directly affect the development of discourse in on- and offline contexts, but they also obfuscate how, when and why these hugely consequential decisions are made. While the US government seems focused on scrutinising foreign-owned platforms like TikTok, American big tech continues to exert significant influence over global public discourse and communication technologies, raising concerns about their unchecked power. In response, efforts to understand and render the opacity of these large transnational organisations more transparent have, in recent years, generated legislative interventions such as the European Union's Digital Services Act (Regulation 2022/2065, 2022) and the AI Act (Regulation 2024/1689, 2024).

The ban's enforcement, swift reprieve, and eventual forced sale of TikTok's US operations are symptomatic of recently re-ignited Sino-American geopolitical tensions, resulting in a unique sociopolitical landscape that has emerged within and around the platform. Amid this uncertainty, users have devised theories that navigate perceived platform censorship, changes to its algorithmic recommendations, and the potential for further disruption.

Objectives/Research Questions

In this article, we examine how TikTok users engaged with and interpreted the ban through videos, exploring how they navigated platform uncertainty and perceived their own agency in response to shifting TikTok's ownership and algorithmic governance (Rouvroy & Berns, 2013).

Rather than treating the 2025 TikTok ban and U.S. buyout as isolated political events, we approach them as moments of infrastructural inversion (Bowker, 1994) that render visible the normally invisible relations among technology, politics, and power. Crises such as this draw public attention to the embedded infrastructure of platforms, exposing how ownership, moderation, and algorithmic design are all implicated within broader geopolitical interests. As TikTok's ownership shifted under U.S. control, what had once been a site of algorithmic otherness—a platform of Chinese origin, marked by distinct cultural practices and user sensibilities—was reconstituted as an extension of American Big Tech. Through this process, we witness the re-hegemonization of the social media landscape and the reassertion of U.S. cultural, political, and technological dominance over the infrastructures that mediate everyday life.

This paper contributes to a nascent body of work that investigates how users relate to algorithms and platforms and enact their agency (Savolainen & Ruckenstein, 2022; Siles, 2023; Siles et al., 2023) how they develop algorithmic folk theories (Karizat et al., 2021; Lin, 2025; Ytre-Arne & Moe, 2021), platform gossip (Bishop, 2019) and engage in algorithmic auditing (Shen et al., 2021). We analyse how TikTok users adapt to platform politics and conceptualise technical affordances by generating novel forms of expression to maintain community and disseminate information in an increasingly volatile and politicised platform environment.

By examining how TikTok users discuss the platform's ban, we can understand how they perceive their own agency in relation to TikTok's platform and algorithmic governance. This methodological approach allows us to study user agency as socially and culturally situated—mediated (Siles et al., 2023). While previous scholars have pointed out how TikTok users perceive the algorithm as good and reliable (Siles, 2023), in this case study, we reveal how users begin to question and distrust TikTok's algorithmic governance in response to the ongoing changes to the sociopolitical landscape within and around the platform. We also incorporate multiple levels of

analysis—examining videos, the broader sociopolitical context and different temporalities—focusing on the conversation before and after the ban.

Methods

Drawing on a corpus of 330 TikTok videos that were collected between December 2024 and October 2025, this article explores how users narrated, theorised, and contested the platform’s transformation during this period. Following our understanding of TikTok’s vernacular, we are aware that (1) many posts do not utilise hashtags, (2) high engagement does not necessarily require a large follower base. These factors informed our data collection strategy, which goes “beyond the hashtag” (Gerrard, 2018). Taking a platform-specific approach, we walked through (Light et al., 2018) TikTok’s interface, leveraging its technical affordances to immerse ourselves in a contemporary conversation, community, and algorithmically curated space.

To collect our dataset, we started from our already existing TikTok accounts, gathering an initial sample of recommended videos related to the TikTok ban. We then created two fresh accounts, where we heavily engaged with the initial set of data - liking, saving, and following creators, as well as searching for keywords and hashtags they might use. Using this initial round of interactions as a first nudge for the platform’s algorithm to refine its recommendations, we explored our feeds, identifying further videos and creators engaged in the conversation.

Our next methodological step was to perform a thematic analysis (still ongoing) (Braun & Clarke, 2019) on the final selected sample of videos and comments, where we explored how users frame and discuss the ban.

Results

At the time of the initial ban in March 2025, the prevailing political narrative framed the move as a defense against Chinese interference and a protection of U.S. citizens’ data. Media coverage framed the issue in terms of privacy and data sovereignty. Yet within TikTok itself, users developed alternative explanations, speculating the ban was less about privacy and more about censorship, and more about consolidating U.S. state control over one of the few remaining digital spaces where young, progressive Americans could discuss politics outside the reach of U.S. corporate platforms.

What emerges from these discussions is not a set of uninformed rumours, but rather a form of vernacular political analysis: the collective attempt to make sense of shifting infrastructures of visibility and control. As state actors publicly suggested that TikTok’s algorithm could be “tuned” to support national interests, these “folk theories” gained retrospective legitimacy. Once dismissed as conspiratorial, they came to mirror realities of algorithmic geopolitics.

Two theoretical implications follow. First, it highlights the methodological potential of infrastructural breakdowns as moments of analytical clarity. In ordinary circumstances, the material infrastructures that underpin social media almost always remain opaque to both users and researchers alike. It is only in these moments of infrastructural breakdown—moments in which systems fail or are forcibly reconfigured—that the contours of power, and the practices and relations that sustain them, become visible. Examining TikTok’s forced sale, and users’ perceptions of this process, provides a rare empirical opportunity to study how geopolitical relations are embedded in, and enacted through social media infrastructures. The platform’s inversion therefore reveals not only how state actors seek to instrumentalise algorithmic systems, but also how users interpret and resist such interventions.

Second, the concept of “folk theory” warrants critical reconsideration. Within media and communication research, folk theories of algorithms are often treated as partial understandings of complex technical systems (Bucher, 2018; Eslami et al., 2015). Yet, our case study demonstrates how these folk theories can also function as legitimate forms of political literacy, different ways of knowing that arise from users’ embedded experiences within platform ecologies and infrastructures. Rather than viewing these articulations as epistemically inferior accounts, we argue that they form vernacular infrastructures of knowledge that both reflect and shape political discourse.

We argue that TikTok’s assimilation into U.S. ownership marks not the end of its political role, but the reassertion of a new infrastructural norm that sees the symbiosis of American Big Tech and the U.S. state as the infrastructural foundation of the contemporary clustered (Gerbaudo, 2024) public sphere.

Future Work

Future research could extend our analysis to other social media platforms, paralleling our methodological approach of using infrastructural inversion moments to understand how users experience, theorise, and resist sociopolitical tensions in diverse digital infrastructures.

References

- Bishop, S. (2019). Managing visibility on YouTube through algorithmic gossip. *New Media & Society*, 21(11–12), 2589–2606.
<https://doi.org/10.1177/1461444819854731>
- Bowker, G. C. (1994). *Science on the Run: Information Management and Industrial Geophysics at Schlumberger, 1920-1940*. MIT Press.
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4), 589–597.
<https://doi.org/10.1080/2159676X.2019.1628806>

- Bucher, T. (2018). *If ... Then: Algorithmic Power and Politics*. Oxford University Press.
- Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., Hamilton, K., & Sandvig, C. (2015). "I always assumed that I wasn't really that close to [her]": Reasoning about Invisible Algorithms in News Feeds. 33rd Annual ACM Conference on Human Factors in Computing Systems, 153–162.
<https://doi.org/10.1145/2702123.2702556>
- Gerbaudo, P. (2024). TikTok and the algorithmic transformation of social media publics: From social networks to social interest clusters. *New Media & Society*, 14614448241304106. <https://doi.org/10.1177/14614448241304106>
- Gerrard, Y. (2018). Beyond the hashtag: Circumventing content moderation on social media. *New Media & Society*, 20(12), 4492–4511.
<https://doi.org/10.1177/1461444818776611>
- Karizat, N., Delmonaco, D., Eslami, M., & Andalibi, N. (2021). Algorithmic Folk Theories and Identity: How TikTok Users Co-Produce Knowledge of Identity and Engage in Algorithmic Resistance. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–44. <https://doi.org/10.1145/3476046>
- Light, B., Burgess, J., & Duguay, S. (2018). The walkthrough method: An approach to the study of apps. *New Media & Society*, 20(3), 881–900.
<https://doi.org/10.1177/1461444816675438>
- Lin, H. (2025). Oscillation Between Resist and to Not? Users' Folk Theories and Resistance to Algorithmic Curation on Douyin. *Social Media + Society*, 11(1), 20563051251313610. <https://doi.org/10.1177/20563051251313610>
- Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.
<https://www.jstor.org/stable/j.ctt13x0hch>
- Regulation 2022/2065 of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and Amending Directive 2000/31/EC (2022). <http://data.europa.eu/eli/reg/2022/2065/oj>
- Regulation 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (2024).
<http://data.europa.eu/eli/reg/2024/1689/oj>
- Rouvroy, A., & Berns, T. (2013). Algorithmic governmentality and prospects of emancipation (E. Libbrecht, Trans.). *Réseaux*, 177(1), 163–196.

- Savolainen, L., & Ruckenstein, M. (2022). Dimensions of autonomy in human–algorithm relations. *New Media & Society*, 14614448221100802.
<https://doi.org/10.1177/14614448221100802>
- Shen, H., DeVos, A., Eslami, M., & Holstein, K. (2021). Everyday algorithm auditing: Understanding the power of everyday users in surfacing harmful algorithmic behaviors. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–29. <https://doi.org/10.1145/3479577>
- Siles, I. (2023). *Living with Algorithms: Agency and User Culture in Costa Rica*. The MIT Press. <https://doi.org/10.7551/mitpress/14966.001.0001>
- Siles, I., Gómez-Cruz, E., & Ricaurte, P. (2023). Fluid agency in relation to algorithms: Tensions, mediations, and transversalities. *Convergence*, 13548565231174586.
<https://doi.org/10.1177/13548565231174586>
- Srnicek, N. (2017). *Platform Capitalism*. John Wiley & Sons.
- Ytre-Arne, B., & Moe, H. (2021). Folk theories of algorithms: Understanding digital irritation. *Media, Culture & Society*, 43(5), 807–824.
<https://doi.org/10.1177/0163443720972314>

8E.3 The Shifting Frontier of Data Access: Research, Restriction, and Rigmarole in the Meta Content Library

Laura Vodden, Axel Bruns (Digital Media Research Centre, QUT)

Introduction

CrowdTangle, a public insights tool once valued for its relative transparency and ease of use, enabled the monitoring and analysis of public content on Facebook and Instagram, accessible via the web, or via the application programming interface (API). Following CrowdTangle's deprecation in August 2024, and its replacement with the Meta Content Library (Meta, 2025c), research into activity on Meta's platforms has become more constrained, with access now limited to a secure, 'cleanroom'-style environment governed by restrictive terms of service. Although social media platform APIs have rarely been intended for academic use (Tromble, 2021), CrowdTangle was an important resource for scholarly research on topics such as the spread of mis- and disinformation (Gruzd et al., 2022; Recuero et al., 2022), news sharing practices (Angus et al., 2023), and political communication during election campaigns (Decker et al., 2024) on Meta's platforms.

Despite its utility, CrowdTangle had several limitations, for example, the lack of an endpoint for gathering comment data. While the inclusion of this functionality represented one advantage of the Meta Content Library, such benefits are overshadowed by the increased restrictions imposed on research into activity across Meta's platforms. In line with a broader trend towards reduced or discontinued API access to social media data (Poudel & Weninger, 2024), researchers wishing to access the Meta Content Library must undergo a lengthy approval process (Meta, 2025b) and navigate numerous platform-imposed technical restrictions, which impact the analytical scope of research on the platform. Tasks that could previously be automated using CrowdTangle's comparatively flexible API, such as tracking trends in engagement over time, or comparing activity across public pages and accounts, now require substantial additional strategic effort, as the Meta Content Library's API quota restrictions, rate and query limits, and mandatory monthly data-deletion policies place practical constraints on the volume of data that can be gathered and analysed within a period of up to one month. Additionally, Meta places strict limitations on what can be imported to and exported from their research environments, making it difficult to replicate analyses and share knowledge between research teams. Regardless, researchers look for ways to work in spite of these limitations, seeking out others working within the Meta Content Library's constraints and sharing experiences and code, although doing so remains challenging.

Objectives/Research Questions

This paper traces our user experience of the Meta Content Library from CrowdTangle's decommissioning to the present. Our objective is to raise awareness of, and stimulate discussion around, the impact of the technical and bureaucratic barriers introduced by Meta's updated access model on social media research. These barriers initially included institutional vetting, increased security and authentication requirements, a lengthy and cumbersome external code and package/model upload process, frequent and sudden changes to the API functionality, out-of-date and conflicting API documentation, and strict limitations and lengthy timeframes for exporting data from the third party-hosted Virtual Desktop Environment (VDE). Further disruption occurred with the (recommended) migration to the Secure Research Environment (SRE), hosted by Meta itself, in late 2025. While the SRE represents a vast improvement to the useability of the system, it came at the cost of a first-of-the-calendar-month complete data deletion policy (Meta, 2025a), and even stricter regulation of data exports. Beyond illustrating the practical challenges of transitioning existing research from CrowdTangle to the Meta Content Library, we situate this development within a broader trend towards the withdrawal of social media data access, and the effects of this on independent scrutiny and scholarly research. By providing a user/researcher perspective and describing the events of the CrowdTangle-to-Meta Content Library transition, we highlight the need for collaboration among approved researchers, in ways that do not breach Meta's terms of service, while recognising that this responsibility falls heavily on individuals and receives little, if any, support from Meta.

Methods

This paper draws on public documentation, as well as email communications from Meta to approved researchers related to release dates, updates, and functional changes to the Meta Content Library. These data were organised to construct a timeline of developments, highlighting key milestones such as API modifications, and policy changes. In addition to documenting release information, we identify recurring pain points and operational challenges encountered by researchers navigating the evolving platform. The very act of documenting and discussing these challenges is itself complicated by the substantial and severe legal and technical restrictions that Meta has placed on researchers, who are prevented, for instance, from taking and showing screen captures of their work in the Meta Content Library VDE or SRE environments, and have only very limited content and code export functionality available to them. We address these restrictions by referring to the information and instructions for users published by Meta, highlighting where this diverges from and conflicts with our own experience of using the Meta Content Library, and presenting our procedural workarounds without including any direct screenshots of the VDE or SRE.

This approach is both descriptive and reflective, combining empirical documentation with a researcher-centered perspective to convey not only the sequence of events but also the operational and methodological implications of these events.

Results

This paper finds that Meta's tightening of access requirements to its platform data has had two primary effects on Meta platform research:

1. Reducing the scope of Meta platform research via the API by imposing strict limits on what and how much data can be gathered and analysed in a one-month time period, and by whom. Journalists and non-governmental organisations (NGOs) who could previously use CrowdTangle are explicitly excluded from using the Meta Content Library. Accredited researchers who do gain approval to work with the Meta Content Library API must respect an asynchronous query rate limit of one query per minute, with the ability to pass a maximum of 250 ids (i.e. user id, post id, etc.) per query, and cannot gather more than 100,000 results per query. Users may gather up to 500,000 posts and 500,000 comments per seven-day rolling window, and all gathered and processed data are deleted systematically on the first day of each month. This limits any analysis to what can be gathered and analysed in a maximum of 30 days. While the SRE allows users to save and re-run queries after the data have been deleted, this counts towards their quota and reduces any additional gathering quota for that month.
2. Siloing platform research by introducing significant barriers to knowledge sharing. Meta's restrictions have limited researchers' ability to collaborate, reproduce findings, and circulate methodological innovations. Nevertheless, researchers have engaged in informal and ad hoc efforts to discuss these constraints, including the sharing of exported and scrubbed Jupyter notebooks through external repositories such as GitHub (e.g. Giglietto, 2026). Code sharing, however, remains substantially constrained by the inability to import external files to the SRE. Instead, to use existing code, researchers must manually transfer the content of a notebook (e.g., copy and paste the contents of each cell) from external repositories into the SRE. Although this can amount to a time-consuming process, it can alleviate the labour involved in developing code within the SRE in the long term, but it does not lend itself to industry-standard version control practices. As a result, knowledge sharing persists, but only in fragmented and inefficient forms that place additional burdens on individual researchers and cannot respond efficiently to ongoing MCL functionality changes.

Future Work

This paper documents and critically examines the impact of the Meta Content Library's access model on social media research from its introduction as a replacement for CrowdTangle to the present, offering a point of reference for researchers navigating the platform's development over the past 18 months. The visibility and public discussion of

these experiences is particularly important in an environment where social media research is becoming increasingly siloed, and where researchers are likely to be addressing similar methodological and technical challenges in isolation. Future work should build on this foundation by further increasing the discoverability of Meta Content Library research practices, workflows, and tools. This is an area that remains within researchers' control despite the many constraints imposed by the platform. Such collective engagement has the potential to ease the impact of the restrictions imposed by Meta on data access and use, to support the development of shared norms and resources for platform research under these constrained conditions, and to exert some greater collective pressure – directly or via research funding and regulatory institutions – on Meta to support the researchers working with its data access platforms.

References

- Angus, D., Bruns, A., Hurcombe, E., Harrington, S., & Tan, X. Y. (Jane). (2023). Computational Communication Methods for Examining Problematic News-Sharing Practices on Facebook at Scale. *Social Media + Society*, 9(3), 20563051231196880. <https://doi.org/10.1177/20563051231196880>
- Decker, H., Angus, D., Bruns, A., Dehghan, E., Matich, P., Tan, J., & Vodden, L. (2024). Topic Diversity in Social Media Campaigning: A Study of the 2022 Australian Federal Election. *Politics and Governance*, 12(0). <https://www.cogitatiopress.com/politicsandgovernance/article/view/8155>
- Giglietto, F. (2026). Fabiogiglietto/mcl-political-reach-study [Jupyter Notebook]. <https://github.com/fabioiglietto/mcl-political-reach-study>
- Gruzd, A., Mai, P., & Soares, F. B. (2022). How coordinated link sharing behavior and partisans' narrative framing fan the spread of COVID-19 misinformation and conspiracy theories. *Social Network Analysis and Mining*, 12(1), 118. <https://doi.org/10.1007/s13278-022-00948-y>
- Meta. (2025a). Data deletion—Meta Content Library and API - Documentation. Meta for Developers. <https://developers.facebook.com/docs/content-library-and-api/content-library-api/guides/data-deletion/>
- Meta. (2025b). Get access—Meta Content Library and API - Documentation. Meta for Developers. <https://developers.facebook.com/docs/content-library-and-api/get-access/>
- Meta. (2025c). Meta Content Library and API | Transparency Center. <https://transparency.meta.com/researchtools/meta-content-library/>
- Poudel, A., & Weninger, T. (2024). Navigating the Post-API Dilemma. *Proceedings*

of the ACM Web Conference 2024, 2476–2484.

<https://doi.org/10.1145/3589334.3645503>

Recuero, R., Soares, F. B., Vinhas, O., Volcan, T., Hüttner, L. R. G., & Silva, V. (2022). Bolsonaro and the Far Right: How Disinformation About COVID-19 Circulates on Facebook in Brazil. *International Journal of Communication*, 16, 24–24.

Tromble, R. (2021). Where Have All the Data Gone? A Critical Reflection on Academic Digital Research in the Post-API Age. *Social Media + Society*, 7(1), 2056305121988929. <https://doi.org/10.1177/2056305121988929>

8E.4 Learning the Algorithm: Algorithmic Grooming and Governance in Social Media Cultures

Nadia Urban (East China Normal University)

Introduction

Across social media platforms, algorithms increasingly govern visibility by ranking, recommending, and moderating what can appear in the feed (Beer, 2017; Bucher, 2018; Gillespie, 2014). While scholarship on algorithmic governance has richly documented the political, economic, and infrastructural dimensions of this process (Gritsenko & Wood, 2022; Just & Latzer, 2017; Katzenbach & Ulbricht, 2019), less attention has been paid to how governance becomes effective through everyday participation. Existing work on algorithmic imaginaries, folk theories, and algorithmic sense-making shows that users actively interpret opaque systems through speculation, discourse, and collective learning (Bucher, 2017; Devito, 2021; Karizat et al., 2021). However, these accounts often treat learning as a matter of knowledge acquisition or literacy. This paper shifts analytical attention from outcomes and understanding to process. It examines how algorithmic governance operates through feedback, ambiguity, and uncertainty, focusing on how users learn to orient themselves toward algorithmic systems despite lacking reliable knowledge or control. Rather than functioning through explicit rules or stable incentives, algorithmic power is shown to operate through anticipatory adaptation. To capture this mechanism, the paper introduces the concept of algorithmic grooming: a process through which users learn to anticipate algorithmic evaluation and adjust their conduct over time, transforming uncertainty itself into a mode of governance.

Objectives/Research Questions

The objective of this paper is to explain how algorithmic governance operates through everyday user interaction under conditions of uncertainty, rather than through explicit rules, stable standards, or transparent control mechanisms. Specifically, the paper seeks to move beyond outcome-oriented accounts of algorithmic power (e.g., bias, removal, amplification) by focusing on the process through which users learn to orient themselves toward algorithmic systems. Fundamentally, this paper has three objectives:

Conceptual: To develop the concept of algorithmic grooming as a mechanism of algorithmic governance that operates through learning, feedback, and uncertainty rather than explicit instruction or stable rules.

Empirical: To examine how users engage in collective efforts to understand, interpret, and adapt to algorithmic systems in environments where outcomes are probabilistic and unpredictable.

Theoretical: To show how algorithmic governance links macro-level structures (platform infrastructures and optimization logics) with micro-level practices of interpretation, affect, and self-regulation.

These objectives are addressed through the following research questions:

How do users learn to orient themselves toward algorithmic systems under conditions of opacity?

What kinds of epistemic and affective labor are involved in attempting to “understand” the algorithm?

How does learning without ability to reliably change outcomes shape user agency, responsibility, and adaptation?

Methods

The study employs a computational–interpretive mixed-methods approach centered on YouTube creator discourse. The dataset consists of creator-generated videos (n=178) discussing platform algorithms, metrics, and visibility, along with associated user comments (n=28,000).

Computational methods are used to map large-scale discursive patterns, including recurring themes related to algorithmic explanation, optimization strategies, and uncertainty. Sentiment analysis is applied to examine affective patterns in audience responses. These quantitative mappings are not treated as ends in themselves but serve as an analytic lens guiding qualitative interpretation.

Qualitative analysis focuses on how creators and audiences articulate understandings of the algorithm, narrate success and failure, and respond affectively to feedback and fluctuating outcomes. This approach allows the study to identify patterned regularities in how algorithmic learning is collectively produced and experienced, while remaining attentive to meaning-making and interpretation.

Results

The findings show that creator communities engage in intensive collective efforts to interpret how algorithmic systems work. Creators exchange explanations, speculate about metrics, and frame visibility outcomes as signals requiring interpretation and adjustment. This reflects a high degree of epistemic agency: users actively attempt to understand and manage algorithmic evaluation.

However, this epistemic effort does not translate into effective agency. Explanations remain provisional, contradictory, and frequently invalidated by changing outcomes. Learning therefore leads to continuous adjustment rather than mastery.

Affective analysis reveals a key asymmetry. Episodic visibility at the level of individual videos is associated with positive affect, while sustained engagement at the channel level correlates with increased negative affect. This suggests that long-term epistemic investment intensifies affective strain rather than providing control or reassurance.

Together, these findings show that learning under algorithmic governance produces a logic of responsabilization: users are positioned as responsible optimizers who must continually adapt to uncertainty, internalize failure, and persist despite the absence of stable rules or guarantees.

Future Work

Future work can extend this analysis to other platform contexts and content domains to examine how algorithmic grooming operates across different sociotechnical environments.

The concept of algorithmic grooming will also be developed further in relation to broader theories of governance, subject formation, and platform power, including its long-term dispositional effects on users' orientations toward risk, value, and legitimacy in algorithmically mediated environments.

References

- Beer, D. (2017). The social power of algorithms. *Information, Communication & Society*, 20(1), 1–13.
- Bishop S (2019) Managing visibility on YouTube through algorithmic gossip. *New Media & Society* 21(11–12): 2589–2606.
- Bucher T (2017) The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. *Information Communication and Society* 20(1). Routledge: 30–44.
- Bucher T (2018) *If...Then*. Oxford University Press.
- Devito MA (2021) Adaptive Folk Theorization as a Path to Algorithmic Literacy on Changing Platforms. *Proceedings of the ACM on Human-Computer Interaction* 5(CSCW2). Association for Computing Machinery.
- Karizat N, Delmonaco D, Eslami M, et al. (2021) Algorithmic Folk Theories and Identity: How TikTok Users Co-Produce Knowledge of Identity and Engage in Algorithmic Resistance. *Proceedings of the ACM on Human-Computer Interaction* 5(CSCW2). Association for Computing Machinery.

- Just, N., & Latzer, M. (2017). Governance by algorithms: Reality construction by algorithmic selection on the Internet. *Media, Culture & Society*, 39(2), 238–258.
- Katzenbach, C., & Ulbricht, L. (2019). Algorithmic governance. *Internet Policy Review*, 8(4).

Session 8F - **PANEL: Performing Platforms: Bodily Performances and Hashtag Cultures in/beyond Performance Studies**

Day 2 · Tues 14 July · 13:50–15:20 · Seminar Suite C, ARC

Miya Shaffer (York University, Toronto), Archer Porter (University of Texas), Ali Na, Eric Chalfant (Queens University)

This panel explores the conceptual and political possibilities that the field of performance studies produces within analyses of social media. As a crucial “stage” for many scholars and practitioners of dance, theater and performance, social media platforms document the theatricalities of everyday life where artists engage their followings to promote live performances or engage dance to activate broad participation and interaction across TikTok and Instagram. Social media cultivates new forms of performance that respond to platform ecologies. Beyond merely analyzing displays of performance, our panel takes up acts of circulation, demonstrating how thinking with performance initiates inquiry into processes of transmission between digital and material worlds, between the assumed containment of the platform and of the theatrical stage. Performance, we suggest, already emphasizes the exchange of ideas between bodies in rehearsal or the sharing of affects between stage and spectator, complicating the “boundary work” that tends to pervade social media practices (Gal 2018).

We specifically analyze how social media (re-)circulates histories of bodily capture and categorization, such as those relevant to race, empire, and their afterlives in contemporary “post-panopticism” (Boyne 2000) . Throughout three distinct yet thematically-related paper presentations from four scholar-artists, our panel demonstrates how performance studies and social media become, in presenters Ali Na and Eric Chalfant’s terms, mutual “kindling” for addressing the continued circulation of racial and imperial histories and examining resistance in both its embodied and digital forms. In sum, we foreground the following objectives: first, we center the body and its activity as sites of political knowledge that can critique and reimagine the body’s capture through social media surveillance and its imagined disembodiment through digital practices like hashtag activism. Second, we demonstrate how frameworks and methods from performance studies, such as choreographic analysis and practice-as-research, can illuminate how power dynamics are continuously lived and shared.

Discourses on social media tend to treat movement and performance as incidental to the networked mediation of content. By foregrounding the body and its choreographies, this panel attends to the nuances of production and the politics of capture that imbue

social media archives. More specifically, the panel engages J. L. Austin's theory of performativity as a conceptual framework for physically and aurally activating—for body-ing forth—social media hashtag cultures and complicating the choreographies therein. Extending scholarly research on the “digital performative,” each panelist attends to critical qualities of how performativity acts in digital culture (Leeker, Schipper, Beyes 2018; Sullivan 2018; Baptiste et al. 2022). Dr. Miya Shaffer analyzes the performativity of racial politics through mixed-race facial reveal videos on TikTok, via hashtags like #wasiancheck. Dr. Archer Porter unpacks the performativity of surveillance by interrogating #caughtdancing videos on Instagram. Drs. Ali Na and Eric Chalfant examine the performativity of the hashtag itself, particularly through the study of #BLM and its transhistorical aggregation of bodies, ideas, and modes of resistance. In each of these instances, the notion of performativity becomes a useful analytic for unpacking the body's relationship to systems of capture, transmission, and networked life. The three papers mobilize performativity not only to reveal how bodies perform on social media, but also how platforms themselves perform, exploring what the body's physicality affords to such performative acts.

In the first presentation, “Moving Face, Choreography Race: TikTok's Facial Choreographies and Resurgent Race Science,” Dr. Shaffer examines the performativity of the human face on TikTok through its performances of racial identification and perception. Video trends concerned with racial self-identification have proliferated since TikTok's early days in the United States. Recent scholarship emphasizes the visibility of self-identified “mixed-race” or multiracial TikTok creators, reflecting apparent demographic changes in the US where populations of self-reported mixed-race people have increased 276% from 2010 (Waring 2023). Scholars have explored how TikTok trends such as “#mixedgirlcheck” or “#wasiancheck” enable opportunities to authenticate racial identity and form digital community, while at the same time homogenizing self-identified multiracial populations as uniformly heterosexual, middle-class, and (Eurocentrically) racially “ambiguous” in appearance (King-O'Riain 2021; Weiß et al. 2025). Although most of this scholarship argues for the centrality of the creator's body to such identity-based video content, few examine the body's choreographic arrangements, such as its physical gestures, sequences of facial expressions, and overall organization throughout the video's temporal duration and spatial framing. Highlighting these attributes, Shaffer's presentation emphasizes the performativity of the video creator's face throughout this style of performance: TikTok's faces become performative apparatuses wherein repetition of gestures, behaviors, and utterances incite action and bring into being politicized worldviews.

Focusing on TikTok's #wasiancheck videos, which were popular on the app between 2020-2022, Shaffer explores how the trend's primary choreography reiterates colonial taxonomies of racial difference and their extension into contemporary resurgences of race science and eugenics. The presentation offers a close-reading of a performance of

facial reveal: several #wasiancheck videos similarly feature the video creator covering and then revealing their face upon a sonic command to “wasian check,” or confirm mixed-racial heritage for the TikTok audience. Engaging dance scholar Sherril Dodds’s (2024) concept of the “choreographic interface” which theorizes the face as a system of meaning-laden codes that necessarily interact with the racialized, gendered, disability-based associations of the body, Shaffer shows how #wasiancheck’s facial conceal-and-reveal produces the face as an external symbol of racialized truth. As such, these videos revive and recirculate histories of physiognomy—the belief that inner desires and character traits manifest in facial features and expressions—and, more importantly, their historic and continued support of beliefs in genetic race and eugenics.

Such an imagining of the face within TikTok’s #wasiancheck trend illuminates, first, the more playful, ordinary iterations of an increasingly popular interest in genetic race within US racial imaginaries and, second, the significant role of the mixed-race Asian-white figure in these resurgent ideas. Revived eugenic notions are especially visible within some factions of US political conservatism and the “tech-right” ideologies associated with the technology industries of Silicon Valley, wherein Asian-white “genetics” are deemed indicative of greater intelligence and beauty (Gebru and Torres 2024; Say Chan 2025). Shaffer’s presentation demonstrates the parallel performance style of TikTok’s #wasiancheck with what she calls the platform’s “optimized” faces, or those demonstrative of physical improvement via “glowing up” and “looksmaxxing.” Through this confluence of facial performances, Shaffer argues that TikTok’s faces mobilize a key tension between neoliberal multiculturalist emphases on self-expression and facial (racial) fluidity versus renewed ideas about genetic essence. The presentation shows how it is specifically the interface of choreography, technological histories, and performance spectatorship through which this tension is illuminated.

The panel’s second presentation, Dr. Porter’s “Candid Dancing and the Aesthetics of Surveillance,” explores a niche category of videos on Instagram. Sometimes tagged #caughtdancing, these videos capture the figure of the unknowing dancer. As short clips typically filmed without the dancer’s knowledge or consent, viewers can observe people moving spontaneously with abandon, seemingly untouched by the capturing forces of the gaze. Though marginal in volume compared to dominant dance trends on Instagram, these “candid dancing” videos are a telling site for understanding the broader conditions of visibility, intimacy, and surveillance that shape embodied experience in a networked life. They transform mundane gestures into cultural texts that aestheticize and circulate embodied privacy, tethering the seemingly unmediated body to social media’s visual economy.

Porter argues that candid dancing videos exemplify a surveillance aesthetic—a visual and affective logic through which social media users aestheticize and normalize acts of

watching and being watched (Finn 2013). The conceit of the “candid” frame hinges on presumed authenticity, a message that is made complete by what dance studies scholar Doran George (2020) calls the “natural body,” or “a representational construct: an image of bodily ease and integration that masked its own social, historical, and institutional production” (pp#). Just as George’s natural body is constructed by devices of visual rhetoric, so too is the authenticity it represents on social media (Taylor 2022). Drawing upon J. L. Austin’s (1962) initial formulation of performativity and Rebecca Scheider’s (2011) reframing of it, this paper proposes that the candid dancing body on social media functions as a form of inverse performativity: its publicization and online circulation negates its innate qualities of invisibility, interiority, and privacy. These media then acquire a particular affective dimension through their online circulation—what is humiliating becomes humorous. The entertaining aspect of such a video then depends on its double bind: the dancer’s unawareness appears as a site of innocence and abandon, even as the camera disciplines, frames, and renders that body into consumable data. In this way, candid dancing mediates the paradox of being both expressive and exposed, spontaneous and surveilled.

Drawing from dance studies, performance studies, and surveillance studies, Porter analyzes a corpus of candid dancing videos posted between 2016 and 2022 on Instagram. As one portion of a larger theory on the aestheticization of intimacy on social media, she argues that candid dancing videos enact a tension between fetishization and discipline, making visible the pleasures of capture while disguising the ethical complexities of consent and the mediatization of “social surveillance” (Marwick 2012). In this regard, privacy is not just political, as Marwick (2024) argues, it is also choreographed.

Ultimately, this presentation contends that candid dancing exposes how surveillance has become an aesthetic condition of everyday life. These videos transform gestures of unselfconscious movement into affective content, rendering privacy not as a withdrawal from visibility but as its performance. The pleasure they elicit on social media—through humor, relatability, and the spectacle of spontaneity—depends on the very infrastructures that compromise the autonomy they appear to celebrate. By turning moments of mundane embodiment into scenes of fetishization and discipline, candid dancing reveals how digital culture choreographs the boundaries between expression and exposure, freedom and capture, intimacy and spectacle.

In the final presentation, “Performing Hashtag Afterlives: Decolonial Geographies of Social Media Protest,” Drs. Na and Chalfant track hashtag afterlives to break down barriers between digital and site-specific protests. They explore the “Staircase of the Colonies” in Marseille, France as a site of #GeorgeFloyd, #JusticeForGeorgeFloyd, #ICan’tBreathe, and #BlackLivesMatter. Seemingly unrelated to George Floyd’s murder in Minneapolis, the French staircase does not conjure the same geographies and

sedimented histories but rather kindles the shared impetus of decolonial placemaking facilitated by social media. How do transnational social media afterlives animate site-specific colonial agitation? In asking this, Na and Chalfant build on scholarship that highlights hashtag activism's capacity for meaningful political engagement and counterpublic formation. However, they also pivot from examining how hashtags can create new online spaces for shared politics to examine how hashtag activism spills over into non-Internet spaces (Halupka 2018; Li, Bernard, and Luczak-Roesch 2021; Penney and Dadas 2013; Wonneberger 2021).

In Marseilles, 2020 #BLM protestors sparked a transnational duration by setting the Botinelly sculpture *Colonies d'Afrique* – depicting a nude African woman reclining on the spoils of African colonization – awash in red paint, writing “Les colonisés d'Afrique niquent la France” (the colonized people of Africa fuck/screw France) (Payan 2020). Presencing the sexual violence inherent in Empire, #BLM protestors upset the colonial fantasy of the taxidermic monuments. In response, “Les Sculptures n'étaient pas blanches” by Mohammed Laouli (Rabat, Morocco) washes the paint from *Colonies d'Afrique*. Laouli dresses in the blue coveralls worn by Moroccan immigrants working in Marseilles. Following the eruption of #BLM rage, Laouli's cloth and water provide compassion to those who bear the legacy of French colonization in Africa. Na and Chalfant analyze Laouli and their own decolonial and anti-ableist 2024 site-specific performance at the staircase to dwell in the chain reactions of hashtag activism.

Performance, Na and Chalfant contend, is critical to re-evaluating the relationship between site-specific performance and hashtag activism. This entails rearticulating #BLM as performative—not to connote artifice as with #PerformativeActivism, #PerformativeMale, or the ubiquitous #Performative, but rather in the sense of an Austinian performative or Butlerian performativity (Austin 1962; Butler 1988; Butler 1993). This move not only helps contextualize the definition of performative but upsets the dichotomy between performance and authenticity—a division at the heart of hashtag activism judgement. For example, in an analysis of #BlackLivesMatter activism, Diana Zulli identifies liveness (e.g. a response to an event of anti-Black violence) as diluting political efficacy, lowering affective attunement, and eliding systemic roots (Zulli 2020, 205). A focus on performance thus helps reconceptualize the Staircase of Colonies as a social media platform operating as a site of content creation, social networking, and interactivity. By entangling the rhetorics of social media with performance, the fields cross-pollinate in a way that refuses the limits of digital platforms. When understood as both platform and site-specific liveness, the staircase alleviates reservations about hashtag activism by tracking long-term transnational political engagement at an affectively charged site, revealing how social media activism participates in confrontations of the legacies of Colonialism that undergird contemporary institutions of State power.

Though distinct in their deployments of performativity, each of these papers articulates one productive framework that bridges performance studies and social media studies, theorizing towards overlapping analytics for both fields. Emphasizing the hashtag as an affordance that initiates activity—that does something—this panel’s relevance lies in its call for revising colloquial approaches to the performative that are apparent throughout ordinary social media usage. By returning the performative to the body and its gestures, as Austin, Butler, and other scholars have theorized, the presenters shift focus from the associations of inauthenticity towards the promises of generativity, towards the “doing” of networked bodies and collectives.

Session 9A - Platform Governance and Regulation (3)

Day 2 · Tues 14 July · 15:35-17:05, RD Suite, ARC

9A.1 The Day-Job Paradox: AI Vulnerabilities of Independent Musicians

Travis Lloyd (Cornell Tech)

Introduction

AI systems can already create content across various mediums—text, sound, still images and video—that previously required skilled creative labor to produce. While these technologies may never eliminate the need for all creative workers, they are already “good enough” to perform certain tasks that historically provided employment opportunities in the cultural industries. Research from both the social sciences and human-centered computing has begun to explore how new AI technologies will impact creative work and workers, focusing on how workers may use these technologies (Caramiaux & Fdili Alaoui, 2022; Galuszka, 2025; Palani & Ramos, 2024; Shelby et al., 2024; Supper, 2025), how workers are responding to them individually and collectively (Grohmann et al., 2025; Halperin & Rosner, 2025; Kaye et al., 2025), and how they may contribute to inequality between workers (Jiang et al., 2023; Kyi et al., 2025; Mateescu et al., 2025). Understanding how AI technology will shift the dynamics of cultural production requires examining the unique challenges and vulnerabilities of distinct groups within this broader field. To this end, this paper focuses on the potential impacts of AI on one specific group of workers: independent musicians within the US music industry.

Objectives/Research Questions

I situate the US music industry as a cultural industry, using Hesmondhalgh’s (2018) term for an industry focused on the creation and dissemination of communicative texts. My conception of the music industry additionally includes self-produced or self-recorded music, even though this activity may challenge the notion of a formal “industry.” While this industry contains many different types of workers—including symphony performers, top-40 pop stars, sound engineers, and social media managers—I focus solely on musicians. To define this group, I take inspiration from Baym’s (2018) study of musicians in the era of social media. Baym acknowledges the “fuzzy” nature of

the “musician” category, given that anyone can produce music, and chooses to focus her study “on career (or would-be career) music workers” (Baym, 13). I do so as well, noting that the difficulty of making a viable living as a musician requires many to rely on other sources of income, complicating simple notions of amateur and professional. Even within the category of musician, there are many subgroups whose different skills, incentives, and income sources make them vulnerable to AI in different ways and to different extents. Part of my contribution is to make these differences salient, in the hopes of illuminating how the impacts from AI technologies will be unevenly distributed, even within a single industry and category of worker. I specifically identify the challenges faced by independent musicians, or musicians who produce recorded music but do not have record deals with one of the “big three” major record labels (Sony Music, Universal Music Group, and Warner Music Group).

Methods

I base my analysis on original empirical research exploring the AI-related concerns around which US musician advocacy organizations are mobilizing and the potential solutions these groups envision. While I also speak with policy groups and unaffiliated musicians, I primarily focus on advocacy groups with musician membership, in order to best understand the equity implications for the workers who contribute to this form of cultural production. Data collection is ongoing, but to date I have completed seventeen semi-structured interviews (spanning five advocacy organizations), conducted eleven months of participant observation, and analyzed the public media produced by three labor organizations that represent working musicians: United Musicians and Allied Workers (UMAW), Music Workers Alliance (MWA), and the American Federation of Musicians (AFM). The AFM is a traditional labor union representing a broad swath of professional musicians, while UMAW and MWA are advocacy organizations specifically focused on independent musicians. I have attended member meetings, public discussions, and protests held by these three organizations, and I am most embedded within UMAW: I have become a member, attend weekly working group meetings, and have participated in several public demonstrations. Outside of my professional identity as a researcher, I am also a working independent musician, which gives me unique access to field sites and participants.

Results

Through this empirical research, I find that many responses to AI currently being discussed within policy spheres, such as copyright protections and licensing deals, do not adequately address the concerns of independent musicians. These musicians are extremely vulnerable to AI for three reasons: their pre-existing economic precarity, their lack of access to legal and organizational protections, and their reliance on side gigs, separate from their primary creative output, to make a living. First and foremost, these musicians find the current platform-mediated streaming economy (Poell et al., 2021)

unsustainable and exploitative, and believe that AI will only exacerbate these problems. Already unable to make a living off streaming revenue, they fear that competition with AI-generated music will drive their earnings even lower. Their primary response is thus not AI-specific but a broader attempt to reform the economics of streaming. Second, remedies that may protect musicians working in more formal arrangements, such as copyright protections and union-negotiated licensing contracts, are unlikely to help independent musicians who are either excluded from these negotiations or operate at too small of a scale to generate meaningful revenue from intellectual property rights. Finally, these workers rely on sustenance income from many different side gigs to fund their primary creative endeavors. This work is often within the music industry and makes use of specific skills and knowledge: for example, making background music for advertisements or producing other musicians' recordings. Independent musicians are not concerned that AI will decrease the appetite for their primary creative output but fear that these side-gigs are the kinds of jobs that will be automated away.

Paradoxically, while AI tools claim to make musicians' lives easier by automating the “non-creative” parts of cultural production, I find that they can actually make their lives more difficult by eliminating sources of sustenance income. I dub this the “day-job paradox” and suggest that it may be a mechanism for inequality in other creative industries as well.

Future Work

Data collection through Interviews and fieldwork will continue over the coming months, with completion planned prior to the Social Media & Society conference in July. Building on this work, I have two future studies planned. First, I am conducting a parallel study examining AI-generated music from a platform governance perspective, using interviews with musicians and listeners to understand their perceptions of streaming platform policies towards AI-generated content (Lloyd et al., 2025; Schaffner et al., 2024). Second, I plan to leverage my relationships with UMAW and MWA to launch a data-driven advocacy project (Ming et al., 2025). This collaborative effort will co-design tools that use streaming platform data to support independent musicians' advocacy efforts. Specifically, I aim to develop a computational tool capable of identifying AI-generated music and artists on streaming platforms, enabling advocates to assess the scope of impact of such content, and to mobilize fans in response.

References

- Baym, N. (2018, July). *Playing to the Crowd*. NYU Press.
<https://nyupress.org/9781479821587/playing-to-the-crowd/>
- Caramiaux, B., & Fdili Alaoui, S. (2022). “Explorers of Unknown Planets”: Practices and Politics of Artificial Intelligence in Visual Arts. *Proc. ACM Hum.-Comput. Interact.*, 6(CSCW2), 477:1-477:24. <https://doi.org/10.1145/3555578>

- Galuszka, P. (2025). The influence of generative AI on popular music: Fan productions and the reimagination of iconic voices. *Media, Culture & Society*, 47(3), 603–612. <https://doi.org/10.1177/01634437241282382>
- Grohmann, R., Rocha, A. C., & Guilherme, G. (2025). Worker-led AI governance: Hollywood writers' strikes and the worker power. *Information, Communication & Society*, 0(0), 1–19. <https://doi.org/10.1080/1369118X.2025.2521375>
- Halperin, B. A., & Rosner, D. K. (2025). "AI is Soulless": Hollywood Film Workers' Strike and Emerging Perceptions of Generative Cinema. *ACM Transactions on Computer-Human Interaction*, 32(2), 1–27. <https://doi.org/10.1145/3716135>
- Hesmondhalgh, D. (2018). *The Cultural Industries*. SAGE.
- Jiang, H. H., Brown, L., Cheng, J., Khan, M., Gupta, A., Workman, D., Hanna, A., Flowers, J., & Gebru, T. (2023). AI Art and its Impact on Artists. *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 363–374. <https://doi.org/10.1145/3600211.3604681>
- Kaye, D. B. V., Pappalardo, K., & Gray, J. (2025). In the style of: Exploring industry, creator and legal implications of copying style through generative AI. *Information, Communication & Society*, 0(0), 1–17. <https://doi.org/10.1080/1369118X.2025.2496905>
- Kyi, L., Mahuli, A., Silberman, M. S., Binns, R., Zhao, J., & Biega, A. J. (2025). Governance of Generative AI in Creative Work: Consent, Credit, Compensation, and Beyond. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–16. <https://doi.org/10.1145/3706598.3713799>
- Lloyd, T., Gosciak, J., Nguyen, T., & Naaman, M. (2025). AI Rules? Characterizing Reddit Community Policies Towards AI-Generated Content. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, 1–19. <https://doi.org/10.1145/3706598.3713292>
- Mateescu, A., West, Z., & Pinto, S. (2025). Fashion's Data Doubles: How AI is Reshaping Modeling Work. <https://hdl.handle.net/1813/117716>
- Ming, J., Tolera, H. H., Tu, J., Yitzhaki, E., Ngai, C. S. E., Sterling, M., Avgar, A. C., Vashistha, A., & Dell, N. (2025). Exploring Data-Driven Advocacy in Home Health Care Work. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, 1–17. <https://doi.org/10.1145/3706598.3713086>
- Palani, S., & Ramos, G. (2024). Evolving Roles and Workflows of Creative Practitioners in the Age of Generative AI. *Proceedings of the 16th Conference on Creativity & Cognition*, 170–184. <https://doi.org/10.1145/3635636.3656190>

- Poell, T., Nieborg, D. B., & Duffy, B. E. (2021). *Platforms and cultural production*. John Wiley & Sons.
- Schaffner, B., Bhagoji, A. N., Cheng, S., Mei, J., Shen, J. L., Wang, G., Chetty, M., Feamster, N., Lakier, G., & Tan, C. (2024). “Community Guidelines Make this the Best Party on the Internet”: An In-Depth Study of Online Platforms’ Content Moderation Policies. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, CHI ’24*, 1–16.
<https://doi.org/10.1145/3613904.3642333>
- Shelby, R., Rismeni, S., & Rostamzadeh, N. (2024). Generative AI in Creative Practice: ML-Artist Folk Theories of T2I Use, Harm, and Harm-Reduction. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–17.
<https://doi.org/10.1145/3613904.3642461>
- Supper, A. (2025). “More Where That Came From”: How the Voice of Randy Travis Helps to Normalize AI. *Popular Music and Society*, 48(4), 397–415.
<https://doi.org/10.1080/03007766.2025.2492500>

9A.2 Governance over hateful ideology: A longitudinal analysis of TikTok's hate speech governance in 2018 - 2025

Yang Xu (University of Helsinki)

Introduction

Hate speech governance on social media platforms remains a contested and evolving domain shaped by legal ambiguity, platform power, and sociotechnical constraints. Despite growing international and regional regulatory efforts—such as the EU Code of Conduct on Countering Illegal Hate Speech Online and UNESCO's Guidelines for the Governance of Digital Platforms—there is no universally agreed-upon definition of hate speech, leaving digital platforms with significant discretion in defining and enforcing their own standards (Sellars, 2016; Gillespie, 2018).

Prior research shows that platforms' hate speech policies have expanded in scope and complexity over time, reflecting increasing regulatory pressure, public scrutiny, and a broader “responsibility turn” in platform governance (Dubois & Reepschlager, 2024; Katzenbach, 2021). However, these developments have also raised concerns about opacity, over-moderation, and disproportionate impacts on marginalized communities (Cavaliere, 2019; Smolarz et al., 2024).

This study focuses on TikTok, a platform whose governance of hate speech has evolved rapidly since its global rise in 2018. Notably, TikTok introduced “hateful ideology” as a distinct regulatory category in 2020, framing certain belief systems as incompatible with platform values. While this move appears to signal a more proactive stance against structural forms of hate, empirical research suggests that TikTok's moderation practices often remain inconsistent, ineffective at protecting vulnerable groups, and at times disproportionately restrictive (Lookingbill & Le, 2024; Tibau et al., 2025). Therefore, this paper examines how TikTok has conceptualized and operationalized hate speech—and particularly “hateful ideology”—over time, and what this reveals about the structural limits of TikTok's platform governance.

Objectives/Research Questions

This study aims to examine the evolution of TikTok's hate speech governance from 2018 to 2025, with a particular focus on the emergence and consolidation of “hateful ideology” as a key regulatory concept. It addresses the following research questions: 1) How has TikTok's conceptualization of online hate evolved from 2018 to 2025? 2) How has the notion of “hateful ideology” been defined, expanded, and operationalized in TikTok's hate speech policy since its introduction in 2020? 3) What structural and epistemic limitations of platform governance are revealed through TikTok's evolving

approach to hate speech? By answering these questions, the study contributes to scholarship on platform governance, hate speech regulation, and the role of ideology in content moderation, offering a platform-specific, longitudinal perspective that complements comparative policy studies.

Methods

The study employs qualitative content analysis to examine TikTok's hate speech governance over time (Barun and Clarke, 2006; Drisko & Maschi, 2015). The dataset consists of historical versions of TikTok's Community Guidelines, transparency reports, and public newsroom communications published between 2018 and 2025. Community Guidelines were selected as the primary policy artifact because they represent TikTok's most comprehensive and public articulation of content moderation rules and platform values (Gillespie, 2018; Chan et al., 2023). Supplementary materials—including transparency reports and newsroom announcements from the United States, European Union, United Kingdom, and Australia—were analyzed to contextualize policy changes and capture governance discourse beyond formal rule texts.

In total, seven versions of Community Guidelines (2018–2025) were analyzed alongside relevant transparency reports and public statements identified through keyword searches (“hate,” “hate speech,” “hateful”). Documents were coded longitudinally to trace changes in definitions, policy structure, enforcement logic, and references to ideology.

Analytical attention was given to (1) shifts in how hate speech is defined, (2) the role assigned to ideology in explaining and governing harmful content, and (3) changes in enforcement mechanisms, particularly the transition from content removal to visibility-based moderation.

Results

The analysis identifies two major shifts in TikTok's hate speech governance: 1) In 2020, TikTok moved from relatively ad hoc prohibitions to a more structured framework that conceptualized hate speech through multiple mechanisms, including direct attacks, slurs, hateful organizations, and—crucially—hateful ideology. Initially framed vaguely, hateful ideology targeted belief systems deemed incompatible with an “inclusive and supportive” community, enabling the platform to act against coded language, symbols, and off-platform ideological movements. 2) Second, from 2023 onward, TikTok restructured its governance architecture by introducing a tiered, visibility-based enforcement system (e.g., For You Feed ineligibility standards). This shift reframed hate speech governance from a binary moderation problem to a platform-design challenge centered on managing content reach—aligning with TikTok's broader model of visibility moderation. Across this period, “hateful ideology” evolved from a sub-category into a central organizing concept that subsumed hate speech, hateful behavior, and extremist

content under a single framework. The concept served multiple governance functions: enabling pre-emptive moderation, facilitating flexible responses to crises (e.g., armed conflicts), and positioning TikTok as a morally responsible actor under regulatory scrutiny.

Conceptually, TikTok's use of hateful ideology reflects a tension between a belief-centered (cognitivist) definition of ideology and a practice-oriented operationalization focused on symbols, discourse, and social meaning. This misalignment mirrors—but also departs from—the legal distinction between *forum internum* and *forum externum*, revealing how platforms reinterpret human rights principles within proprietary governance systems.

Future Work

Future research can further examine the normative implications of governing hate through ideology-based frameworks. In particular, the study points to three areas for future work:

First, TikTok's single-axis, attribute-based conception of hateful ideology may be ill-suited to addressing intersectional forms of hate, suggesting a structural mismatch between scalable moderation systems and the complexity of social harm.

Second, the analysis raises questions about over-moderation and epistemic limits, as abstract ideological categories risk becoming performative “boundary objects” that enable governance flexibility while obscuring accountability.

Third, the findings invite comparative analysis across platforms and legal regimes to assess whether ideology-based governance reflects a broader trend in platform regulation or a TikTok-specific response shaped by geopolitical, commercial, and regulatory pressures. More broadly, the study suggests that the limitations observed are not accidental flaws but structural features of contemporary platform governance, highlighting the need for regulatory and conceptual innovations beyond existing content moderation paradigms.

References

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
<https://doi.org/10.1191/1478088706qp0630a>
- Cavaliere, P. (2019). Digital platforms and the rise of global regulation of hate speech. *Cambridge International Law Journal*, 8(2), 282–304.
- Chan, N. K., Su, C. C., & Shore, A. (2023). Shifting platform values in community guidelines: Examining the evolution of TikTok's governance frameworks. *New*

Media & Society, 14614448231189476.

<https://doi.org/10.1177/14614448231189476>

Dubois, E., & Reepschlager, A. (2024). How harassment and hate speech policies have changed over time: Comparing Facebook, Twitter, and Reddit (2005-2020). *Policy & Internet*, 16(3), 523–542.

Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.

Haslanger, S. (2017). Racism, ideology, and social movements. *Res Philosophica*, 94(1), 1–22.

Sellers, A. (2016). *Defining Hate Speech* (SSRN Scholarly Paper No. 2882244). Social Science Research Network. <https://doi.org/10.2139/ssrn.2882244>

Smolarz, A. V., Vinther-Jensen, E., Ikanović, E. L., & Bornakke, T. (2024). Preventing “Torrents of Hate” or Stifling Free Expression Online?: An Assessment of Social Media Content Removal in France, Germany, and Sweden. *The Future of Free Speech*. <https://justitia-int.org/en/14062/>

Zeng, J., & Kaye, D. B. V. (2022). From content moderation to visibility moderation: A case study of platform governance on TikTok. *Policy & Internet*, 14(1), 79–95. <https://doi.org/10.1002/poi3.287>

9A.3 Who Should Be in Charge? Safety and Agency in the Public Imagination of EU Platform Governance

Claire Stravato Emes (SciencesPo)

Introduction

Public discourse on platform governance has shifted from the early optimism of “connective social media” to a diagnosis of decay and extraction. This shift is captured by “enshittification,” a term popularized by Cory Doctorow to describe how platforms, after securing users and business partners, progressively degrade the user experience as they intensify value extraction (Doctorow, 2023). The term has also become a widely legible cultural diagnosis: in 2024, Macquarie Dictionary selected enshittification as Word of the Year, explicitly framing it as capturing a deterioration “many of us feel” in the services we depend on (Shepherd, 2024).

At the same time, the platform economy is entering a new phase in which AI-mediated interfaces increasingly promise “seamless” decisions and automated assistance. Recent infrastructure forecasts suggest massive capital amounts for AI-based applications, reinforcing a sense that AI is reorganizing the conditions of digital intermediation (Noffsinger et al., 2025). The launch of “agentic commerce” standards such as Google’s Universal Commerce Protocol (UCP) is presented as an open standard for agentic commerce journeys across consumer surfaces, retailers, and payment providers, potentially shifting more decision and transaction moments into AI-mediated interfaces (Google, 2026a, 2026b). These developments sharpen a central governance dilemma: should regulation primarily protect users from harm in adversarial environments, or should it build capability, the practical conditions for choice, exit, and contestation?

One recurring orientation is protection, framing users as vulnerable subjects (especially children) and legitimating harm-reduction governance through age gating, restrictions, and safety-by-design obligations. The youth mental health debate is a key site where this protective posture becomes politically salient (Haidt, 2024). A second orientation is capability (or agency): governance should enable users to act meaningfully in digital environments by lowering the cognitive and procedural costs of control. danah boyd’s distinction between “risks” and “harms” captures this move, emphasizing governance that supports navigation and agency rather than simple exclusion (boyd, 2024). The EU stack embodies both imaginaries. The Digital Services Act (DSA) is routinely narrated as an “online safety” rulebook, but it also contains explicit capability hooks: Article 25 restricts interface designs that “materially distort or impair” users’ ability to make “free and informed decisions” (European Union, 2022). Yet whether these

capability-facing elements are publicly perceived as meaningful remains an empirical question. Past evidence suggests that formal rights do not automatically translate into felt empowerment: a Dutch panel survey finds that high awareness among citizens of the GDPR also comes with strong doubts about its effectiveness in empowering them (Strycharz et al., 2020).

Objectives/Research Questions

The objective is to map how mainstream news discourse frames the EU platform regulatory stack (DSA/DMA/GDPR/AI Act) along a spectrum between protection/paternalism and capability/agency, and to assess how these framings are evaluated (supported, criticized, or treated as symbolic).

RQ1: What narrative frames dominate?

RQ2: How are users/citizens positioned?

Methods

We conduct a qualitative textual analysis of mainstream press discourse to study how public discourse articulates normative ideals and imagined futures of the EU digital governance. Here, the focus is not “platform futures” but governance futures: what EU regulation is expected to do, for whom, and with what implied model of the user. Articles will be collected from ProQuest and Media Cloud using transparent, replicable keyword searches anchored in the EU stack (DSA/DMA/GDPR/AI Act), enforcement institutions (e.g., European Commission; Digital Services Coordinators), and major platforms. Sampling will concentrate on high-salience governance moments (e.g., DSA general applicability since February 2024), with a documented audit trail of inclusion/exclusion decisions. This corpus construction strategy follows best practice guidance on the consequences of query design and validation in news-text research (Barberá et al., 2021).

Results

The project will produce a qualitative mapping of the EU “governance imagination” in mainstream discourse. We anticipate that coverage will often privilege a protection/safety rationale (risk, minors, harmful/illegal content, enforcement) while treating capability-oriented provisions (e.g., interface impairment limits; non-profiling recommender options) as secondary, technical, or symbolic. In contrast, DMA coverage may more frequently invoke capability-adjacent language through market/choice narratives (switching, interoperability, gatekeeper constraints), though not necessarily in “agency” terms.

Across the stack, we expect recurring discursive tensions: (a) governance praised as necessary protection versus criticized as paternalistic or rights-restrictive; (b) capability tools praised as empowerment versus dismissed as performative compliance; and (c)

“enshittification” nostalgia functioning either as an argument for stronger intervention or as resignation about the possibility of restoring user autonomy. The analysis will identify which actors are mobilized to legitimate each frame and how the “user” is positioned within these narratives.

Future Work

Future work will extend these findings into the project’s broader agenda by theorizing the institutional imaginary that underpins EU platform governance and testing how far its governing “resources” (e.g., GDPR information rights, transparency duties, notices, dashboards) actually convert into user capabilities in practice. In other words, the study will motivate and inform a research program on capability conversion: identifying the cognitive, social, and interface “conversion barriers” that prevent users from learning, acting, and developing agency, and specifying design- and policy-relevant pathways through which rights and disclosures could become usable freedoms under AI-amplified platform power.

References

Barberá, P., Boydston, A. E., Linn, S., McMahon, R., & Nagler, J. (2021). Automated text classification of news articles: A practical guide. *Political Analysis*, 29(1), 19-42.

boyd, d. (2024, October 8). Risks vs. harms: Youth & social media. Substack.

<https://zephoria.substack.com/p/risks-vs-harms-youth-and-social-media>

Butkowski, C., & Corry, F. (2025). Social Media’s Midlife Crisis? How Public Discourse Imagines Platform Futures. *Social Media+ Society*, 11(2), 20563051251351493.

Doctorow, C. (2025). *Enshittification: Why everything suddenly got worse and what to do about it*. Verso Books.

European Union. (2022). Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act). *Official Journal of the European Union*, L 277, 1–102. <https://eur-lex.europa.eu/eli/reg/2022/2065/oj/engGoogle>. (2026a, January 11). New tech and tools for retailers to succeed in an agentic commerce future. Google Blog. <https://blog.google/products/ads-commerce/agentic-commerce-ai-tools-protocol-retailers-platforms/>

Haidt, J. (2024). *The anxious generation: How the great rewiring of childhood is causing an epidemic of mental illness*. Penguin Press.

Handa, A., & Gupta, A. (2026, January 11). Under the hood: Universal Commerce Protocol (UCP). Google Developers Blog. <https://developers.googleblog.com/under-the-hood-universal-commerce-protocol-ucp/>

Noffsinger, J., Patel, M., & Sachdeva, P. (2025, April 28). The cost of compute: A \$7 trillion race to scale data centers. McKinsey & Company.

<https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-cost-of-compute-a-7-trillion-dollar-race-to-scale-data-centers>

PwC. (2026, January 19). CEO confidence in revenue outlook hits five-year low – as AI becomes a defining divide between leaders and laggards: PwC 2026 Global CEO Survey [Press release]. <https://www.pwc.com/gx/en/news-room/press-releases/2026/pwc-2026-global-ceo-survey.html>

Shepherd, T. (2024, November 26). “What many of us feel”: Why “enshittification” is Macquarie Dictionary’s word of the year. The Guardian. <https://www.theguardian.com/science/2024/nov/26/enshittification-macquarie-dictionary-word-of-the-year-explained>

Strycharz, J., Ausloos, J., & Helberger, N. (2020). Data protection or data frustration? Individual perceptions and attitudes towards the GDPR. Eur. Data Prot. L. Rev., 6, 407.

9A.4 “It is not about vaccines, it is about THEM!”: Othering and Discursive Violence in Anti-Vaccine Telegram Groups

Raquel Recuero (Universidade Federal de Pelotas)

Introduction

Social media platforms enable individuals with shared interests to form relatively stable online communities, while simultaneously facilitating the production and normalization of symbolic violence through boundary-making between “us” and “them.” This article examines how discursive violence is constructed in Brazilian anti-vaccine communities on Telegram by analyzing how out-groups are named and described. Grounded in the concepts of othering and symbolic violence, the study adopts a mixed-methods approach to investigate who is constructed as the “Other” and how this Other is subjected to discursive forms of violence. The analysis draws on Critical Discourse Analysis (Fairclough, 1995; Wodak, 2015) and research on digital and platformed discourse (van Dijck, 2013; Papacharissi, 2015), as well as Social Media Critical Discourse Studies (KhosraviNik, 2017).

Theoretically, we treat discourse as social practice. In CDA, discourse is understood as both shaped by and shaping social structures, contributing to the reproduction of power relations and social inequalities (Fairclough, 1995). In platformed environments, these processes are intensified by persistence, scale, and low-cost engagement mechanisms that facilitate rapid alignment and reduce the social cost of hostile participation (Papacharissi, 2015; KhosraviNik, 2017). Within this context, othering is understood as a discursive operation through which out-groups are constructed as inferior, immoral, or threatening, enabling exclusion and hostility to appear legitimate (van Leeuwen, 2008; Wodak, 2015). Symbolic violence operates through these representations by naturalizing domination via naming, classification, and stigma (Bourdieu, 1991; Goffman, 1963), while fear-oriented narratives frame the Other as a danger, mobilizing urgency and legitimizing action against perceived enemies (Wodak, 2015).

Objectives/Research Questions

This study is guided by two central research questions: (1) who is constructed as the “Other” in Brazilian anti-vaccine Telegram communities, and how are different types of out-groups discursively named and organized; and (2) how is this Other described through forms of symbolic and discursive violence. Accordingly, the study has three main objectives: first, to map and classify the dominant categories of out-groups mobilized in anti-vaccine discourse; second, to examine how different forms of discursive violence (delegitimization, stigmatization, dehumanization, and threat) are

articulated in relation to these categories; and third, to assess how othering operates as a structuring mechanism that links heterogeneous discursive domains and contributes to the legitimization of radical and antagonistic narratives within these communities.

Methods

The first analytical step addressed the question “who is the Other?” by operationalizing nomination patterns into five categories of out-groups. Four substantive categories were identified through iterative reading of the corpus: ideological enemies, referring to actors framed as politically corrupting society or threatening the moral and political order (e.g., “the left,” socialism/communism, political figures as proxies for an ideological bloc); epistemic enemies, including actors constructed as controlling or falsifying knowledge (e.g., scientists, experts, universities, medical professionals, and mainstream media); inferiorized enemies, referring to actors portrayed as socially, cognitively, or morally inferior, including women and minorities; and systemic enemies, encompassing abstract or diffuse structures of domination such as “the state,” “the government,” or “the system.” A fifth residual category (no named Other) was used when messages did not explicitly name an out-group. These categories are conceptually consistent with CDA approaches to social actor representation (van Leeuwen, 2008) and discourse-historical analyses of in-group/out-group construction (Wodak, 2015).

To classify the full dataset, lexical dictionaries were developed for each category, including morphological variants and orthographic distortions typical of Telegram discourse. Each message received a single label based on a frequency rule: matches were counted per category, and the category with the highest count was assigned. This procedure ensured comparability across categories while accounting for the centrality of ideological antagonism in structuring the discourse. Reliability was assessed using a random sample of 300 messages independently coded by a human coder and by the automated procedure, yielding a Cohen’s Kappa of 0.74, indicating substantial agreement and supporting application to the full corpus.

The second analytical step addressed “how the Other is described.” Given that discursive violence often operates through implicit evaluation, metaphor, irony, and presupposition, a purely lexical approach was insufficient. A qualitative analysis was therefore conducted on a stratified random sample of 120 messages (30 per category of Other), following a saturation-oriented logic. Messages were coded according to four types of discursive violence: delegitimization, stigmatization, dehumanization, and threat. Each message was assigned a dominant violence type; when no violence was identifiable, it was coded as non-violent. This procedure ensured analytical comparability across categories while preserving contextual interpretation consistent

with CDA. Finally, quantitative and qualitative findings were cross-compared to address the research questions.

Results

The quantitative results show that the ideological Other is the most frequent out-group in the corpus, surpassing the epistemic Other. In this study, ideological Others refer to actors framed as political enemies (e.g., “the left,” communists, political elites); epistemic Others to actors associated with knowledge production and mediation (e.g., scientists, experts, and the media); systemic Others to abstract institutions or structures of domination (e.g., “the state” or “the system”); and inferiorized Others to groups portrayed as socially or morally inferior, including women and minorities. This distribution indicates that anti-vaccine discourse in these Telegram communities is structured primarily as political conflict, rather than as a dispute over scientific evidence.

Vaccination is framed less as a public health issue and more as a symbol of ideological struggle, with science, institutions, and expertise repositioned as extensions of a political enemy. The presence of systemic and inferiorized Others further suggests that these groups function not only as spaces for debating vaccination, but also as arenas for circulating and legitimizing symbolic violence against institutions and marginalized groups.

Qualitative analysis shows that delegitimization is the most prevalent form of discursive violence, targeting institutions, experts, science, courts, and the state, often through associations with the “left,” socialism/communism, and criminality. Stigmatization appears most strongly in ideological and inferiorized Others, naturalizing stereotypes of corruption, dishonesty, manipulation, and moral degeneracy, particularly in representations of women and minorities. Dehumanization and threat, while less frequent, concentrate on ideological and inferiorized Others, including animal metaphors, contamination narratives, and portrayals of systemic enemies as dangerous or oppressive.

Taken together, the findings highlight three dynamics. First, othering functions as a mechanism of discursive articulation, connecting heterogeneous discourses—such as public health, politics, and conspiracy narratives—through the construction of common enemies, even where thematic coherence is weak. Second, this articulation is sustained by fear-oriented constructions, especially threat, stigmatization, and dehumanization, which portray the Other as morally corrupt and socially dangerous, legitimizing exclusion (Wodak, 2015). Third, these patterns resonate with scholarship on right-wing populist discourse, particularly the moralized division between a homogeneous “people” and a corrupt or dangerous enemy, and the erosion of trust in democratic institutions (Mudde, 2004; Laclau, 2005). Overall, the results suggest that

anti-vaccine Telegram discourse may operate as an entry point for broader exclusionary and authoritarian narratives, in which vaccine skepticism facilitates the legitimization of symbolic violence against institutions and minorities.

Future Work

Future research will build on these findings by more explicitly connecting forms of othering to toxicity as an analytical and measurable dimension of platformed discourse. In particular, we aim to examine how different othering strategies contribute to the normalization and circulation of toxic communicative practices across online environments. In addition, this research agenda will be expanded to include other types of groups and thematic communities beyond anti-vaccine contexts, allowing for comparative analysis across issues, platforms, and political orientations. By doing so, future work seeks to better understand the role of othering not only in shaping antagonistic identities, but also in facilitating processes of political radicalization and the legitimization of exclusionary and authoritarian narratives in digital spaces.

References

- Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. London, England: Longman.
- Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. Englewood Cliffs, NJ: Prentice-Hall.
- KhosraviNik, M. (2017). Social media critical discourse studies (SM-CDA). In J. Flowerdew & J. E. Richardson (Eds.), *The Routledge handbook of critical discourse studies* (pp. 582–596). Routledge.
- Laclau, E. (2005). *On populist reason*. London, England: Verso.
- Mudde, C. (2004). The populist zeitgeist. *Government and Opposition*, 39(4), 541–563. <https://doi.org/10.1111/j.1477-7053.2004.00135.x>
- Papacharissi, Z. (2015). *Affective publics: Sentiment, technology, and politics*. Oxford, England: Oxford University Press.
- van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford University Press.
- van Leeuwen, T. (2008). *Discourse and practice: New tools for critical discourse analysis*. Oxford, England: Oxford University Press.
- Wodak, R. (2015). *The politics of fear: What right-wing populist discourses mean*. London, England: SAGE.

Session 9B - Uses & Users (2)

Day 2 · Tues 14 July · 15:35-17:05, Seminar Suite A, ARC

9B.1 Paths towards the far-right: the role of social media in young people's political socialization

Laia Corxet Solé (Scuola Normale Superiore)

Introduction

Increasing evidence seems to suggest that young people – most specifically young men – across the western world are coming to hold regressive and far right values (Abou-Chadi, 2024; Milosav et al., 2025; Off et al., 2025). In a context of rising salience of far right actors in western politics, the normalization of reactionary and regressive values seems to be hitting the youth hard. For example, multiple sources indicate that misogynistic attitudes are increasingly present among young men around Europe (Off, et al., 2022; Campbell and Cowper-Coles, 2025; Milosav et al., 2025). In recent European elections, the vote or vote intention of the youth for the far right has increased considerably throughout, from France, to Portugal, or Germany (Abou-Chadi, 2024; Azmanova, 2024; Cokelaere, 2024; Ifop, 2024; Off, et al., 2025). When once ‘the youth’ as a demographic was regarded as a progressive stronghold, being the face of movements and protests throughout the ages (Hooghe, 2004; Inglehart & Welzel, 2001), this seems to no longer hold true.

As already stated, this phenomenon appears to be heavily gendered, with young women generally moving in the opposite ideological direction than young men (Bos et al., 2022). As a consequence, scholars are discussing the existence of an ideological gender gap among the youth, the existence of which, while seemingly obvious in some places, is still debated. However, many of the authors who explore this ideological gender gap suggest, at some point or other, the influence of social media in this phenomenon (Campbell & Cowper-Coles, 2025; Milosav et al., 2025; Renström et al., 2020)– which is precisely what this paper wishes to investigate.

This paper is part of a broader project; my PhD thesis. Consequently, it encompasses one of the avenues through which I am investigating the ways social media influences the development of far right ideas among the youth. Thus, this paper will be based on the interviews conducted with young individuals who are both social media users and hold far right beliefs, in order to explore their paths of political socialization through their own lived experiences with social media and the online sphere.

Objectives/Research Questions

The paper intends to explore the ways in which social media shapes the political socialization of young individuals in contemporary society. Through the framework of (gendered) political socialization (Bos et al., 2022; Hooghe, 2004), and focusing on the current phenomenon discussed by media and scholars alike of a growing extremization towards the right among young men in many western contexts, the paper inquires about the role(s) social media plays in the processes of political learning and identity formation among young people who populate the far right's online milieu.

Consequently, the main research question is the following:

What role(s) does social media play in the political socialization of young people towards the far right?

This paper wishes to explore the “other side” of social media, focusing on its users, rather than remaining rooted on what is visible on online. It argues that we cannot restrict our inquiry into the complex and impactful relationships between politics and social media to the online sphere, but rather advocates for exploring the phenomenon through the lived experiences of social media users. Understanding the online and the offline as extensions of one another, this paper aims to provide insights about the ways in which young people narrate their own experiences with social media and their own understanding of the source of their political views. Doing so, it is also aimed to better understand how the far right is benefiting from the prevalence of social media in young people's lives for political gain, letting users themselves highlight what has been relevant in their own socialization processes.

Drawing from Spain as a case study – owing to the considerable manifestation of the gender ideological gap phenomenon in the country (CIS, 2025; El País, 2025) –, therefore, this paper will, more specifically, explore the following question:
How do young Spanish people who hold far right views narrate their lived experiences with social media in relation to their processes of political socialization?

Methods

This paper will be based on a set of semi-structured qualitative interviews, which allow for the inquiry into personal meaning-making and lived experiences (Magnusson & Marecek, 2015). Set within a broader interpretative research approach, the use of interviews allows us to recognise the agency of individuals in their lives, also in their socialization processes. It allows us to regard people as active and cognisant of their own experiences, focusing on the way individuals narrate their own lives. Consequently, it is the method best adept to explore how social media content is perceived and received by users, as well as to understand how these users interact and relate to the discourses therein embedded.

These qualitative interviews will be conducted with young Spanish social media users who hold far right views. Following the ways in which previous scholars have approached the phenomenon of the gender ideological gap among the youth, I define young people as belonging to Generation Z; people born between 1997 and 2012 (Abou-Chadi, 2024; Milosav et al., 2025; Off et al., 2022).

In order to guarantee that the interview subjects are both social media users and hold far right views, they have been located, selected, and contacted alongside an ongoing online ethnographic process of the Spanish online far right milieu being carried out on two different social media platforms highly populated by young people: Instagram and TikTok. Therefore, the interview subjects are persons who are active in such spaces, populated by far right actors in their multiplicity – from political organizations to influencers or meme accounts – and therefore are assumed to hold a degree of sympathy for the discourses and ideas that populate such (online) space.

Results

While the process of conducting interviews is, as of yet, still ongoing, some early findings can be drawn from the conversations already had. So far, all respondents have been male – aligning both with the gendered nature of the ideological phenomenon as well as the gendered identity of the majority of users who populate the online far right milieu; and their gender had emerged as a critical element in the processes of political socialization taking place online. Respondents tend to highlight both their gender identity as a source of political interest in early teenage years, as well as it playing a role in the type of content consumed and the kinds of social relationships established online. They highlight the masculinized nature of online far right spaces, and draw links between male-catered interest and political discourse online, such as gaming or sports. Additionally, all respondents express a sense of belonging and brotherhood among men of their generation, drawing parallels between themselves and other men around them: friends (online and offline), classmates, or simply “boys my age”.

Interestingly, all respondents so far have a middle or lower-middle class background and a generally non-politicised environment. Those who are able to identify a political lenience among their familial background indicate that they mostly belong to a progressive and left-leaning family. Indeed, several respondents express having identified as left-leaning at the early stages of developing a political identity, and cite social media as a critical source of information and influence in their socialization towards the far right.

Future Work

As already mentioned, this paper contains a part of my PhD thesis project. This project not only considers the narrations of social media users, but also observes the social dynamics of the far right online milieu through an online ethnography. The insights

gained from the interview that compose the data for this paper will be complemented with the observations from the non-participant ethnographic process, comparing the experiences of users with the observed behaviours.

Additionally, also as part of my PhD thesis, the research done in Spain will also be conducted in Italy, a country that, unlike Spain – and most other western societies – does not seem to, as of yet, present evidence of an ideological gender gap among the youth, nor an increasing prevalence of far right values in comparison to previous generations (Calvi, 2024; de Sio & Cataldi, 2024; Ricci, 2025). Consequently, following a comparative framework, both contexts will be contrasted in order to inquire whether any differences in socialization processes and social media content, diffusion, or consumption can provide some justification to the differences between countries.

References

- Abou-Chadi, T. (2024). A gendered far-right wave among young voters in Western Europe? *European Journal of Politics and Gender*, 1–6.
<https://doi.org/10.1332/25151088Y2024D000000065>
- Azmanova, A. (2024, June 14). How far-right parties seduced young voters across Europe. *The Guardian*.
<https://www.theguardian.com/commentisfree/article/2024/jun/14/far-right-seduced-young-voters-europe-elections>
- Bos, A. L., Greenlee, J. S., Holman, M. R., Oxley, Z. M., & Lay, J. C. (2022). This One's for the Boys: How Gendered Political Socialization Limits Girls' Political Ambition and Interest. *American Political Science Review*, 116(2), 484–501.
<https://doi.org/10.1017/S0003055421001027>
- Calvi, G. (2024, June 4). I giovani europei e la destra. *InsideOver*.
<https://it.insideover.com/politica/europa-dalla-germania-al-portogallo-i-giovani-guardano-a-destra-ma-in-italia.html>
- Campbell, R., & Cowper-Coles, M. (2025). From a gap to a gulf? Gen Z and the gender gap in political attitudes in post-industrial societies.
<https://doi.org/10.1332/25151088Y2025D000000093>
- CIS. (2025, May 8). Avance de resultados estudio 3497 'Calidad de la democracia (III)'. CIS. <https://www.cis.es/-/avance-de-resultados-del-estudio-3497-calidad-de-la-democracia-iii->
- Cokelaere, H. (2024, May 29). It's not just boomers, young people are voting far right too. *POLITICO*. <https://www.politico.eu/article/europe-young-people-right-wing-voters-far-right-politics-eu-elections-parliament/>

de Sio, L., & Cataldi, M. (2024, June 10). Chi ha votato chi? Gruppi sociali e voto - cise.
<https://cise.luiss.it/2024/06/10/chi-ha-votato-chi-gruppi-sociali-e-voto/>

Hooghe, M. (2004). Political Socialization and the Future of Politics. *Acta Politica*, 39(4), 331–341. <https://doi.org/10.1057/palgrave.ap.5500082>

Ifop. (2024). Les jeunes et les élections européennes de 2024. IFOP.
<https://www.ifop.com/publication/les-jeunes-et-les-elections-europeennes-de-2024/>

Inglehart, R., & Welzel, C. (2001). *Modernization, Cultural Change, and Democracy: The Human Development Sequence* (1st edn). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511790881>

Magnusson, E., & Marecek, J. (2015). *Doing Interview-based Qualitative Research: A Learner's Guide*. Cambridge University Press.

Milosav, Đ., Dickson, Zachary, Hobolt, Sara B., Klüver, Heike, Kuhn, Theresa, & Rodon, T. (2025). The youth gender gap in support for the far right. *Journal of European Public Policy*, 0(0), 1–25. <https://doi.org/10.1080/13501763.2025.2481181>

Off, G., Alexander, A., & Charron, N. (2025). Is there a gender youth gap in far-right voting and cultural attitudes? <https://doi.org/10.1332/25151088Y2025D000000077>

Off, G., Charron, N., & Alexander, A. (2022). Who perceives women's rights as threatening to men and boys? Explaining modern sexism among young men in Europe. *Frontiers in Political Science*, 4, 909811. <https://doi.org/10.3389/fpos.2022.909811>

País, E. (2025, March 3). Consulte todos los datos internos de la encuesta de EL PAÍS: Cuestionarios, cruces y respuestas individuales. El País.
<https://elpais.com/espana/2025-03-03/consulte-todos-los-datos-internos-de-la-encuesta-de-el-pais-cuestionarios-cruces-y-respuestas-individuales.html>

Renström, E. A., Bäck, H., & Knapton, H. M. (2020). Exploring a pathway to radicalization: The effects of social exclusion and rejection sensitivity. *Group Processes & Intergroup Relations*, 23(8), 1204–1229. <https://doi.org/10.1177/1368430220917215>

Ricci, M. (2025, February 25). Il voto radicale dei giovani sposta i maschi a destra e le donne a sinistra—Il Diario del Lavoro. <https://www.ildiariodelavoro.it/il-voto-radicale-dei-giovani-sposta-i-maschi-a-destra-e-le-donne-a-sinistra/>

9B.2 Feeling climate: how emotional responses to data visualisation shape political engagement

Monika Fratzczak (Manchester Metropolitan University)

Introduction

The visual communication of data has risen significantly in recent years. Data visualisations are increasingly people's primary source of information about complex social issues, shaping understanding and potentially guiding everyday decisions. Despite their ubiquity, particularly online, the affective and political dimensions of data visualisations remain largely underexplored in social media research. While prior work has examined how data visualisations enhance comprehension, shape public discourse, and evoke emotions (see D'Ignazio & Klein, 2016; Kennedy & Hill, 2017; Hill, 2017), less attention has been paid to the role of these emotional responses in shaping political engagement. Similarly, although literature on data visualisation and democracy highlights its potential to foster political engagement and support civic activities (Nærland, 2020; Nærland and Engebretsen, 2021), the affective dimensions of such engagement remain under-theorised and under-researched. This is notable given that emotions are recognised as central to social and political experience alongside rational judgement (Ahmed, 2004; Wahl-Jorgensen, 2018) and as a key driver of user engagement on social media (Wollebæk et al, 2019; De León & Trilling, 2021; Xu et al., 2022; Pennington & Palagi, 2023). This gap is particularly striking given the saturation of data and their visualisations in recent public crises and online campaigns, from climate change and the Covid-19 pandemic to the UK Brexit referendum and US elections, where data visualisation have been used not only to inform but also to evoke emotions, persuade and mobilise. This study is therefore the first to empirically examine, from the users' perspective, how emotional responses to data visualisations encountered on social media influence different forms of political engagement, from online sharing to more formal offline political actions.

The paper examines climate data visualisation as an urgent and emotionally charged domain on social media, where abstract and politicised climate risks (Boykoff et al., 2022) are communicated through simplified representations that shape public understanding and meaning-making (Pearce et al., 2020).

Objectives/Research Questions

Responding to the limited academic empirical interest in social media users' experiences of data visualisation and recognising the significance of this for political engagement, the primary aim of this paper is to better understand diverse users'

emotional responses to data visualisation on social media and their significance for online political engagement and more formal offline political participation.

Within this aim, the paper addresses the following questions:

1. What kinds of emotional responses correspond to different modes of engagement with climate data visualisations on social media (e.g., sharing, discussing with others, or disengaging)?
2. How do these emotional responses translate into broader forms of offline political engagement, including everyday practices and more formal political actions?

Methods

The main discipline researching how users engage with data visualisation is computer science (as noted by Kennedy et al., 2016), which often focuses on the visual text and does not provide much information about the study participants. Consequently, these research projects mostly try to quantitatively assess the effectiveness of data visualisation and its visual elements. Therefore, this sociological approach takes into account social factors outside the visualisation text, such as “who are the users and what are their circumstances” (Kennedy et al., 2016, p.5), presenting a picture that is far more complex than other studies mentioned above can offer.

To investigate how diverse users respond to climate data visualisations through a qualitative lens, the research was conducted in two contrasting national settings, Poland and the United Kingdom, chosen for their distinct political, environmental, and media cultures surrounding climate change. This comparative approach enabled an examination of how localised discourses, historical trajectories, and environmental politics shape emotional responses to climate data visualisations, while simultaneously decentring dominant Western perspectives in data studies (Milan & Treré, 2019).

This research draws on 34 semi-structured interviews and 13 follow-up interviews with social media users from the UK (N=18) and Poland (N=16), selected through purposive and convenience sampling. It investigates responses to 13 climate data visualisations shared on social media by six environmental organisations: Greenpeace (UK/Poland), WWF (UK/Poland), Carbon Brief (UK), and Nauka o klimacie (Poland). The visualisations were selected to capture a diverse range of visual forms, design styles, and climate-related themes circulating online. The interview discussions were structured around these examples. I recruited participants primarily through the social media channels of the aforementioned organisations. My sampling aimed to capture a range of social backgrounds, with attention to age, gender, ethnicity and race, education, nationality, and occupation. Participants were not selected on the basis of strong environmental activism or alignment with pro-climate positions. My analytic approach was grounded

in thematic analysis (Braun & Clarke, 2006). All interviews were transcribed and analysed using qualitative coding software. Transcripts were inductively coded in Nvivo, and a codebook was developed to identify emerging themes, one of which, ‘The impact of emotions on political engagement’, is the focus of this paper. Coding was carried out iteratively, with themes refined through ongoing reflection and memo-writing.

Results

The findings show that different emotional valences—positive, negative, and ambivalent—towards data visualisations correspond to distinct modes of political engagement: (a) sharing and discussing data visualisation with others, (b) changes in everyday practices, and (c) more formal political actions. I argue that although engagement with data visualisation on social media rarely leads to formal political mobilisation, it can foster individualised and emotionally-mediated forms of everyday engagement.

Data visualisation on social media thus operates not merely as conveyors of numeric information but as affective mediators that connect individuals to wider publics, spark conversations, and embed environmental issues in everyday practices. They blur boundaries between online and offline, between informal and formal, between knowledge and feeling. In doing so, they enable what Dahlgren (2002) calls the civic potential of informal conversation, and what Marres (2016) theorises as material participation. At the same time, mobilisation is uneven and context-dependent. Not all participants were able to translate emotions into action, and some visualisations failed to connect to lived realities. Nonetheless, the evidence suggests that data visualisation encountered on social media can be powerful triggers of affectively charged engagement, sustaining everyday civic practices that, while modest, collectively can contribute to broader cultural and political shifts.

The paper provides an empirical analysis of how emotional responses to climate data visualisations on social media shape different forms of political engagement, offering both theoretical insights into affective digital media practices and practical implications for climate change communication. It makes an original contribution to the social media research, sociology of emotions, critical data studies, the limited existing research on data visualisation in society, and the emerging sociology of data, by advancing understanding of experiences of data visualization published on social media from an audience ‘bottom up’ perspective (Couldry and Powell, 2014) and focusing on factors that affect diverse users’ engagement and political participation.

Future Work

The paper concludes by calling for further empirical research into the affective dimensions of data visualisation and their implications for political engagement,

particularly in the context of online misinformation, where emotions influence both the interpretation of data visualisation and subsequent action.

Further research could also aim for a more demographically diverse sample and to include more marginalised voices and communities. Some measures were employed to recruit people from a range of backgrounds and efforts were made to recognise differences and possible inequalities during analysis. Nevertheless, the participants from this study were, for the main part, White British and White Polish and tended to be educated to degree level and above. This may be partly related to the fact that climate organisations/groups/movements themselves have a history of being understood as white, middle-class movements – and that this issue is compounded by datavis also being seen as a relatively elitist way of disseminating knowledge.

References

- Ahmed, S. (2004). *The Cultural Politics of Emotions*. Edinburgh: Edinburgh University Press.
- Boykoff, M., Aoyagi, M., Ballantyne, A. G., Benham, A., Chandler, P., Daly, M., Doi, K., Fernández-Reyes, R., Hawley, E., Hwang, K., Lee, K., Lyytimäki, J., McAllister, L., McNatt, M., Mervaala, E., Mocatta, G., Nacu-Schmidt, A., Oonk, D., Osborne-Gowey, J., ... Ytterstad, A. (2022). World Newspaper Coverage of Climate Change or Global Warming, 2004-2022: Media and Climate Change Observatory Data Sets. Dataset, <https://doi.org/10.25810/4c3b-b819>.
- Couldry, N. & Powell, A. (2014). Big Data from the bottom up. *Big Data & Society*, 1.
- Dahlgren, P. (2002). In Search of the Talkative Public: Media, Deliberative Democracy and Civic Culture. *Javnost - The Public*. 9(3), pp.5-25.
- De León, E., & Trilling, D. (2021). A Sadness Bias in Political News Sharing? The Role of Discrete Emotions in the Engagement and Dissemination of Political News on Facebook. *Social Media + Society* <https://doi.org/10.1177/20563051211059710>.
- D'Ignazio, C., & Bhargava, R. (2020). Data Visualization literacy: A feminist starting point. In: Engebretsen, M. & Kennedy, H. eds. *Data Visualization in Society*. Amsterdam: Amsterdam University Press.
- D'Ignazio, C., & Klein, L. (2017). Feminist Data Visualization. In: IEEE VIS Conference, Baltimore, Maryland.
- Gray, J. (2020). The data epic: Visualization practices for narrating life and death at a distance. In: Engebretsen, M. and Kennedy, H. eds. *Data Visualization in Society*. Amsterdam: Amsterdam University Press.

- Hill, R.L. (2017). The political potential of numbers: data visualisation in the abortion debate. *Women, Gender & Research*. [online]. 26(1), pp.83-96.
- Kennedy, H., Hill, R., Lucy, Allen, W., & Kirk, A. (2016). Engaging with (big) data visualizations: Factors that affect engagement and resulting new definitions of effectiveness. *First Monday*, 21(11).
- Kennedy, H., & Hill, R. (2017). The feeling of numbers: Emotions in everyday engagements with data and their visualization. *Sociology*.
- Kennedy, H., Steedman, R., & Jones, R. (2020). Approaching public perceptions of datafication through the lens of inequality: a case study in public service media. *Information, Communication & Society*, 24:12.
- Marres, N. (2016). *Material participation: Technology, the environment and everyday publics*. Basingstoke: Palgrave Macmillan
- Milan, S., & Treré, E. (2019). Big Data from the South(s): Beyond Data Universalism. *Television & New Media*. 20, pp.319-335.
- Nærlund, T.U. (2020). The political significance of data visualization: Four key perspectives. In: Engebretsen, M. & Kennedy, H. eds. *Data Visualization in Society*. Amsterdam: Amsterdam University Press. pp.40-47.
- Nærlund, T. U., and Engebretsen, M. (2021) Towards a critical understanding of data visualisation in democracy: a deliberative systems approach. *Information, Communication & Society*, 26(3), 637–655.
- Pearce W., Özkula S. M., Greene A. K., Teeling L., Bansard J. S., Omena J. J., Rabello E. T. (2020). Visual cross-platform analysis: Digital methods to research social media images. *Information, Communication & Society*, 23(2), 161–180.
<https://doi.org/10.1080/1369118X.2018.1486871>.
- Pennington, N., & Palagi, J. (2023). Examining How Social and Emotional Factors Inform Response to Cross-Cutting Political Views on Social Media. *Social Media + Society*, 9. <https://doi.org/10.1177/20563051231207851>
- Wahl-Jorgensen, K. (2018). *Emotions, Media and Politics*. Polity Press.
- Wollebæk, D., Karlsen, R., Steen-Johnsen, K., & Enjolras, B. (2019). Anger, Fear, and Echo Chambers: The Emotional Basis for Online Behavior. *Social Media + Society*, 5. <https://doi.org/10.1177/2056305119829859>.
- Xu, Z. Laffidy, M., & Ellis, L. (2022). Clickbait for climate change: comparing emotions in headlines and full-texts and their engagement. *Information, Communication & Society*.

9B.3 Playing with Fate: Play Framing, Context Collapse, and Personal Data Disclosure in AI-Mediated Bazi Divination

Hui Lin (King's College London), Yingwen Wang (University of The Arts London)

Introduction

This study examines how Chinese users disclose sensitive personal data when using generative AI chatbots for Bazi (八字) divination on Xiaohongshu. Bazi is a traditional Chinese metaphysical system that derives personality traits and life predictions from the precise year, month, day, and hour of birth, requiring users to input highly specific biographical information. Unlike tarot or Western horoscopes, which require minimal personal data, Bazi readings depend on exact temporal coordinates that can serve as identifying information. On Xiaohongshu, users share step-by-step prompt instructions for obtaining Bazi readings from ChatGPT, DeepSeek, and other large language models, while comment sections evaluate which prompts yield the most convincing results. This prompt-sharing ecology raises questions about what enables users to disclose sensitive data so readily to commercial AI systems whose data retention and usage practices remain opaque.

Scholarship on digital divination has primarily examined how platforms become sites of spiritual meaning-making. St. Lawrence (2024) documented how TikTok users in the Witchtok community treat the platform's recommendation algorithm as a divinatory partner, interpreting algorithmically delivered content as spiritually significant. This work illuminates how users imbue algorithmic curation with sacred meaning, yet it focuses on content consumption rather than active data disclosure. The Chinese AI Bazi phenomenon differs in a crucial respect: users do not merely interpret algorithmically delivered content but actively submit intimate biographical data in exchange for personalised predictions. Understanding what facilitates this shift requires attention to both privacy decision-making and the social practices surrounding AI fortune-telling.

Privacy calculus theory posits that individuals rationally weigh perceived benefits against privacy risks when deciding whether to disclose information (Dinev and Hart, 2006). Yet empirical research consistently finds that this rational model inadequately explains actual disclosure behaviour. Gerber et al. (2018) demonstrated in their systematic review that perceived benefits predict disclosure more strongly than risk perceptions, while Kehr et al. (2015) showed that affective processing frequently overrides deliberative assessment. Trang and Weiger (2021) extended this line of inquiry by examining gamification, finding that gamified services increase users' willingness to disclose personal information by inducing cognitive absorption that depletes resources

for careful privacy evaluation. Their work suggests that playful contexts may fundamentally alter how users approach disclosure decisions.

Objectives/Research Questions

We build on these insights to argue that AI Bazi divination on Xiaohongshu operates within the play frame: an interactional context constituted by social proof metrics (likes, saves, comments), iterative prompt experimentation, community validation of accuracy, and low-stakes repeatability. This play frame produces functional context collapse, which we define as a situation where users perceive themselves within one informational context (sacred-therapeutic consultation) while their data flows into a structurally incompatible context (commercial platform infrastructure). Unlike Marwick and boyd's (2011) original formulation of context collapse as audience flattening, functional context collapse describes a disjuncture between perceived and operational contexts that remains invisible to users precisely because the play frame provides a coherent experiential logic. This study asks: (1) What platform affordances and community practices constitute the play frame around AI Bazi divination? (2) How do users understand the interactional context when disclosing personal data? (3) What disclosure strategies do users employ, and what logics justify these choices?

Methods

We employed a qualitative design combining platform observation with semi-structured interviews. Between September and December 2025, we observed approximately 200 Xiaohongshu posts retrieved using keywords including DeepSeek fortune-telling, Divination prompt, and Bazi prompt. We sampled both high-engagement posts (over 500 likes) and lower-engagement posts to capture variation in the prompt-sharing ecology. Observation focused on prompt structures, visible personal information in screenshots, and comment discussions about accuracy. We recorded fieldnotes but did not archive screenshots containing user data. This observation informed interview protocol development rather than serving as primary data.

We recruited interview participants through purposive sampling via Xiaohongshu and WeChat, seeking users who had used generative AI for Bazi divination at least three separate sessions. At the time of writing, we have completed 10 of a target 20 interviews, conducted in Mandarin between December 2025 and January 2026 via video call or in person (34–78 minutes each). The protocol explored initial encounters with AI divination, comparisons with human practitioners, emotional experiences, and perceptions of privacy. All interviews were transcribed verbatim and analysed using reflexive thematic analysis (Braun and Clarke, 2006). Coding was conducted in Mandarin to preserve meaning before translation of illustrative quotations. This study received ethical approval from related University Ethics Committee. Participants

provided informed consent and are identified by pseudonyms. We did not collect participants' actual birth data; we asked them to describe what information they typically enter without reproducing it.

Results

Three interconnected themes emerged from our analysis. First, participants framed AI Bazi divination as recreational exploration rather than serious spiritual consultation. Unlike consulting human practitioners, which participants described as requiring psychological preparation and financial commitment, AI divination could be initiated spontaneously. XiaoYang (32, freelancer) explained: "If a post has lots of likes, saves, and comments saying this prompt is accurate, I will try it, even if I have no questions right now." Meiling (28, graduate student) similarly described trying new prompts "for fun" when scrolling through her feed. This orientation, consistent with Trang and Weiger's (2021) finding that playful contexts alter privacy decision-making, transformed disclosure from deliberate act to incidental aspect of exploration.

Second, participants positioned AI simultaneously as fortune-teller and emotional support, a dual role human practitioners rarely occupy. XiaoYang explained her preference for AI when anxieties recurred: "If I have talked to a friend many times about the same thing, I do not want to bother them again, so I choose AI". When asked to characterise what AI Bazi consultation provided, she responded: "Therapeutic conversation, I suppose". This therapeutic framing situated AI within an intimate context even as data potentially flowed to commercial infrastructure, exemplifying the disjuncture between perceived and operational contexts that characterises functional context collapse.

Third, participants exhibited selective opacity: disclosing precise birth data while withholding names. Few participants reported using pseudonyms or English names while providing exact birth times. XiaoYang explained: "I do not need to let these public websites know who I am". When we noted that birth data combined with other traces could potentially identify her, she distinguished between active and passive disclosure: "If they passively use my data for something, I do not care. But actively, I will not let them use my privacy to do things". This logic reveals how the play frame reconstitutes disclosure: birth data becomes a functional input for the game rather than sensitive personal information. One participant, however, expressed discomfort: Jiahui (31, technician) reported that he "did not quite dare" to enter his real birth time after reading about data breaches, instead rounding to the nearest hour.

Crucially, participants acknowledged that AI predictions were unreliable even as they continued seeking them. XiaoYang stated: "Emotionally it helps, but rationally I know this is unreliable". What mattered was receiving a concrete answer to temporarily alleviate anxiety, not predictive accuracy. Participants also developed verification

practices, asking AI to analyse past events before trusting future predictions, which paradoxically deepened engagement by making platform comparison itself a form of play.

Future Work

Our preliminary findings suggest that the play frame does not simply lower disclosure thresholds but reconstitutes the meaning of disclosure itself. Within this frame, sharing birth data becomes a move in a game rather than an act with surveillance implications. This reconstitution depends on functional context collapse: users experience AI as confidant and diviner while their data may be retained, logged, or repurposed according to provider policies they have not examined. We emphasise may because we have not traced actual data flows; how AI companies handle birth data remains empirically unexamined and constitutes a priority for future research. The prompt-sharing ecology on Xiaohongshu warrants focused analysis as a peer-to-peer network that may normalise disclosure through community validation. We are currently expanding our interview sample to include users who have reduced or ceased AI divination use, which will strengthen claims about disclosure consequences.

References

- Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology* 3(2): 77-101.
- Dinev, T. and Hart, P. (2006) An extended privacy calculus model for e-commerce transactions. *Information Systems Research* 17(1): 61-80.
- Gerber, N., Gerber, P. and Volkamer, M. (2018) Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers and Security* 77: 226-261.
- Kehr, F., Kowatsch, T., Wentzel, D. and Fleisch, E. (2015) Blissfully ignorant: The effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal* 25(6): 607-635.
- Marwick, A.E. and boyd, d. (2011) I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media and Society* 13(1): 114-133.
- St. Lawrence, E. (2024) The algorithm holy: TikTok, technomancy, and the rise of algorithmic divination. *Religions* 15(4): 435.
- Trang, S. and Weiger, W.H. (2021) The perils of gamification: Does engaging with gamified services increase users' willingness to disclose personal information? *Computers in Human Behavior* 116: 106644.

9B.4 “you are treshhhhh player” - Toxicity in Twitch Livestreams

Tamara Bodden (Universität Koblenz), Maite Taboada (Simon Fraser University)

Introduction

Twitch is a digital platform primarily used by gamers that reaches over 2 million users daily and assembles an international audience of different age groups. With their interactive mix of 'hot' (live videos) and 'cold' media content (chat functions) (Sjöblom et al. 2016, 4), live streams on Twitch offer a unique communication structure in which meaningful communities emerge and form specific practices (Diwanji et al. 2020, 7). Like any other platforms, Twitch however is facing high levels of toxicity i.e. antisocial behaviour in interpersonal communication, especially tying up to multiplayer online gaming (e.g. Kwak & Blackburn 2014, Martens et al. 2015). Anker Nexø (2024, 3) identifies online video-based approaches as a current desideratum in toxicity research on competitive multiplayer gaming, Twitch streams make audiovisual gameplay data publicly available and can serve as an information-dense multimodal research object. While current research approaches on toxicity on Twitch mainly focus on chats (e.g. Groen 2020, Dreier & Pirker 2023), the example setting of streamers (Seering et al. 2017, Poyane 2019) or user's perspectives (Cai & Wohn 2019) the multimodal interplay between the layers of a live stream have not been investigated yet either. The project therefore develops methodological approaches to investigate live streaming but also aims to fill two current 'blind spots' in toxicity research.

Because of the immense amount of audiovisual data on Twitch (with an average of 93.2K concurrent streamers and 2.17M concurrent viewers in 2025; Twitchtracker 2026) Twitch relies on a mixture of centralized and decentralized moderation tools. This can lead to a lack of transparency in moderation actions, the potential of false flagging (see Zolides 2023, 134) and abuse of a low-paid workforce (like Amazon Turk workers) for sensitive decisions about violations (Seering 2020, 5-6). Users however seem to see the responsibility for content moderation mostly with streamers and community moderators instead of the platform (Cai & Wohn 2019: 166). While 96,76 % of Streams use moderation (Twitch 2022), especially larger communities tend to rely on bots instead of human moderators (Seering 2020, 15) using automated commands and block lists.

While these lists of offensive or objectionable words can help prevent some toxic comments, they can also be easily circumvented by using different spellings or homophones (Saleem et al. 2016, 3) and need constant modification (Nobata et al. 2016, 146). Taking the multimodality on Twitch into account, emoticons can be a part of

toxic communication (Kim et al. 2022), but toxicity can also occur in the streamed games (Anker Nexø 2024 for an overview) or in nonverbal/paraverbal actions of the streamer (e.g. aggressive voice, offensive gestures) which demands a multimodal approach for understanding toxicity in streams more extensively.

Objectives/Research Questions

The paper explores antisocial behaviour, platform culture and online communities on Twitch with a qualitative and multimodal framework asking what forms of multimodal toxicity can be found in Twitch streams (RQ1) and how the multimodal toxicity on the levels of game, stream and chat influence one another (RQ2).

While many different aspects of toxicity in lives streams and gaming have been analysed before, they were not put into relation to one another so far. To close this desideratum, we have developed a classification scheme for multimodal toxicity in live streams that can help distinguishing all potential occurrences of toxicity including for example gameplay actions, gestures and tonality of the streamer or emotes in addition to spoken and written communication.

As Taboada (2021) noted, toxicity is often trivial and not necessarily intended or obvious to all participants of a toxic interaction. The classification scheme and results of the study therefore can enable researchers, policy makers, platforms and content creators to become more aware of different forms of toxic behaviour. While the influence of toxic communication was already discussed in the context of example setting of streamers (Seering et al. 2017, 112) a ‘toxicity map’ of the stream was generated by annotating all potentially toxic actions and identifying trigger events in the different layers of the stream to visualize the spreading of toxic actions and their interconnectedness.

Methods

Since live streams cannot be described with a simple “sender-receiver model, but one that is circular and characterized by information exchange between all parties” (Diwanji et al. 2020, 7), toxicity can occur at different multimodal stages of the stream that influence one another: the layers of game (including group communication in MOGs), streamer and chat. A transcript of a massive Twitch stream (110 minutes video material, over 18.000 chat messages) that includes all spoken and written communication as well as a description of both visual and auditive layers (tying up to and expanding Recktenwald’s 2017 transcription scheme) is the basis for the pragma-linguistic analysis. The transcript consists of seven layers (visual + auditive descriptions of game, stream, chat along a time axis) that was generated by extracting the Twitch chats, transcribing all auditive utterances in game and stream and describing all relevant visible actions. Using a python script, all usernames in chat messages (except for bots) were automatically anonymized with persistent identifiers.

The sample was chosen through discourse analysis of online texts discussing toxic streamers and communities and features a gameplay stream of the competitive online game DotA2 by a popular Twitch streamer. According to the investigated online discourse his community is considered extremely toxic, nonetheless he is one of the most successful Twitch streamers, and people return to watch his content and interact in the Twitch chat.

Knowledge about the communities is necessary to understand and recognize specific linguistic patterns but also helps to differentiate the self-classification of toxic language by users. The same expression can be considered acceptable or toxic in different groups (Nobata et al. 2016, 146). Consequently, toxicity is conceptualized following Wittgenstein (1967, PU 23) as Sprachspiel that adheres to specific rules and is embedded in a context, requiring a pragma-linguistic approach. The study therefore annotates toxicity triggers (Almerekhi et al. 2019) and potential toxicity, taking the perspective of community practices into account. For this the qualitative annotation program MAXQDA (Verbi) was used in a semi-automated process through manual annotation and automated detection of repeated toxic Emotes in the transcript of the stream. This data (over 2500 annotations) enables a quantitative and qualitative evaluation of the material by tracking targets of toxic actions (generalized vs. directed), topics and content (such as sexism, ableism or curse words) and especially the impact of one layer of the stream on another.

Results

The paper identifies different types of multimodal toxicity that did not receive much attention in research yet and therefore are not likely to be detected by current approaches. It combines a wide overview of potentially toxic actions in live gaming streams in a multimodal annotation scheme. This can serve as a methodological basis for future research but can also be used for optimizing automated detection. A ‘toxicity map’ following the time axis of the stream and separating the layers was created to furthermore visualize the results.

While game actions like loss or death can be identified as primary triggers for toxic behaviour of the streamer (in accordance with the research on online gaming), other factors like character taunts (e.g. teammates congratulating on a bad playing move) or background music (chat messages referring to “99 noobs”, with the song ‘99 Red Balloons’ playing in the background) can also play a role connecting different layers of the stream.

The chat seems to amplify any toxic utterance of the streamer directed at other players (creating a toxic echo chamber) but also directs toxicity at the streamer himself (e.g. Emotes with his face titled as ‘Loser’ or ‘RetardJAM’), which is often not distinguishable between playful bantering and actual attacks. This ambiguity of a bantering but

potentially toxic tone alongside generalized toxicity (for example sexist or ableist emotes that are circulated in the community or bots creating porn images in chat) create a potentially toxic atmosphere for specific groups. However, it seems as if this toxicity is not undetected (and could be prevented) but a deliberately shared practice (Saleem et al. 2016 on self-identified toxic communities) that is tolerated or invited by the streamer despite of the infringements of Twitch's community guidelines. This leads to open questions for future work on the awareness of toxicity within these groups and their motivations for toxic behaviour.

Future Work

Taking the first analysis as a basis, the project continues to investigate a wider sample of Twitch streams (currently about 16 h video material, 10 streamers) in different sizes, with different game genres and varying genders of the streamers to extend and verify the preliminary results with an extended corpus. The goal is to utilize current toxicity detection programs and compare their results (including false positives and undetected toxicity) with the developed multimodal approach and identify toxic patterns that can be included in automated research (for example Twitch/gaming specific terms, Emotes).

Another future goal is the (sociolinguistic) investigation of toxicity as a positive, community building factor that can bond teams in competitive online gaming as well as communities that explicitly self-define through shared toxic practices.

References

- Almerekhi, H., Kwak, H., Salminen, J., & Jansen, B. J. (2019). Detecting Toxicity Triggers in Online Discussions. Proceedings of the 30th ACM Conference on Hypertext and Social Media (HT'19), 291–292. <https://doi.org/10.1145/3342220.3344933>
- Anker Nexø, L. (2024). Toxic Behaviours in Esport: A Review of Data-Collection Methods Applied in Studying Toxic In-Gaming Behaviours. International Journal of Esports, 1(1). <https://www.ijesports.org/article/127/html>
- Cai, J., & Wohn, D. Y. (2019). What are Effective Strategies of Handling Harassment on Twitch? Users' Perspectives. Companion Publication of the 2019 Conference on Computer Supported Cooperative Work and Social Computing (CSCW '19 Companion). Association for Computing Machinery, 166–170. <https://doi.org/10.1145/3311957.3359478>
- Diwanji, V., Reed, A., Ferchaud, A., Seibert, J., Weinbrecht, V., & Sellers, N. (2020). Don't just watch, join in: Exploring information behavior and copresence on Twitch. Computers in Human Behavior 105(2020) 106221.
- Dreier, L., & Priker, J. (2023). Toxicity in Twitch Live Stream Chats: Towards Understanding the Impact of Gender, Size of Community and Game Genre. 2023

- IEEE Conference on Games (CoG), 21.08.2023, 1.
<http://dx.doi.org/10.1109/CoG57401.2023.10333159>
- Groen, M. (2020). Digital Governmentality: Toxicity in Gaming Streams. In M. Groen, N. Kiel, A. Tillmann, & A. Weßel (Eds.), *Games and Ethics. Digitale Kultur und Kommunikation* (pp. 97–111). Springer VS. https://doi.org/10.1007/978-3-658-28175-5_7
- Kim, J., Wohn, D. Y., & Cha, M. (2022). Understanding and identifying the use of emotes in toxic chat on Twitch. *Online Social Networks and Media*, 27, 100–180.
<https://doi.org/10.1016/j.osnem.2021.100180>
- Kwak, H., & Blackburn, J. (2014). Linguistic analysis of toxic behavior in an online video game. *Social Informatics*. Springer, 209–217.
- Martens, M., Shen, S., Iosup, A., & Kuipers, F. (2015). Toxicity detection in multiplayer online games. 2015 International Workshop on Network and Systems Support for Games (NetGames), 1–6.
- Nobata, C., Tetreault, J., Thomas, A., Mehdad, Y., & Chang, Y. (2016). Abusive language detection in online user content. *Proceedings of the 25th International Conference on World Wide Web*, ser. WWW '16. International World Wide Web Conferences Steering Committee, 145–153.
<https://doi.org/10.1145/2872427.2883062>
- Poyane, R. (2019). Toxic Communication on Twitch.tv. Effect of a Streamer. In D. Alexandrov, A. V. Boukhanovsky, A. V. Chugunov, Y. Kabanov, O. Koltsova & I. Musabirov (Eds), *Digital Transformation and Global Society. DTGS 2019. Communications in Computer and Information Science* (pp 414–421). Springer.
https://doi.org/10.1007/978-3-030-37858-5_34
- Recktenwald, D. (2017). Toward a Transcription and Analysis of Live Streaming on Twitch. *Journal of Pragmatics*, 115, 68–81.
<https://doi.org/10.1016/j.pragma.2017.01.013>
- Saleem, H. M., Dillon, K. P., Benesch, S., & Ruths, D. (2016). A web of hate: Tackling hateful speech in online social spaces. *Proceedings of Text Analytics for Cybersecurity and Online Safety (TA-COS)*, LREC Portoroz, 1–9.
- Seering, J., Kraut, R., & Dabbish, L. (2017). Shaping Pro and Anti-Social Behavior on Twitch Through Moderation and Example-Setting. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, ser. CSCW '17. Association for Computing Machinery, 111–125.
<https://doi.org/10.1145/2998181.2998277>

- Seering, J. (2020). Reconsidering community self-moderation: The role of research in supporting community-based models for online content moderation. Proceedings of the ACM on Human-Computer Interaction, 4, Article 107. <https://doi.org/10.1145/3415178>
- Sjöblom, M., Hamari, J., (2016). Why Do People Watch Others Play Video Games? An Empirical Study on the Motivations of Twitch Users. SSRN Electron. J. <http://dx.doi.org/10.2139/ssrn.2779543>
- Taboada, M. (2021, October 4). The banality of online toxicity. <https://policyoptions.irpp.org/magazines/october-2021/the-banality-of-online-toxicity/>
- Twitch (2022). Transparenzbericht für das 2. Halbjahr 2022. <https://safety.twitch.tv/s/article/H2-2022-Transparency-Report?language=de>
- Twitchtracker (2026). Twitch Statistics and Charts. <https://twitchtracker.com/statistics>
- Wittgenstein, L. (1984). Philosophische Untersuchungen. In Tractatus logico-philosophicus. Tagebücher 1914-1916. Philosophische Untersuchungen. Suhrkamp.
- Zolides, A. (2023). Toxic community policing: Weaponizing moderation tools on Twitch. In J. Brewer, B. Ruberg, A. L. L. Cullen, C. J. Persaud (Eds.), Real Life in Real Time: Live Streaming Culture (pp.131–144). MIT Press.

Session 9C - Politics, Policy, & Regulation (2)

Day 2 · Tues 14 July · 15:35-17:05, Seminar Suite B, ARC

9C.1 Do Electoral Stages Shape Digital Campaigning? Meta Strategies in Uruguay's 2024 Elections

*Juan Pérez Martirené, Rafael Piñeiro Rodríguez (Universidad Católica del Uruguay),
Fernando Rosenblatt (University of Manchester)*

Introduction

Campaigns no longer begin with party events or press conferences; they begin with an ad that appears on the screens of the electorate's devices. Unlike traditional media, advertising on social media decentralizes communication and broadens access to diverse audiences. Political advertising on social media is a central component of electoral campaigns, influencing political competition and the democratic discourse quality (Enli and Skogerbø 2013; Stromer-Galley 2014). Platforms such as Facebook and Instagram offer flexible and lower-cost access, allowing small and medium-sized campaigns to reach audiences that were previously unattainable (Gibson et al. 2025; Fowler et al. 2021).

Despite recent advances, empirical research on political advertising on social media in Latin America remains limited, with most studies focusing on official candidate's accounts, rather than providing systematic multi-level analyses. Uruguay provides a compelling case to study these dynamics. The study of advertising on Meta platforms by political parties and candidates is relevant for understanding how the institutional incentives of the Uruguayan electoral cycle influence campaigns, particularly the use of digital tools. The scale and characteristics of Uruguay's party system (level of nationalization and institutionalization) and electoral system (simultaneous primary elections) make it possible to analyze all the advertising produced by all electoral actors across all the stages of the campaign (13,438 ads).

Uruguay's presidential elections unfold in three electorally distinct stages with differentiated audiences and incentive structures: primaries (voluntary voting), national elections (compulsory voting), and, if required, presidential runoff. This electoral format allows us to identify patterns of mobilization and persuasion in advertisements depending on the stage of the election and to evaluate the opportunities and risks that platforms generate for democratic quality. In a context where algorithmic segmentation redefines audiences and competition, it becomes crucial to generate evidence about

campaigns on these platforms for the debate on the regulation, transparency, and political impact of digital campaigns.

Objectives/Research Questions

The goal of this study is to examine whether digital strategies of candidates advertising in Meta align with incentive structures of the electoral system of Uruguay. We analyze whether Meta electoral ads vary across the three different stages of the national electoral cycle: primaries, general elections, and the presidential runoff. The research questions addressed in this study are:

RQ1: How do message typologies vary across the three electoral stages: primaries, general elections, and the presidential runoff?

RQ2: How do strategic incentives shape digital communication across different levels of electoral competition (national and local)?

These objectives and questions are grounded in stage-specific theoretical expectations. In primaries, where voting is voluntary, voter mobilization and grassroots outreach are expected to predominate (Denton et al., 2008; Stromer-Galley et al., 2021). In the general election, compulsory voting reduces the need for mobilization, and campaigns are expected to emphasize policy-based persuasion targeting the median voter. In the runoff, characterized by head-to-head competition between presidential tickets, we expect an increase in negative or contrastive messaging as candidates attempt to reach crossover voters.

The analysis also considers Uruguay's multilayered electoral competition. Parties compete not only at the national level (presidency and Senate) but also through factions and local-level candidates contesting at departmental districts (representatives). In this way, it is possible to obtain a more complete picture of the electoral advertising to which voters are subjected in an election (which usually involves more than one level of competition) and, at the same time, it enables us to assess whether message strategies differ across national parties, factions, district level candidates, complementing the stage-based analysis.

Methods

This study employs a text-as-data approach (Grimmer et al. 2022) to classify political advertising on Meta throughout Uruguay's 2023–2024 electoral cycle. We retrieved 18,833 ads categorized as “social issues, elections or politics” using the Meta Ad Library API and retained 13,438 ads after excluding non-electoral content through manual review. Each ad was linked to its corresponding party, candidate, and constituency (national or departmental) using page-name identification, keyword searches, and inspection of visual content.

To construct a unified textual corpus, multiple sources of textual information were integrated. We used the ad body text provided by the API, audiovisual ads were processed by extracting and transcribing audio tracks automatically, and texts embedded in images were retrieved using optical character recognition. This multimodal preprocessing ensured full textual coverage across heterogeneous ad formats.

For classification, we implemented a semi-supervised machine learning approach using a BERT model (Devlin et al. 2019) fine-tuned on manually labeled data. We selected ROUBERTa (Filevich et al. 2024), a model pretrained on Uruguayan news text and optimized to capture local expressions, slang, and context-specific names. Trained coders independently labeled a stratified subset of 1,236 ads following a detailed coding manual; disagreements were adjudicated to create a consensus gold-standard set used for training and calibration. Model performance, evaluated using precision, recall, and F1 scores, was stable across categories, with macro and micro F1 scores of 0.70 and 0.75, respectively.

Ads were classified according to the typology proposed by Stromer-Galley et al. (2021), which distinguishes between advocacy, attack, image, issue, call-to-action, and ceremonial messages. The typology captures the diversity of functions assumed by political ads on Meta. It differentiates between strategies focused on persuasion, using either personal attributes or programmatic proposals, from messages aimed at mobilizing or strengthening symbolic links with different audiences.

Results

The results show that the proportion of ad types throughout the stages of the electoral cycle remains stable, with similar proportions in all elections. This means that, contrary to what the literature suggests, the specific rules and objectives of each stage do not significantly influence the types of ads. Overall, the campaign was characterized by programmatic promotional messages and widespread use of calls to action. Attack ads were in the minority throughout the cycle, indicating a relatively less confrontational tone.

The differences in the distribution of ad types between parties were very small. More meaningful variation emerged when distinguishing advertisers by constituency. Ads with national reach showed a greater emphasis on programmatic messages, whereas subnational advertisers displayed higher proportions of image-based and call-to-action content, consistent with their need to mobilize geographically targeted audiences. Across all parties and stages, the platform's ad-delivery algorithm did not systematically privilege specific message types. The ratio of spending/impression closely mirrored the underlying distribution of classified ads, suggesting that algorithmic allocation reproduced, rather than reshaped, campaign strategies.

Overall, these findings align with the characteristics of Uruguay's political system, where relatively low affective polarization and strong programmatic competition reduce incentives for negative campaigning. The stability observed across electoral stages and actor types indicates that institutional differences in voting rules and campaign objectives did not translate into substantial changes in digital messaging strategies. Instead, political actors relied on a consistently programmatic and mobilization-oriented communication style throughout the electoral cycle.

Future Work

This project is part of an ongoing research agenda, and several extensions are planned. First, future electoral cycles will offer an opportunity to assess whether the stability of message strategies observed here persists over time or responds to changes in party competition, institutional reforms, or shifting patterns of voter engagement. Comparative analyses with countries that exhibit higher polarization, different primary systems, or alternative regulatory frameworks would further clarify whether the observed dynamics are specific to Uruguay's institutional incentives or reflect broader tendencies in digital campaigning.

Second, as platform governance and transparency standards continue to evolve, future research could evaluate how changes in disclosure requirements, ad-library completeness, or targeting rules affect the strategic behavior of parties and candidates. Understanding whether these regulatory shifts alter the visibility, distribution, or framing of political ads is essential for debates on electoral fairness and democratic accountability.

Finally, ongoing work will refine the typology and the classifier used in this study. Improving performance for less frequent categories requires revisiting the labeled training set and integrating additional contextual features. Enhancing model metrics will allow more precise identification of strategic differences across electoral stages and actor types, strengthening the analytical leverage of text-as-data approaches for studying digital campaigning in institutionalized party systems.

References

- Denton, R. E., Voth, B., Trent, J. S., & Friedenberg, R. V. (2008). Political campaign communication: Principles and practices (Sixth edition). Rowman & Littlefield.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In J. Burstein, C. Doran, & T. Solorio (Eds), Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers) (pp. 4171–4186). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N19-1423>

- Enli, G. S., & Skogerbø, E. (2013). Personalized Campaigns in Party-Centred Politics: Twitter and Facebook as Arenas for Political Communication. *Information, Communication & Society*, 16(5), 757–774.
<https://doi.org/10.1080/1369118x.2013.782330>
- Filevich, J. P., Marco, G., Castro, S., Chiruzzo, L., & Rosá, A. (2024). A Language Model Trained on Uruguayan Spanish News Text. In F. Gaspari, J. Moorkens, I. Aldabe, A. Farwell, B. Altuna, S. Piperidis, G. Rehm, & G. Rigau (Eds), *Proceedings of the Second International Workshop Towards Digital Language Equality (TDLE): Focusing on Sustainability @ LREC-COLING 2024* (pp. 53–60). ELRA and ICCL.
<https://aclanthology.org/2024.tdle-1.5/>
- Fowler, E. F., Franz, M. M., Martin, G. J., Peskowitz, Z., & Ridout, T. N. (2021). Political Advertising Online and Offline. *American Political Science Review*, 115(1), 130–149. <https://doi.org/10.1017/S0003055420000696>
- Gibson, R., Southern, R., Vaccari, C., Smyth, P., & Musayev, J. (2025). Does digital campaigning matter, and if so, how? Testing a broadcast versus network effects model of candidates Twitter use. *Journal of Information Technology & Politics*, 22(3), 357–372. <https://doi.org/10.1080/19331681.2024.2387634>
- Grimmer, J., Roberts, M. E., & Stewart, B. M. (2022). *Text as data: A new framework for machine learning and the social sciences*. Princeton University Press.
- Stromer-Galley, J. (2014). *Presidential campaigning in the Internet age*. Oxford University Press.
- Stromer-Galley, J., Rossini, P., Hemsley, J., Bolden, S. E., & McKernan, B. (2021). Political Messaging Over Time: A Comparison of US Presidential Candidate Facebook Posts and Tweets in 2016 and 2020. *Social Media + Society*, 7(4).
<https://doi.org/10.1177/20563051211063465>

9C.2 Enabling informed voter decisions via parsimonious factually falsifiable social media design

Patrick McHugh, Tommy Barton, Sami Peretz (Brown University)

Introduction

This research arose from a simple question considered by the authors when contemplating the U.S. election process: “Why is it a more rigorous process to apply for university admissions than it is to apply for a position in congress?” From this initial question many others followed from a comparison of political elections and academic admission processes: Where are candidate CVs? Where are candidate integrity statements? What happens if candidates lie to voters? Where is the standard “application” format to enable easy comparison across candidates? Does a lack of accessibility, trust and standardization for candidate consideration impact voter satisfaction and elected official quality?

This study considered whether voters made quick, heuristic-based candidate selections and, if yes, whether social media design choices in communications about candidates could lead to greater systematic thinking. The heuristic systematic model (HSM) of decision making, known in the trade literature as “thinking fast and slow” (Chaiken, 1980 & Kahneman, 2011). noted the positive impact of motivation and ability on the selection of systematic thinking by decision makers. HSM along with the communication theory, informational social influence, informed a series of hypotheses tested in this study and a proposed “parsimonious factually falsifiable social media design” (PF2SM) methodology. The study investigated if subjects provided with a clearer understanding of their role, the elected candidate’s role, relevant candidate data in a format that could easily be cross analyzed, and assurances of trust resulted in statistically significant shifts in candidate election outcomes.

Objectives/Research Questions

The heuristic systematic model distinguishes between quick heuristic driven and slower systematic driven decisions (Chaiken, 1980). In disciplined hiring or admissions processes committees are established to review all candidates in a structured manner against relevant criteria consistently gathered for ease of candidate comparison. These processes and structures were designed to facilitate thoughtful, i.e. systematic, decision making. At an abstract level the voting and hiring or admission processes are similar in that voters need to review and compare candidate capabilities in the context of the requirements of the political appointment the candidate is “applying for”.

Nawara and Bailey (2021) noted in a study of repeated scandals committed by political candidates that, supported by cognitive load theory, voters become overwhelmed, thus ignoring the cumulative “messages” candidate data provides. These findings suggest that the systematic decision making process may be best enabled with a parsimonious, easy to compare, communication design leading to the following hypothesis:

H1 Simple clear candidate data arranged to facilitate easy comparison will be preferred by voters to current methods of candidate data access.

Perceived information quality was found to be important to social media purchase intent (Fu, Lu, Chen, Farn, 2020) and warning labels were found to reduce observer’s perceived credibility of fake news posts (Koch, Frischlich, Lermer, 2023). These findings suggest the following hypothesis.

H2 Affirmation statements will be impactful to voter candidate selection.

Social influences, both normative and informational, were found to have a greater impact on consumer shopping intent than perceived information quality (Fu, Lu, Chen, Farn, 2020) and a subject’s political attitude was found to be significantly impactful in Koch, Frischlich, and Lermer’s (2023) study done in the context of climate change news. These findings suggest the following hypothesis.

H3 Political party affiliation will be impactful to voter candidate selection.

Methods

This research utilized vignettes incorporated in an initial exploratory study with a convenience sample of 52 eligible U.S. voters sourced by the authors to evaluate the impact of: PF2SM design to improve candidate comparison, candidate truthfulness, and party affiliation. Subject data was analyzed utilizing binomial tests given the binary outcomes resulting from the simulated election decisions of the subjects and their preference for the presentation of candidate information tested in this study compared to their existing methods of candidate evaluation. Candidate CV data was juxtaposed for easy comparison and candidate behaviors were modeled via vignette scenario changes. As scenarios were adjusted a Cochran’s Q test was utilized to evaluate impacts on candidate selection and Pairwise McNemar tests utilized to discern differences between various scenario results.

Results

H1, subject’s perceptions of the PF2SM design approach to candidate data presentation versus their existing methods were analyzed with 31/52 (59.6%) preferring the PF2SM approach to candidate comparison. However, the binomial test against 50% was not significant (one-sided $p=0.11$). The 95% CI for the preference rate was 45.1% to 73.0%.

H2, affirmation statement impact and H3, political party affiliation impact was next analyzed. Subjects first selected a preferred candidate from two profiles, then reconsidered their choice under three scenarios. The proportion who switched their selection was 53.8% (28/52) when the chosen candidate refused to sign an affirmation statement, 61.5% (32/52) when the candidate signed but was found to be meaningfully lying, and 84.6% (44/52) in the general-election scenario where the opposing candidate's CV was clearly superior and switching implied changing party. A Cochran's Q test indicated that switching rates differed significantly across scenarios, $Q(2) = 10.95$, $p = .004$. Pairwise McNemar exact tests (Holm-adjusted for three comparisons) showed no reliable difference between the refused-to-sign and signed-but-lied scenarios ($p_{\text{Holm}} = .597$). Switching in the general-election/superior-opponent scenario was significantly higher than both the refused-to-sign scenario ($p_{\text{Holm}} = .0046$) and the signed-but-lied scenario ($p_{\text{Holm}} = .0236$).

Future Work

Additional data needs to be collected to validate the initial findings observed in this exploratory study and to expand the diversity of subjects. Chat GPT was used in conducting the initial computational analysis and these results need to be validated by the authors using Python or SPSS. Qualitative information on the subject's design preferences was also collected during this data collection, focused on their perceptions of the PF2SM design, providing rich input for the future discussion section.

Post this ongoing analysis sample questions for future research consideration follow. If PF2SM design can be impactful in triggering systematic thinking, what are other contexts where this framework could be applied to improve decisions and outcomes? Subjects demonstrated a willingness to adjust candidate selection when presented with superior candidates or candidate transgressions provided in an easy to interpret format. Poor primary candidate selection needs to be better understood. Why are poor quality candidates advanced, who benefits from such "non-Darwinian" selection, and can PF2SM designs positively impact this selection process in practice?

References

- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*. 39(5) 752-766.
- Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, 51(3), 629.

Fu, J., Lu, I., Chen, J., Farn, C. (2020) Investigating consumers' online social shopping intention: An information processing perspective, *International Journal of Information Management*. 54.

Gottfried, J., Park, E. (2025). American's social media use 2025. Pew Research Center. https://www.pewresearch.org/wp-content/uploads/sites/20/2025/11/PI_2025.11.20_Social-Media-Use_REPORT.pdf

Kahneman, D. (2011). *Thinking, fast and slow*. Penguin Books.

Kahneman, D., Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. *Econometrica*, 47(2) 263-291.

Koch, T., Frischlich, L., Lerner, E. (2023). Effects of fact-checking warning labels and social endorsement cues on climate change fake news credibility and engagement on social media. *Journal of Applied Social Psychology*. 2023 53:495–507. DOI: <https://onlinelibrary.wiley.com/doi/10.1111/jasp.12959>

Nawara, S., Balley, M. (2021). Scandal-ridden campaigns: the relationship between cognitive load and candidate evaluation. *The Social Science Journal*. 65(1)

Pew Research Center. (2025). Social media and news fact sheet. <https://www.pewresearch.org/journalism/fact-sheet/social-media-and-news-fact-sheet/>

9C.3 A wholesome insult, an uplifting threat: Making sense of toxicity amid 'good vibes' digital cultures

Esteban Morales (University of Groningen)

Introduction

Digital platforms increasingly mediate how we experience violence (Recuero, 2024), shaping how we connect with others (Walther, 2024) and make sense of our surroundings (Author, blinded). The critical role of this platformization of violence is particularly evident on Reddit, the focus of this paper, where approximately one out of every 100 comments is estimated to be toxic in nature (Kumar et al., 2023). This toxicity is not at the fringes of Reddit—indeed, it can be found within mundane conversations, from news and politics to travel and history (Tufa et al., 2024). Given the ease with which hateful speech is enacted within and expanded through Reddit's infrastructures, Massanari (2017) has described some environments on this platform as toxic technocultures.

But to define Reddit as exclusively a toxic space due to the presence of platformized violence would paint an incomplete picture. On the one hand, previous literature has shown that toxicity is not only about harming others—it can also promote political discussion (McCosker, 2014), be a strategy for belonging (Author, blinded), foster creative insurgency (Tovares, 2018), and even act as a form of affective resistance (Sundén & Paasonen, 2018). On the other hand, terms such as 'toxicity' are not neutral categorizations of deviant behaviour, but rather intentional political and discursive constructions of acceptability, values, and norms (Carmi, 2022; Gibson et al., 2024).

Objectives/Research Questions

Attending to both the seemingly pro-social uses of toxicity and the inherent power in defining what is toxic in the first place calls for a critical examination of the social and cultural processes involved in the deployment of antisocial behaviour. To achieve this, in this paper, I focus on the use of toxicity in spaces where people seek the opposite of a toxic culture—that is, in communities that strive to define themselves in terms of 'good vibes.' The term 'vibes' is used here to denote a collective affective orientation of the communities, aimed at fostering shared positive moods and sensations (James, 2024; Lupinacci, 2024). By focusing on the use of toxicity amid these 'good vibes' digital cultures, I do not aim to erase its real and consequential ethical impacts (Jane, 2015). Instead, I aim to foreground the tension between the social and affective gratifications that motivate people to enact harm and its antisocial consequences for online communal life.

Methods

To address this research objective, I draw on a digital ethnography of Reddit communities (Postill & Pink, 2012), using this approach as a methodological lens to foreground the ebb and flow of everyday life on digital platforms (Markham, 2017). I selected three subreddits among the top 100 most popular communities that explicitly promote ‘good vibes’: r/MadeMeSmile, r/WholesomeMemes, and r/UpliftingNews. My analysis is based on participant observation, field notes, and content analysis, focusing on how toxicity operates in the everyday practices of these groups. Drawing on scholarship on platformization (van Dijck et al., 2019) and violence (Dwyer, 2017), I conceptualize toxicity as that which harms the social body (Tirrell, 2017), a socially constructed manifestation of platformized violence, highlighting how practices, cultures, and structures of harm are mediated through digital infrastructures. Accordingly, I adopt a broad analytical lens on what counts as toxic, from overt insults and threats to more subtle forms of exclusion and dehumanization.

Results

Despite their stated positive orientation, these spaces exhibit levels of toxicity comparable to other parts of Reddit. A toxicity analysis using Detoxify through Communalistic (Gruzd & Mai, 2025) shows that between 6.98% and 7.28% of comments contain highly toxic language, exceeding prior estimates for the platform (Kumar et al., 2023). Such toxicity is visible even in otherwise mundane interactions. For instance, one user remarked: “What a dumb, gross lady.” If the landscape of violence in these communities broadly resembles that of other spaces, how, then, can we make sense of its presence within good-vibes digital cultures? Here, we argue that understanding toxicity requires moving beyond its mere presence or frequency—and instead focusing on how toxic discourse is mobilized within specific normative and sociocultural contexts.

Following this approach, we focus then on how toxic discourse functions as a mechanism of social positioning. Drawing on positioning theory (Moghaddam & Harré, 2010), our analysis shows how users employ violence to negotiate belonging, morality, and political alignment within these good vibes communities. Here, two analytically distinct forms of positioning emerge. Performative acts use toxicity to reinforce communal norms and delineate moral boundaries, often aligning the group against perceived outsiders. For instance, in a thread about Tucker Carlson and Ukraine, a highly upvoted comment read: “Tucker Carlson was unable to reply as his tongue is stuck in Putin’s ass,” illustrating how toxicity operates as a boundary-making practice. Accountability acts, by contrast, mobilize toxicity to challenge what is deemed acceptable within the community. For example, a racist comment in r/MadeMeSmile (“What the Chinese propaganda!? They’re desperately trying to make more slaves lol”)

was heavily downvoted and criticized, demonstrating how toxic discourse is also used to police the limits of good-vibes cultures.

Our analysis further shows that platform affordances—particularly visibility mechanisms such as upvotes, downvotes, and awards—play a central role in modulating whether toxicity functions performatively or accountably. On Reddit, amplified toxic comments tend to reinforce dominant community positions, while downvoted comments signal collective rejection of attempts to challenge them. For instance, a comment about Trump supporters (“This is how fucking stupid they are”) received over 1,600 upvotes and an award, becoming the most visible contribution in the thread. Through this amplification, toxicity functioned performatively, solidifying the position that such political alignments were unwelcome in the community. In this way, the (in)visibility of violence shapes the communities’ structures of feeling (Duncombe, 2020; Papacharissi, 2016) through processes that are both collective and intentional.

Future Work

In brief, toxicity in so-called good vibes digital cultures cannot be understood simply as a breakdown of civility or an aberration of platform everyday life. Rather, toxic discourse operates as a central mechanism of social positioning, through which users negotiate belonging, morality, and political alignment, among others. As the analysis shows, toxicity can coexist with—and even sustain—the ethos of good vibes when it functions through positioning acts that reinforce communal norms rather than undermine them. Moreover, by distinguishing between performative and accountability uses of toxicity, the paper advances a conceptual approach for understanding the complex and often ambivalent roles of harmful speech in digital culture. These findings also point to important directions for future research, particularly in examining toxicity as a political and discursive force that operates across multiple levels—from the community dynamics illustrated here to the broader regimes through which platforms such as Reddit define, amplify, and regulate what counts as toxic.

References

- Almerekhi, H., Jansen, S. B. B. J., & Kwak, C.-S. B. H. (2020). Investigating Toxicity Across Multiple Reddit Communities, Users, and Moderators. Companion Proceedings of the Web Conference 2020, 294–298.
<https://doi.org/10.1145/3366424.3382091>
- Are, C. (2024). Flagging as a silencing tool: Exploring the relationship between de-platforming of sex and online abuse on Instagram and TikTok. New Media & Society. <https://doi.org/10.1177/14614448241228544>

- Carmi, E. (2022). Don't Be Antisocial: The Politics of the "Anti-Social" in "Social" Media. In D. Rosen, *The Social Media Debate* (1st ed., pp. 38–52). Routledge.
<https://doi.org/10.4324/9781003171270-4>
- Duncombe, C. (2020). Social media and the visibility of horrific violence. *International Affairs*, 96(3), 609–629. <https://doi.org/10.1093/ia/iaa055>
- Dwyer, P. (2017). Violence and its histories: Meanings, methods, problems. *History and Theory*, 56(4), 7–22. <https://doi.org/10.1111/hith.12035>
- Gibson, A. D., Docherty, N., & Gillespie, T. (2024). Health and toxicity in content moderation: The discursive work of justification. *Information, Communication & Society*, 27(7), 1441–1457. <https://doi.org/10.1080/1369118x.2023.2291456>
- Gruzd, A., & Mai, P. (2025). Communalystic – A Research Tool For Studying Online Communities and Online Discourse. [Computer software].
<https://communalystic.org/>
- Harris, P., Fisher, R., Wickramaarachchi, T. I., & Brock, C. H. (2024). Examining Varieties of Positioning: Considerations From the Past to the Present and Into the Future. In *The Routledge International Handbook of Positioning Theory*. Routledge.
- James, R. (2024, August 23). Introduction to GOOD VIBES ONLY [Substack newsletter]. Its Her Factory Newsletter. https://itsherfactory.substack.com/p/introduction-to-good-vibes-only?utm_medium=web
- Jane, E. A. (2015). Flaming? What flaming? The pitfalls and potentials of researching online hostility. *Ethics and Information Technology*, 17(1), 65–87.
<https://doi.org/10.1007/s10676-015-9362-0>
- Kumar, D., Hancock, J., Thomas, K., & Durumeric, Z. (2023). Understanding the Behaviors of Toxic Accounts on Reddit. *Proceedings of the ACM Web Conference 2023*, 2797–2807. <https://doi.org/10.1145/3543507.3583522>
- Lupinacci, L. (2024). Mixed feelings: The platformisation of moods and vibes. *AoIR Selected Papers of Internet Research*.
<https://doi.org/10.5210/spir.v2024i0.13992>
- Markham, A. (2017). Ethnography in the digital internet era. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE Handbook of Qualitative Research*.
- Massanari, A. (2017). #Gamergate and The Fappening: How Reddit's algorithm, governance, and culture support toxic technocultures. *New Media & Society*, 19(3), 329–346. <https://doi.org/10.1177/1461444815608807>

- McCosker, A. (2014). Trolling as provocation: YouTube's agonistic publics. *Convergence: The International Journal of Research into New Media Technologies*, 20(2), 201–217. <https://doi.org/10.1177/1354856513501413>
- Moghaddam, F. M., & Harré, R. (2010). Words, conflicts and political processes. In F. M. Moghaddam & R. Harré (Eds.), *Words of Conflict, Words of War: How the Language We Use in Political Processes Sparks Fighting* (pp. 1–28). Praeger.
- Papacharissi, Z. (2016). Affective publics and structures of storytelling: Sentiment, events and mediality. *Information, Communication & Society*, 19(3), 307–324. <https://doi.org/10.1080/1369118X.2015.1109697>
- Postill, J., & Pink, S. (2012). Social Media Ethnography: The Digital Researcher in a Messy Web. *Media International Australia*, 145(1), 123–134. <https://doi.org/10.1177/1329878X1214500114>
- Recuero, R. (2024). The platformization of violence: Toward a concept of discursive toxicity on social media. *Social Media + Society*, 10(1), 20563051231224264. <https://doi.org/10.1177/20563051231224264>
- Sundén, J., & Paasonen, S. (2018). Shameless hags and tolerance whores: Feminist resistance and the affective circuits of online hate. *Feminist Media Studies*, 18(4), 643–656. <https://doi.org/10.1080/14680777.2018.1447427>
- Tirrell, L. (2017). Toxic speech: Toward an epidemiology of discursive harm. *Philosophical Topics*, 45(2), 139–161. <https://doi.org/10.5840/philtopics201745217>
- Tovares, A. V. (2018). Trolling as creative insurgency: The carnivalesque delegitimization of Putin and his supporters in online newspaper commentary. In A. S. Ross & D. J. Rivers (Eds.), *Discourses of (De)Legitimization: Participatory Culture in Digital Contexts* (1st ed.). Routledge. <https://doi.org/10.4324/9781351263887>
- Tufa, W. T., Markov, I., & Vossen, P. (2024, April 29). The constant in HATE: Analyzing toxicity in Reddit across topics and languages. *Proceedings of the Fourth Workshop on Threat, Aggression & Cyberbullying @ LREC-COLING-2024*. <https://aclanthology.org/2024.trac-1.1/>
- van Dijck, J., Nieborg, D., & Poell, T. (2019). Reframing platform power. *Internet Policy Review*, 8(2). <https://doi.org/10.14763/2019.2.1414>
- Vizcaíno-Verdú, A., & Abidin, C. (2022). Music challenge memes on TikTok: Understanding in-group storytelling videos. *International Journal of Communication*, 16(0), Article 0.

- Walther, J. B. (2024). Making a case for a social processes approach to online hate. In J. B. Walther & R. E. Rice (Eds.), *Social Processes of Online Hate* (1st ed.). Routledge. <https://doi.org/10.4324/9781003472148>
- Warren, Z., & Moghaddam, F. M. (2016). Positioning Theory and Social Justice. In *The Oxford Handbook of Social Psychology and Social Justice* (pp. 319–330).

9C.4 “The culture is CHANGING!!!”: #MAHA, RFK, and the TikTok health influencer

*Hannah Artman (University of Missouri), Rutendo Chimbaru (University of Miami),
Bonnie Deal (Phoenix College)*

Introduction

In November of 2024, then President-elect Trump announced his nomination of RFK Jr. for secretary of health (The American Presidency Project, 2025). RFK has long been criticized for his skepticism of vaccines, and since his nomination in November, a social media trend known as “#makeamericahealthyagain” or #MAHA (a play on Trump’s campaign slogan, Make America Great Again or #MAGA) emerged and gained significant traction both on- and offline. Since RFK’s time in office, the administration has touted their accomplishments in launching initiatives to phase out synthetic food dyes and encouraging fast food chains to move away from using seed oils in their deep-fried foods (Demopoulos, 2025). While some of these initiatives are supported by nutritionists, the #MAHA movement has also led to the reversal of COVID-19 vaccination protocols and further lifting of vaccine requirements. The administration and followers of the #MAHA movement have largely celebrated these changes as victories, yet many concerns have arisen from the general public, and more notably from medical practitioners, as the goals towards lowering childhood obesity have now uncomfortably blended with the questioning of long held scientific consensus. As these debates play out in the public sphere, it is important to understand how the movement gained so much traction online, largely through TikTok, where many mothers, influencers, and everyday users have formed a community that resists dominant health and nutrition narratives (Demopoulos, 2025).

This study examines the most viral #MAHA TikTok posts over a six-month period from the announcement of RFK’s appointment in November 2024 to June 2025 in order to identify who are the most prominent voices leading the movement on the platform and what discursive framing techniques drive their influence. As there are no studies to date regarding this topic, it is worth examining the makeup of the popular #MAHA movement, and its (mis)alignment with broader conservative ideologies. Such research helps our understanding of platformized societies and the divisions that result from polarized social media discourse.

Objectives/Research Questions

TikTok and other social media platforms have become central sites for political discussion, debate, and expression. The multimodality of the platform allows for social movements, such as the 2022 Iranian protests, to coalesce by emphasizing visual cues

that symbolize the movement (Walsh, 2024). Forms of “reactionary” politics have also found large audiences on these platforms, as they allow creators an ability to repurpose media content and contest mainstream narratives through their video responses and coverage (Lewis, 2020).

TikTok has also emerged as a key source for health information, provided through short, engaging videos which are created by a range of users including health professionals, general users and organizations. Despite the opportunities for health education, there is increased risk of misinformation on the platform. Scholars find that most health content on the platform is misleading or lacks scientific backing (Kong et al., 2021; Hill et al., 2023; Yeung et al., 2022). Furthermore, videos from non-professionals often get higher engagement and go viral in comparison to higher quality evidence-based content from professionals who have less reach (Kirkpatrick & Lawrie, 2024; Turuba et al., 2023; Kong et al., 2021).

The technological affordances of social media platforms allow users to create digital communities beyond their physical and immediate surroundings, referred to as “networked publics” (boyd, 2010), thus shaping unique dynamics of public participation and discourse. Whereas networked publics were theorized to understand digital “hang out” spaces, studies have revealed how networked counterpublics utilize platform affordances, and in particular hashtags, to generate community activism around a certain issue. Networked counterpublics rely on the digital community to propel an issue forward by garnering platform visibility. This can be seen in the #MeToo movement (Trott, 2021), Black Lives Matter (Jackson & Foucault Welles, 2016), or the 2022 protests in Iran (Walsh, 2024). These communities may weaponize a hashtag by reappropriating its original meaning and bringing visibility to their movement.

The #MAHA movement complicates this classification of digital communities, as it is made up of both political and apolitical actors, on a platform with a reputation of leaning liberal. So, what kind of public constitutes the conservative leaning movement such as #MAHA? In addition to the documenting evidence of how the #MakeAmericaHealthyMovement gained traction on TikTok, we offer theoretical contributions to our understanding of networked (counter)publics as they manifest in today’s polarized digital era.

RQ1: What are the most viral #MAHA videos on TikTok, and what publics do they represent?

RQ2: What discursive framing techniques are used in the #MAHA movement on TikTok?

Methods

Videos with the hashtags “#MakeAmericaHealthyAgain” and/or “#MAHA” were scraped using the TikTok API. The sample collection period was from November 2024 (marking

the announcement of RFK as the Secretary of Human Health and Services) to June 2025. A six-month period was chosen to focus on the initial rise of the #MAHA movement so that we can appropriately assess the most influential actors and messaging to use as a foundation for future studies regarding political and lifestyle trends on TikTok. The total sample of videos scraped is $n=4132$, however, we selected the 100 most viral videos (most “liked” versus most commented or shared, though there are further studies that can be done using these video analytics) and offer a descriptive account of the makeup of these 100 videos, as well as a thematic analysis (Braun & Clarke, 2006; Joffe, 2011) of the most salient messaging.

Results

Of our sample, many videos are expressly health oriented, others politically oriented, and only a small few demonstrating an ability to appeal to both. Viewers watch as influencers prepare meals while discussing international and historical policies regarding food regulation. Other videos feature individuals, typically female, reading off ingredient labels in their kitchen and in grocery stores, aiming to seize the cultural moment and gain visibility. Doctors and nurses participate in creating #MAHA content that avoids political overtones, instead focusing content on what RFK’s nomination might mean for psych meds, obesity, food dyes, etc. They rely on their expertise in the medical field to deliver information in “explainer” style narratives.

Conversely, many posts featured political statements in support of Trump, RFK, and the Republican party in general. A post just after the 2024 Presidential Election results were announced reads “Sleep tight tonight my babies. You’re safe. Mommy fought hard for you & all of the other babies” with the caption, “Rest easy, mamas. We did it”. The accompanying image and sound appeal to a soft, femininized aesthetic, demonstrating how the intertextuality of music, image, text, and caption work together to appeal to motherhood as it is situated in the conservative movement. Though it is a video-centric platform, the multimodality of TikTok allows users to adapt by creating non-video content that still garners engagement.

Despite the well-known “#MAHAMoms” or “momfluencers”, who have been attributed to widespread support for RFK, very few videos explicitly link health and political narratives. These “MAHA moms” are mostly conservative women who center their traditional gender role as a homemaker and childbearer concerned with the nutrition of their children (Demopoulos, 2025). In one video, a woman films herself in her car with the text overlay, “I woke up./ I’m not a feminist/ I’m not fat/ I eat lots of red meat / Trump is president / Life is good”. She connects anti-feminism, body-shaming, a high protein diet, and support for Trump in one short clip. They position themselves as a political and cultural minority by resignifying leftist slogans and platforms to propel their movement and their personal brand forward.

Others utilize a concept known as “hashtag jacking” (Abidin, 2021a), where users co-opt social media movements by using a popular hashtag (such as #MAHA) alongside oppositional content. While most content from medical professionals was not expressly political, it did express disagreement with some of the core tenets of the movement. Other users call attention to the cognitive dissonance of #MAHA supporters, such as, “First Lady Michelle Obama suggested we eat healthier and you people called her a socialist.” These users employ similar social media multimodal production techniques such as humor and cross-posting across platforms, to call attention to their political opposition of RFK and Trump.

Future Work

To our knowledge, no studies or reports have analyzed the makeup of the MAHA movement’s driving forces and messages. The thematic analysis of the most viral videos revealed a combination of health-focused, political- focused, uniquely #MAHA (dual-focused), and counterpublic movements that utilize this hashtag to gain visibility. Future studies should analyze specifically health-related messages using a deductive framework to understand how health information circulates on TikTok, particularly when it is coupled with political messages. Similarly, studies should also analyze strictly politically related messages to critically understand how platform affordances contribute to populist movements and power negotiations. As the circulation of both health and political messages on social media constitute a form of populist expertise, future studies should also examine source transparency and credibility through content analyses or interviews. Finally, TikTok offers rich multimodality for researchers to analyze a variety of texts including comments and network analysis to provide deeper insight into the participatory politics of social media movements.

References

- Abidin, C. (2018). *Internet celebrity: Understanding fame online*. Emerald Publishing Limited.
- Abidin, C. (2021a). From “networked publics” to “refracted publics”: A companion framework for researching “below the radar” studies. In *Social Media + Society* (Vol. 7, Issue 1, pp. 1–13). <https://doi.org/10.1177/20563051209844>
- Abidin, C. (2021b). Mapping internet celebrity on TikTok: Exploring attention economies and visibility labours. In *Cultural Science Journal* (Vol. 12, Issue 1, pp. 77–103).
- Banet-Weiser, S., & Miltner, K. M. (2016). # MasculinitySoFragile: Culture, structure, and networked misogyny. *Feminist Media Studies*, 16(1), 171–174.
- Berdón-Prieto, P., Herrero-Izquierdo, J., & Reguero-Sanz, I. (2023). Political polarization and politainment: Methodology for analyzing crypto hate speech on TikTok. *Profesional de La Información*, 32(6). <https://doi.org/10.3145/epi.2023.nov.01>

- Bishop, S. (2025). Influencer creep: How artists strategically navigate the platformisation of art worlds. *New Media & Society*, 27(4), 2109–2126.
<https://doi.org/10.1177/14614448231206090>
- boyd, danah. (2010). Social network sites as networked publics: Affordances, dynamics, and implications. In *A networked self* (pp. 47–66). Routledge.
<https://www.taylorfrancis.com/chapters/edit/10.4324/9780203876527-8/socialnetwork-sites-networked-publics-affordances-dynamics-implications-danah-boyd>
- Canary, H. E., Wellman, N., & Martinez, L. S. (2023). COVID-19, genetics, and risk: Content analysis of Facebook posts early in the coronavirus pandemic. *Health Communication*, 38(8), 1654–1665.
- Cervi, L. (2023). TikTok use in municipal elections: From candidate-majors to influencerpoliticians. *Más Poder Local*, 53, 8–29.
- Demopoulos, A. (2025, January 13). Rise of the Maha moms: The momfluencers embracing RFK’s push for a ‘healthier’ America. *The Guardian*.
<https://www.theguardian.com/lifeandstyle/2025/jan/13/rfk-jr-moms-health-trump>
- Fairclough, N. (2008). Critical discourse analysis in transdisciplinary research. In R. Wodak & P. Chilton (Eds.), *Discourse Approaches to Politics, Society and Culture* (Vol. 13, pp. 53–70). John Benjamins Publishing Company.
<https://doi.org/10.1075/dapsac.13.06fai>
- Fraser, N. (1992). Rethinking the public sphere: A contribution to the critique of actually existing democracy. In *Public Space Reader* (pp. 34–41). Routledge.
- Gaddini, K. C. (2021). ‘Wife, mommy, pastor and friend’: The rise of female evangelical microcelebrities. In *Religions* (Vol. 12, Issue 9, p. 758).
- Glatt, Z. (2022). We’re all told not to put our eggs in one basket”: Uncertainty, precarity and cross-platform labor in the online video influencer industry. *International Journal of Communication*, 16, 3853–3871.
- Glatt, Z. (2024). The intimacy triple bind: Structural inequalities and relational labour in the influencer industry. *European Journal of Cultural Studies*, 27(3), 424–440.
<https://doi.org/10.1177/13675494231194156>
- Harris, B. C., Foxman, M., & Partin, W. C. (2023). “Don’t Make Me Ratio You Again”: How Political Influencers Encourage Platformed Political Participation. *Social Media + Society*, 9(2), 20563051231177944.
<https://doi.org/10.1177/20563051231177944>

- Hill, G., Manuell, A., & Willemsen, A. (2023). A rapid review of quality of health information on TikTok. *The Journal of Health Design*, 8(2).
- Jackson, S. J., & Foucault Welles, B. (2015). Hijacking #myNYPD: Social Media Dissent and Networked Counterpublics. *Journal of Communication*, 65(6), 932–952. *Communication & Mass Media Complete*. <https://doi.org/10.1111/jcom.12185>
- Jackson, S. J., & Foucault Welles, B. (2016). # Ferguson is everywhere: Initiators in emerging counterpublic networks. In *Information, Communication & Society* (Vol. 19, Issue 3, pp. 397–418).
- Jones, C. (2024, February 27). These Influencers Are Making Content to Make You Angry— And It’s Working. *Rolling Stone*.
<https://www.rollingstone.com/culture/culturefeatures/what-is-rage-bait-influencers-making-people-angry-1234976621/>
- Jong-Fast, M. (2022, January 13). Owning the Libs Is the Only GOP Platform. *The Atlantic*.
<https://www.theatlantic.com/newsletters/archive/2022/01/owning-the-libs-is-the-only-gop-platform/676692/>
- Kirkpatrick, C. E., & Lawrie, L. L. (2024). TikTok as a source of health information and misinformation for young women in the United States: survey study. *JMIR infodemiology*, 4(1), e54663.
- Kong, W., Song, S., Zhao, Y. C., Zhu, Q., & Sha, L. (2021). TikTok as a health information source: Assessment of the quality of information in diabetes-related videos. *Journal of Medical Internet Research*, 23(9), e30409.
- Lau, N., Srinakaran, K., Aalfs, H., Zhao, X., & Palermo, T. M. (2025). TikTok and teen mental health: An analysis of user-generated content and engagement. *Journal of Pediatric Psychology*, 50(1), 63-75.
- Lewis, R. (2020). “This is what the news won’t show you”: YouTube creators and the reactionary politics of micro-celebrity. In *Television & New Media* (Vol. 21, Issue 2, pp. 201–217).
- MacKinnon, K. R., Kia, H., & Lacombe-Duncan, A. (2021). Examining TikTok’s potential for community-engaged digital knowledge mobilization with equity-seeking groups. *Journal of Medical Internet Research*, 23(12), e30315.
- Martinez, L. S., Savage, M. W., Jones, E., Mikita, E., Yadav, V., & Tsou, M. H. (2022). Examining vaccine sentiment on Twitter and local vaccine deployment during the COVID-19 pandemic. *International Journal of Environmental Research and Public Health*, 20(1), 354.

- Marwick, A. E., & Partin, W. C. (2024). Constructing alternative facts: Populist expertise and the QAnon conspiracy. In *New Media & Society* (Vol. 26, Issue 5, pp. 2535–2555). <https://doi.org/10.1177/14614448221090201>
- Nair, I., Patel, S. P., Bolen, A., Roger, S., Bucci, K., Schwab-Reese, L., & DeMaria, A. L. (2023). Reproductive health experiences shared on TikTok by young people: Content analysis. *JMIR Infodemiology*, 3, e42810.
- Newman, N. (2024). What do we know about the rise of alternative voices and news influencers in social and video networks? | Reuters Institute for the Study of Journalism. <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024/risealternative-voices-and-news-influencers-social-and-video-networks>
- O'Donnell, N., Jerin, S. I., & Mu, D. (2023). Using TikTok to educate, influence, or inspire? A content analysis of health-related EduTok videos. *Journal of Health Communication*, 28(8), 539-551.
- Papacharissi, Z. (2016). Affective publics and structures of storytelling: Sentiment, events and mediality. In *Information, communication & society* (Vol. 19, Issue 3, pp. 307– 324).
- Quick, M., & Maddox, J. (2024). Us, Them, Right, Wrong: How TikTok's Green Screen, Duet, and Stitch help shape political discourse. *First Monday*.
- Sattora, E. A., Ganeles, B. C., Pierce, M. E., & Wong, R. (2024). Research on health topics communicated through TikTok: A systematic review of the literature. *Journalism and Media*, 5(3), 1395-1412.
- Sharevski, F., Vander Loop, J., Jachim, P., Devine, A., & Das, S. (2024, September). 'Debunk-It-Yourself': Health Professionals Strategies for Responding to Misinformation on TikTok. In *Proceedings of the New Security Paradigms Workshop* (pp. 35-55).
- Song, S., Zhao, Y. C., Yao, X., Ba, Z., & Zhu, Q. (2021). Short video apps as a health information source: An investigation of affordances, user experience and users' intention to continue the use of TikTok. *Internet Research*, 31(6), 2120-2142.
- Trott, V. (2021). Networked feminism: Counterpublics and the intersectional issues of #MeToo. *Feminist Media Studies*, 21(7), 1125–1142.
- Tudehope, L., Sofija, E., & Harris, N. (2024). VentTok: Exploring the mental health narrative on TikTok. *Stigma and Health*.
- Turuba, R., Cormier, W., Zimmerman, R., Ow, N., Zenone, M., Quintana, Y., Jenkins, E., David, S-B., Raimundo, A., Marcon, A., Mathias, S., Henderson, J. L., & Barbic, S. (2024). Exploring how youth use TikTok for mental health information in British

Columbia: Semistructured interview study with youth. *JMIR Infodemiology*, 4(1), e53233.

Walsh, T. (2024). TikTok as a site of social protest in Iran's Gen-Z uprising. *Discourse & Society*, 35(5), 625–650. <https://doi.org/10.1177/09579265241234351>

Yeung, A., Ng, E., & Abi-Jaoude, E. (2022). TikTok and attention-deficit/hyperactivity disorder: a cross-sectional study of social media content quality. *The Canadian Journal of Psychiatry*, 67(12), 899-906.

Zenone, M., Ow, N., & Barbic, S. (2021). TikTok and public health: A proposed research agenda. *BMJ Global Health*, 6(11).

Session 9D - Digital Selfhood and Social Experience

Day 2 · Tues 14 July · 15:35-17:05, Studio 2, ARC

9D.1 Navigating Agency and Constraint: Racialised Emotions and Resistant Subjectivities in TikTok Anti-Asian Activism

Xin Zhao, Yunfei Li, Xiao Ma (Bournemouth University)

Introduction

In recent years, the Asian community has increasingly turned to social media platforms as a collective means of resisting racism, particularly in response to the heightened anti-Asian discrimination during the COVID-19 pandemic. Mirroring offline movements such as “Stop Asian Hate,” this digital activism has gained notable traction, with the hashtag #StopAsianHate appearing in approximately 337,500 TikTok posts as of February 2025. Many of these posts feature creators expressing emotions as they recount or reflect on experiences of racial discrimination.

Bonilla-Silva (2019: 3) conceptualises these emotions as racialised emotions, referring to “socially engendered emotions in racialized societies”, encompassing both feelings and affective states. Individuals across different racial groups develop distinct emotional repertoires according to their perceived “location in the racial order,” thereby forming and reflecting their “emotional subjectivity” (Bonilla-Silva, 2019: 2). Social media platforms such as TikTok provide spaces for articulating these emotions and constructing subjectivities in response to real-life events, including lived experiences of racial discrimination (Lee and Lee, 2023; Zhao and Abidin, 2023). This phenomenon challenges conventional assumptions that emotions are counterproductive to serious societal discussions such as those on racial equality and the corresponding tendency to suppress or disregard emotionally charged discourse in the public domain (Kim, 2016).

The role of emotions, such as anger, in offline and online activism aiming to challenge unequal social norms has long been an academic focus (Chon, 2023; van Zomeren et al., 2008). Specific to anti-racism, Kim (2016: 439) argues that discussing racial equality accompanied by emotion may “revitalize the public discourse on racial equality.” While existing scholarship on the interplay between emotion and digital activism has primarily

focused on the facilitative role of emotions in advancing activist goals, a gap remains in understanding other forms of emotional display, particularly emotion regulation, that operate alongside emotional agency, weaving into the complex and multifaceted racialised subjectivities of online actors.

Drawing on the scholarship on racialised emotions, emotion regulation, and digital anti-racism, this study examines the manifestation of the agency and regulation of racialised emotions in 164 TikTok videos featuring anti-racism content related to the Asian community. It aims to shed light on the formation of emotional subjectivities and the extent to which they challenge the racial order shaping the experiences of the Asian community. By tracing the emotional dynamics conveyed in these videos, this study illuminates how TikTok functions as a site where racialised emotions both reproduce and resist racial hierarchies, thereby advancing scholarship on digital anti-racism and affective solidarity.

Objectives/Research Questions

How do the agency and regulation of racialised emotions manifest in TikTok videos responding to perceived anti-Asian racial discrimination?

Methods

This study analysed a sample of 164 TikTok videos featuring seven resistance-related hashtags (i.e. #stopasianhate, #asianlivesmatter, #stopasianhatecrimes, #asianawareness, #modelminoritymyth, #protectasianelders, and #asianrepresentation). We conducted a qualitative content analysis (Hsieh and Shannon, 2005) to code (1) outward-directed emotions expressed in resistance to anti-Asian racism and (2) forms of emotion regulation. Codes of emotion regulation were further differentiated between instances that resisted and those that embraced regulatory norms. Both coding schemes were applied concurrently to all videos, allowing a single video, for example, to simultaneously challenge response modulation, a form of emotion regulation, while expressing pride as an outward-directed emotion. In line with institutional ethical approval and best practices for researching racism-related computer-mediated communication (Lee and Lee, 2023), this study ensured the full anonymisation of all content creators. To further protect identities and minimise the risk of traceability, quoted textual elements (e.g. captions, text overlays, and narrations) were not reproduced verbatim, and the visual content of the videos was not described in full detail (McCosker and Gerrard, 2021).

Results

The findings demonstrate that TikTok operates as a complex affective space in which racialised emotions are neither uniformly liberatory nor consistently constraining. Instead, emotions function in multifaceted ways.

On the one hand, outward-directed emotional expressions, including righteous anger, sadness directed at the wrongness of racism, humour, pride, hope and solidarity, as well as practices that challenge emotion regulation, enable racialised subjects to break silence, contest dominant racial narratives, and articulate alternative subjectivities. These expressions are afforded and amplified by TikTok's platform infrastructure, which facilitates the visibility and circulation of affective responses to racial injustice.

On the other hand, the study also identified how emotion regulation operates as an internalised force within a racialised social order through practices such as response modulation, situation management, cognitive reappraisal, and attentional deployment. While these strategies may serve as coping mechanisms or forms of self-protection, they can simultaneously reproduce emotional norms that discipline racialised subjects into silence, endurance, or self-blame. Their circulation on TikTok may further contribute to the normalisation of these regulatory practices. There are also inherent limitations to leveraging emotions in digital anti-racism, including the volatility of the sense of community among participants in online activism and the potential ethical concerns surrounding the treatment of alleged perpetrators.

Future Work

Methodologically, it is first important to note that the emotions examined in this study are displayed, performed, and mediated, and are often short-lived, spontaneous, or one-off in nature. This characteristic may be viewed as a limitation of social media content analysis, as researchers cannot access narrators' underlying emotions as directly as those employing methods such as ethnography or interviews. However, since this study is concerned with the societal impact of emotional displays in these videos, it is appropriate for the researchers to adopt the position of viewers rather than focusing on the creators' production processes. Future research could explore the underlying reasons behind content creators' specific displays of emotion.

Second, this study treats emotional displays collectively rather than categorising them according to the narrators' nationality, ethnicity, generation, immigration status, or gender. This methodological choice does not imply that all Asians manage racialised emotions, or perceive and experience the racial order, in the same way. Rather, the study focuses on how the Asian community responds to racialised emotions in a collective manner. Given that different groups within the Asian community are likely to have distinct racialised experiences and respond differently to racialised emotions, it would be valuable for future research to examine these patterns in greater depth.

References

- Bonilla-Silva, E. (2019). Feeling race: Theorizing the racial economy of emotions. *American Sociological Review*, 84(1), 1–25.
<https://doi.org/10.1177/0003122418816958>

- Chon, M.-G. (2023). The role of social media in empowering activism: Testing the integrative model of activism to anti-Asian hate crimes. *Asian Journal of Communication*, 33(6), 511–528.
<https://doi.org/10.1080/01292986.2023.2251131>
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288.
<https://doi.org/10.1177/1049732305276687>
- Kim, J. Y. (2016). Racial emotions and the feeling of equality. *University of Colorado Law Review*, 87(2), 437–500.
- Lee, J. J., & Lee, J. (2023). #StopAsianHate on TikTok: Asian/American women’s space-making for spearheading counter-narratives and forming an ad hoc Asian community. *Social Media + Society*, January-March, 1–11.
<https://doi.org/10.1177/20563051231157598>
- McCosker, A., & Gerrard, Y. (2021). Hashtagging depression on Instagram: Towards a more inclusive mental health research methodology. *New Media & Society*, 23(7), 1899–1919. <https://doi.org/10.1177/1461444820921349>
- van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin*, 134(4), 504–535.
<https://doi.org/10.1037/0033-2909.134.4.504>
- Zhao, X., & Abidin, C. (2023). The “Fox Eye” challenge trend: Anti-racism work, platform affordances, and the vernacular of gesticular activism on TikTok. *Social Media + Society*, January-March, 1–16. <https://doi.org/10.1177/20563051231157590>

9D.2 Who owns #RealEnglish? Influencer Teachers' Construction of Authenticity on Instagram

Ali Fuad Selvi (University of Alabama), Erhan Aslan (University of Reading)

Introduction

Despite sustained (inter)disciplinary interest and engagement with the notion of authenticity across fields, authenticity remains a construct marked by conceptual ambiguity and theoretical fragmentation. This lack of consensus is particularly consequential in language education, where authenticity has been a powerful anchor point, carrying normative references, absolute values, and projected ideals across three interrelated reference points:

- 1) language, understood as a bounded, standardized, and often idealized linguistic system (e.g., “real,” “natural,” or “native-like” English);
- 2) language use, anchored in sociocultural contexts in which the ‘target’ language is imagined to be used ‘naturally,’ typically in the Anglosphere; and
- 3) language users, implicitly or explicitly positioned in relation to an idealized ‘native’ speaker norm, which functions as a proxy toward legitimacy, authority, and expertise (Lowe & Pinner, 2016).

These longstanding tensions are further intensified within contemporary networked, algorithmic, and platformized communicative environments. In such contexts, previously stable demarcations between authenticity and inauthenticity) are now increasingly blurred. Social media platforms do not merely host content; rather, they actively shape what counts as visible, valuable, and authentic through platformized affordances and algorithmic logics.

Within this shifting landscape, a growing number of ordinary (English) language teachers have emerged as self-entrepreneurial influencer teachers on platforms such as Instagram, TikTok, and YouTube. They leverage platform affordances and produce “recreational” (Chik & Ho, 2017) or “bite-sized” (Aslan, 2024) multimodal content to attract attention, sustain engagement, and foster a vast follower base (Selvi, 2025) and cultivate forms of “microcelebrity” status (Senft, 2013). Their content not only reflects but also actively constructs, performs, negotiates, and circulates authenticity claims regarding what counts as “real English,” how English should be used, and who can legitimately represent English and its speakers. These authenticity discourses are entangled with platform logics, neoliberal self-branding practices, and the globalized status of English, and therefore warrant critical scholarly scrutiny (Aslan & Selvi, under review).

Situated at the nexus of influencer culture, attention economy, and the “platformized ELT ecosystem” (Selvi, forthcoming), this exploratory qualitative study addresses the following question:

How do influencer language teachers on Instagram construct authenticity in relation to language, language use, and language users?

This study contributes to ongoing scholarly discussions by theorizing authenticity as a platformized, discursively constructed, and technologically mediated phenomenon within an emerging algorithmic, transnational, and neoliberal educational landscape

Methods

This exploratory qualitative study examines how transnational macro and mega-level social media influencer teachers construct authenticity in relation to the English language, language use, and language users through multimodal content on Instagram. Drawing on a content-analytical approach (Selvi, 2019), it conceptualizes authenticity as a discursively constructed, multimodally ensembled, and ideologically loaded phenomenon situated at the intersections of the attention economy, influencer culture, and language ideologies. Data were collected using a purposive, hashtag-based sampling strategy, consistent with established practices in social media research (Mayr & Weller, 2016). Instagram was particularly well-suited to this inquiry due to its global reach and prominence within influencer-driven content ecosystems, with over 2 billion active users worldwide (Statista, 2025). Seventeen authenticity-oriented hashtags (e.g., #RealEnglish, #EverydayEnglish, #EverydayEnglish) were identified through iterative platform searches. For each hashtag, the first 25 Reels returned by Instagram’s search function were collected and screened using the following inclusion criteria: (1) explicit framing of English, English use, or user as authentic, (2) English-language content, (3) production by macro- and mega-level influencer teachers, and (4) adopting multihastagging that reinforced authenticity positioning. After removing duplicate videos, the final dataset comprises 209 unique Reels, yielding a transnational corpus.

Data analysis was informed by social semiotics (Kress & van Leeuwen, 2020), foregrounding multimodality and sign-maker agency. An iterative coding process yielded a tripartite analytical framework encompassing agents (identity representations and positioning practices), language (linguistic varieties, features, and functions), and platformized multimodality (modes, digital affordances, communicative spaces, and intertextuality). This article reports findings from the agents and language categories.

Results

Regarding language, linguistic variety is largely homogenized, with American English functioning as an unmarked and naturalized norm, indexed symbolically (through flags, accent imitation, and pop culture references) rather than explicitly named.

Regarding language use, lexical choices featured in influencer content foreground informality, idiomaticity, and colloquial expressions, reinforcing an ideology of authenticity grounded in everyday usage and social immersion. Grammatical features receive comparatively little explicit attention and are often framed contrastively, privileging fluency and pragmatic effect over formal accuracy. Phonological features are mobilized performatively through accent imitation and stylization, operating as semiotic shortcuts to authenticity and cultural belonging rather than as objects of instructional discourse. Register is consistently organized around a binary opposition between formal or “school” language and informal or sociolinguistically marked “street” language, with the latter positioned as flexible, socially effective, and desirable. Across the dataset, communicative functions are predominantly demonstrative and performative, with language presented as content to be showcased, imitated, and evaluated rather than taught.

Regarding language users, the data highlights a recurrent pattern in which speakers are framed through simplified and contrastive roles, most prominently native versus non-native speakers, experts versus novices, and “real” language users versus institutional learners. Influencers consistently position themselves as legitimate, embodied authorities on language use, while pedagogical or classroom-based forms of English are implicitly devalued.

Future Work

These findings have broader implications for the understanding of authentic language, expertise, and learning in social media environments. First, they illustrate a shift in linguistic authority, whereby influencers are positioned as legitimate language experts through embodied, performative displays rather than institutional credentials. This reconfiguration challenges traditional models of expertise and reshapes how language knowledge is evaluated and circulated online. Second, the implicit normalization of American English as a global default reinforces existing linguistic hierarchies, marginalizing alternative varieties and limiting representations of linguistic diversity. Third, the prioritization of informality, fluency, and affective performance over grammatical explanation reflects an ideological shift in what it means to “know” a language in digital spaces. Language competence is framed less in terms of structural knowledge and more in terms of sounding natural, confident, and socially aligned, which may reduce affective barriers for learners while obscuring the complexity of language use. That being said, deterministic approximations and idealizations about a single variety and an idealized native speaker of English may also induce anxiety among learners. In addition, the idealized treatment of phonology and register undermines sociolinguistic variation, turning accents and styles into recognizable semiotic cues rather than context-sensitive practices. Taken together, these implications suggest that social media not only mediate language learning content but also actively shape

prevailing ideologies of authenticity, expertise, and legitimacy in contemporary digital communication.

References

- Aslan, E. & Selvi, A. F. (under review). #RealEnglish: How do influencer teachers curate authenticity through platformized multimodality on Instagram?
- Aslan, E. (2024). Bite-sized language teaching in the digital wild: Relational pedagogy and micro-celebrity English teachers on Instagram. *System*, 121. <https://doi.org/10.1016/j.system.2024.103238>
- Chik, A., & Ho, J. (2017). Learn a language for free: Recreational learning among adults. *System*, 69, 162-171. <https://doi.org/10.1016/j.system.2017.07.017>
- Kress, G. R., & van Leeuwen, T. (2020). *Reading images: The grammar of visual design*. Routledge.
- Lowe, R. J., & Pinner, R. (2016). Finding the connections between native-speakerism and authenticity. *Applied Linguistics Review*, 7(1), 27-52.
- Mayr, P., & Weller, K. (2016). Think before you collect: Setting up a data collection approach for social media studies. In L. Sloan & A. Quan-Haase (Eds.), *The SAGE handbook of social media research methods* (pp. 107-124). SAGE Publications. <https://doi.org/10.4135/9781473983847.n8>
- Selvi, A. F. (forthcoming). The platformized English Language Teaching ecosystem: Navigating professional opportunities, challenges, and the future of the profession. *TESOL Quarterly*.
- Selvi, A. F. (2019). Qualitative content analysis. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 440-452). Routledge. <https://doi.org/10.4324/9780367824471-37>
- Selvi, A. F. (2025). Teachers as influencers and personal brands on Instagram: Double-tapping language ideologies in the attention economy. *Applied Linguistics*. <https://doi.org/10.1093/applin/amaf039>
- Senft, T. M. (2013). Microcelebrity and the branded self. In J. Hartley, J. Burgess, & A. Bruns (Eds.), *A companion to new media dynamics*, (pp. 346-354). Blackwell Publishing. <https://doi.org/10.1002/9781118321607.ch22>
- Statista. (February 5, 2025). Most popular social networks worldwide as of February 2025, by number of monthly active users (in millions) [Graph]. Statista. <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>

9D.3 Seeing what your friends are up to: FOMO through group chats

Carolyn Lehmann, Sonja Utz (Leibniz-Institut für Wissensmedien (IWM) Tübingen)

Introduction

Many long-distance friends rely on group chats on messengers like WhatsApp to stay in touch. As asynchronous and multimodal channels, they enable continued contact in a lightweight, low-effort manner. Yet, research remains inconclusive on whether group chats evoke genuine feelings of connectedness. Critics argue that short, often mundane messages lacking contextual cues hamper nuanced expressions of emotions (Yus, 2017). However, passively reading friends' updates can create ambient awareness and closeness (Yus, 2017). Thus, the question remains: Do group chats transmit connectedness, and what psychological processes mediate this feeling?

Group chats may also have adverse effects. Even though the fear of missing out (FOMO) is still mainly associated with social media platforms, there is evidence that certain scenarios in messenger communication can elicit FOMO as well ((Marengo et al., 2021)). A potent psychological driver of FOMO is the perception of missed opportunities for bonding (Rifkin et al., 2024) and a fear of exclusion (Hefner et al., 2018). Accordingly, if long-distance friends miss out on a group activity due to the distance, they may experience FOMO. However, FOMO levels may depend on the rationale individuals provide for why they missed the activity: Lutz et al. (2024) showed that the effects of non-participation vary depending on whether it is self-caused or other-caused. Subjects in their study felt their needs were more threatened when they hadn't been invited (other-exclusion) than when they had chosen not to participate (self-exclusion). Applied to long-distance friendships, this suggests that individuals who moved away themselves (self-exclusion) experience less FOMO than those whose friends moved away (other-exclusion). When non-participation is attributed to friends rather than to one's own decisions or external circumstances, feelings of exclusion and FOMO become more likely.

Group chat messages about group activities that some individuals have not attended offer room for interpretation. Such messages can be attributed to intentions of inclusion: Absent members may interpret them as a sign that the group is keeping them updated, creating ambient awareness. Conversely, the same messages can be interpreted as a reflection that the group has disregarded its distant members and carelessly posted messages in which they are not involved. When applying this attribution, individuals may feel excluded.

A scenario that is conceptually related to other-exclusion is discovering that friends have a separate subgroup chat. Although widespread, subgroup chats can undermine group cohesion (Mannell, 2020). Feelings of exclusion intensify when subgroup chats are used to plan joint activities that omit non-members. In contrast to spontaneous friend meetups, where exclusion appears more accidental, the deliberate creation of a subgroup chat can be seen as intentional exclusion. Intentional exclusion depletes needs and should amplify FOMO.

Objectives/Research Questions

In two vignette experiments, we tested how messages about missed activities influence connectedness and FOMO among long-distance friends.

Study 1 assessed whether connectedness and FOMO vary depending on subjective interpretations of why participants missed out, i.e., whether it was them (self-exclusion) or their friends (other-exclusion) who had moved away and separated the group.

H1: Experiencing friends moving away is associated with a) more FOMO and b) lower connectedness than moving away oneself.

Additionally, Study 1 examined the role that message attributions play in shaping emotional responses. We assessed the extent to which messages about missed activities were attributed to 1) friends' efforts at including absent members and 2) their disregard, i.e., that they have forgotten about absent members.

H2: The effect of experiencing friends moving away (vs. moving away oneself) on a) FOMO and b) connectedness is mediated by attributions.

Study 2 tested the impact of a special form of exclusion, the formation of subgroup chats. In the study scenario, participants were told either that the activity they missed out on was organized in a separate subgroup chat among their friends (intentional exclusion) or that it emerged from a spontaneous meeting of two friends (accidental exclusion).

H1: Missing activities organized through subgroup chats is associated with a) more FOMO and b) lower connectedness than missing chance encounters.

Study 2 also assessed whether FOMO depended on the missed activity's attractiveness: Given that FOMO involves anticipated regret and envy toward others having fun (Tandon et al., 2021), rewarding experiences should be more likely to evoke FOMO than unattractive ones.

H2: Missing attractive experiences is associated with higher FOMO than missing unattractive experiences.

RQ1: Is there an interaction effect of intentionality of exclusion and experience attractiveness on FOMO?

Methods

In Study 1, a vignette experiment with two conditions, participants were asked to imagine having a geographically dispersed group of friends communicating through a group chat. They imagined receiving messages about a group gathering, which they couldn't attend due to distance. In the 'stayed' condition, all group members except the participant had left their hometown; in the 'left' condition, the participant was the only one who had moved away. The study was preregistered under

https://osf.io/j69mz/overview?view_only=6fb1b87870bc4cc19de5ac1aaabdd0ee

Study 2 involved a 2 (exclusion: intentional/unintentional) x 2 (attractiveness: high/low) vignette experiment. Here, participants imagined receiving messages about a group activity they had not known about. Participants were either told that their friends had instigated the activity during a chance meeting (unintentional) or via a second group chat (intentional). Attractiveness was manipulated by telling participants that they either missed a video call or a pub night with their friends. The study was preregistered under https://osf.io/qp4db/overview?view_only=e4e0ebf85d6d404f92a5c2c7364aebe3

Results

In both studies, the scenario of receiving group messages about missed activities was rated as realistic.

Study 1

H1: One-tailed independent t-tests showed that those who imagined having moved had significantly less FOMO ($M = 3.05$, $SD = 0.89$) than those who imagined having stayed, ($M = 3.18$, $SD = 0.89$), $t(624) = 1.78$, $p = .04$, $d = .14$. Connectedness did not differ between those who imagined having stayed ($M = 2.66$, $SD = 0.8$) and those who imagined having left ($M = 2.73$, $SD = 0.8$), $t(624) = -1.11$, $p = .13$.

To test H2, we ran mediation models using Hayes' (2017) PROCESS macro model number 4, with attribution to inclusion as mediator 1 and attribution to disregard as mediator 2. See the appendix for mediation path diagrams.

For FOMO, there was a significant total effect of the manipulation, but no significant direct effect. A significant indirect effect via attribution to disregard was found, but not via attribution to inclusion. When friends move away, individuals may feel uncertain about the continuity of the friendship. This, in turn, may make them more sensitive to exclusionary cues, and this sensitivity to exclusion fosters FOMO.

For connectedness, no significant total effect emerged. There was a significant indirect effect of the manipulation ('who moved away') via attribution to disregard, but not via attribution to inclusion.

Thus, connectedness was not directly influenced by who had relocated. Yet, experiencing friends' relocations led individuals to attribute messages to disregard, and this attribution tendency reduced their connectedness.

Study 2

H1a): A two-factorial ANOVA revealed a main effect of intentionality on FOMO, $F(1, 391) = 13.51$, $p < .001$, with a small effect size, $\eta^2_p = .03$. FOMO was higher when the missed activity was organized through a subgroup chat ($M = 3.47$, $SD = 1$) than when organized at a chance encounter ($M = 3.09$, $SD = 1.04$). FOMO is amplified when friends organize activities in a separate group chat, likely due to feelings of exclusion.

H1b): In the ANOVA, intentional exclusion was associated with less connectedness than unintentional exclusion, $F(1, 391) = 11.52$, $p < .001$. The effect was small, $\eta^2_p = .03$. When the group activity was instigated in a subgroup chat (intentional), participants felt less connected ($M = 2.83$, $SD = 0.58$) than when it came about through a chance encounter (unintentional; $M = 3.03$, $SD = 0.61$). Likely, hearing about others' subgroup chats causes feelings of exclusion, which diminish connectedness.

For H2, attractiveness had a significant positive effect on FOMO, $b = 0.34$, $F(1, 391) = 45.07$, $p < .001$, indicating that activities perceived as attractive are associated with higher FOMO. The effect size was medium, $\eta^2_p = .11$. This is in line with the definition of FOMO as centered on rewarding experiences.

RQ1: The interaction term in the ANOVA was not significant, $F(1, 391) = 0.32$, $p = .57$. Thus, exclusion intentionality and activity attractiveness independently influence FOMO.

Future Work

Future studies may inquire into the psychological predictors of the attribution effects identified, i.e., what factors motivate individuals to employ benevolent versus maladaptive message attributions. Potential predictors may be rooted in personality, e.g., attachment security (Holte et al., 2024), or in group norms, e.g., the tendency to be permanently connected (Knop-Huelss et al., 2017).

Subgroup chats not only evoke feelings of social exclusion but can also create perceptions of information exclusion. Both 'out-of-the-loop' experiences and feelings of social exclusion foster FOMO (Hefner et al., 2018; Nesbit & Lole, 2025). To systematically investigate their relative impact on FOMO, experimental designs could manipulate missed information, e.g., study-related updates in a group of students) versus missed social interactions (activities organized in the group).

References

- Hayes, A. F. (2017). *Introduction to Mediation, Moderation, and Conditional Process Analysis, Second Edition: A Regression-Based Approach*. Guilford Publications.
- Hefner, D., Knop, K., & Vorderer, P. (2018). “I wanna be in the loop!” – The role of Fear of Missing Out (FoMO) for the quantity and quality of young adolescents’ mobile phone use: Current perspectives on media use and effects (pp. 39–54).
<https://doi.org/10.5771/9783845280455-39>
- Holte, A. J., Nixon, A., & Cooper, J. (2024). Attachment anxiety mediates the relationship of need to belong and Fear of Missing Out (FoMO). *Discover Psychology*, 4(1), 141. <https://doi.org/10.1007/s44202-024-00250-2>
- Knop-Huelss, K., Winkler, J. R., & Penzel, J. (2017). Being POPC together: Permanent connectedness and group dynamics. In *Permanently Online, Permanently Connected*. Routledge.
<https://www.taylorfrancis.com/chapters/edit/10.4324/9781315276472-13/being-popc-together-katharina-knop-huelss-julia-winkler-jana-penzel>
- Lutz, S., Büttner, C. M., & Neumann, D. (2024). A window to what we missed: Effects of self- versus other-exclusion on social media users’ fundamental needs, emotional responses, and online coping behaviors. *Media Psychology*, 27(3), 401–427. <https://doi.org/10.1080/15213269.2023.2242769>
- Mannell, K. (2020). Plural and porous: Reconceptualizing the boundaries of mobile messaging group chats. *Journal of Computer-Mediated Communication*, 25(4), 274–290. <https://doi.org/10.1093/jcmc/zmaa007>
- Marengo, D., Settanni, M., Fabris, M. A., & Longobardi, C. (2021). Alone, together: Fear of missing out mediates the link between peer exclusion in WhatsApp classmate groups and psychological adjustment in early-adolescent teens. *Journal of Social and Personal Relationships*, 38(4), 1371–1379.
<https://doi.org/10.1177/0265407521991917>
- Nesbit, T., & Lole, L. (2025). A qualitative investigation of social media fear of missing out (FoMO) and needs fulfilment among young adults. *International Journal of Human–Computer Interaction*, 0(0), 1–17.
<https://doi.org/10.1080/10447318.2025.2543998>
- Rifkin, J. R., Chan, C., & Kahn, B. E. (2024). Anxiety about the social consequences of missed group experiences intensifies fear of missing out (FOMO). *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/pspa0000418>
- Tandon, A., Dhir, A., Talwar, S., Kaur, P., & Mäntymäki, M. (2021). Dark consequences of social media-induced fear of missing out (FoMO): Social media stalking,

comparisons, and fatigue. *Technological Forecasting and Social Change*, 171, 120931. <https://doi.org/10.1016/j.techfore.2021.120931>

Yus, F. (2017). Contextual constraints and non-propositional effects in WhatsApp communication. *Journal of Pragmatics*, 114, 66–86.
<https://doi.org/10.1016/j.pragma.2017.04.003>

9D.4 For a Theory of Onlife Practices: Studying Media with Bourdieu to Interpret Practices in the Interaction between Autochthonous Contexts and Digital Platforms

Roberto Graziano (Università degli studi di Napoli Federico II)

Introduction

What tools do we, as social scientists, have to understand and interpret social action and cultural phenomena that are continuously mediated by digital devices? How can we analyse the complex mechanism of practices, consumption, style, taste, and aesthetic dispositions in the onlife infosphere, if when we act we already take into account, within our embodied and “unconscious” practical sense, participation in the network? In order to grasp practical sense in the onlife infosphere, I argue that once again Bourdieu’s sociology (1978, 1980, 1982, 2002) can contribute to understanding such a delicate moment, which pushes us to revise the methods of sociological inquiry (Cohen & Wartofsky, 2013). Therefore, in this epistemological reflection, I seek to bring into dialogue the key concepts of Bourdieusian sociology, considered together as analytical tools for understanding complex digitised society. Too often, against Bourdieu’s intentions (Lindell, et al. 2025), his analytical concepts have been fragmented and analysed in isolation, also because of the complexity of each of them (Paulucci, 2014), but this partiality prevents a comprehensive reading and favours myopic deterministic interpretations (Gerli 2021). One of the best-known cases is the individualised interpretative reading of the concept of habitus, extracted from the complex mechanism that led Bourdieu to formulate the theory of practice, which instead requires habitus to be intrinsically connected to capital and field (Gerli 2021). In reality, Bourdieu’s sociology, as shown by Passeron in the book dedicated to Bourdieu’s death *Death of a Friend, Disappearance of a Thinker* and by Wacquant in *Bourdieu Goes to the City* (2018), should be interpreted as a comprehensive sociology of domination, aimed at unveiling the misrecognised dynamics of domination.

Objectives/Research Questions

The paper pursues four main objectives. First, it aims to employ and adapt Pierre Bourdieu’s theory of practice to the study of digital platforms, exploring how social practices increasingly unfold in hybrid configurations between autochthonous contexts and algorithmically mediated environments. Second, it seeks to highlight how habitus, capital, and field can be fruitfully analysed in a relational way within media and platform studies, as an alternative to fragmented or overly deterministic readings. Third, it revisits the concept of field in the context of platformization, analysing digital platforms as micro-fields endowed with specific rules while remaining structurally

connected to digital capitalism and to the social contexts in which contents are produced and interpreted. Finally, the paper contributes to the development of an interpretive perspective on onlife practices, with particular attention to taste regimes and the construction of masculinity on platforms, offering conceptual tools to support the interpretation of trends, virality, and symbolic power in digital environments.

Methods

The article is organised into three sections. In the first, I rework the results of a previous study (Graziano, 2025), in continuity with Wacquant (2010), proposing the use of habitus not only as an object of investigation but also as a methodological tool, reintroducing it into qualitative digital sociology. I show that this concept can be used as an analytical instrument to orient the selection of cases in digital empirical research. Taking the concept of habitus into account, in relation to the scheme $[(\text{habitus}) \cdot (\text{capital}) + \text{field}] = \text{practice}$, I used this tool for selecting the cases of analysis.

However, aware that habitus is a relational tool capable of linking subjective reality with objectified reality, and that, according to one of the last definitions provided by Bourdieu in *The Weight of the World*, habitus is the acquisition of forms of capital that exist in an embodied state (Bourdieu et al. 1993), I updated the classical scheme outlined in *Distinction* (1979) by introducing new forms of capital, used as indirect indicators to describe and categorise subjects and their hybrid interaction on digital platforms and in autochthonous contexts. For this reason, the study integrates new forms of capital with the classical ones identified by Bourdieu, in particular erotic capital (Hakim, 2010) and autochthonous capital (Chamboredon 1982). In this section I explain the reasons for this integration.

In the second section, I reintegrate and deepen the concept of field, developed in previous work, as a methodological conceptual tool for media studies, namely as a relatively autonomous microcosm, with its own rules, in which individuals compete through the different forms of capital they possess. In this sense, I analyse platforms individually, for example, TikTok, as a micro-field within the broader field of platforms in general, as suggested by van Dijck et al. (2016), and I methodologically select the object of analysis in a strongly sectorialised way. As I show in the concluding section, I give the example of studying consumption on TikTok by focusing on gourmand food as a sectorial category, rather than food in general. In this way, even if the definition of field may seem misleading in interpreting reality, given that nothing is completely autonomous, it acquires methodological meaning because it allows us to sectorialise the object of study. Field is a key concept that links individuals' action, through habitus, to social structures, showing how social practices result from the interaction between agents' embodied dispositions and the objective conditions of a specific arena. Given the complexity also of a digitised social arena, this aspect pushes us to elaborate a new theory of onlife practices, updated to the complex dynamics of contemporary society.

Results

To understand practices in the age of algorithmic hegemony we must take into account some elements identified by post-Bourdieuian sociology, renewing the formula of practice as a methodological tool for reading complex cultural phenomena. Thus, the classical formula identified by Bourdieu, $(\text{habitus})(\text{capital}) + \text{field} = \text{practice}$, is made more complex in relation to the pervasiveness of the digital in everyday life.

In this sense, for a theory of onlife practice, in which digital and physical experience merge inseparably, practices of consumption between online and offline and modes of behaviour shaped by digital life can be read as the interaction between habitus, the embodied history of subjects who bring online their class history, their dispositions to act, their tastes, and machine habitus, the habitus of algorithms, a historical memory of culture incorporated into the technological devices that underlie digital platforms.

Habitus is the acquisition of forms of capital that exist in an embodied state, while forms of capital are the resources that subjects acquire or inherit in order to move within specific arenas. However, the classical forms identified by Bourdieu are necessary but not sufficient to operate in the complex digitised society. Therefore, other forms of capital are required in the onlife arena. The principles of capital, convertibility, acquisition, and exchange also play a central role in this context. What is needed is therefore a capital elevated to the digital ($\text{capital}^{\text{digital}}$), or, more conceptually, capital in its hybrid interaction transformed, enhanced, and reconfigured by the digital, that is, the set of classical and new forms of capital in the form they take when they are reconfigured and made operational in platformised fields. But what are platformised fields? They are microcosms with relatively autonomous rules, such as platform layouts and design, which are inseparably connected to the autochthonous contexts in which digital contents are created and uploaded. In the literature this relational phenomenon between the autochthonous and the digital is known as the infosphere. This complex mechanism that indissolubly connects autochthonous reality with digital reality exemplifies onlife practices. The classical formula is thus articulated as:

$$[(\text{habitus} + \text{machine habitus}) \times (\text{capital}^{\text{digital}})] + (\text{field} + \text{digital field}) = \text{Onlife Practice}$$

This formula can be used as an interpretive key for understanding the formation of trends and their memetic reproduction, highlighting the sociological dimensions that precede digital technologies. This complex mechanism between habitus, structures, and capital provides an explanation of practices exactly as Bourdieu used it to understand how subjects are pushed to behave in regulated ways, showing regularity even when their behaviour, in most cases, does not obey formalised and codified rules present in their consciousness.

Future Work

Clearly, this rereading of the formula of practice inherited from Bourdieu's sociological work does not aim to become a deterministic theory capable of explaining every practice or trend that spreads on platforms. The purpose of synthesising this complex scheme is to offer a socially determined interpretive key for practices between the digital and autochthonous reality. It is a tool for interpreting social action shaped by cultural positioning, by the culture learned by algorithms, and by forms of capital operating within digital fields of struggle.

In the third section, after describing in detail the reworking of this formula, I highlight its analytical usefulness for the sociological investigation of the digital by referring, on the one hand, to secondary sources that have employed Bourdieusian theory to study online practices and, on the other, to two illustrative analytical cases, which will be discussed more extensively during the conference presentation. These cases concern, respectively, the study of consumption practices related to popular gourmand gastronomy and the analysis of the construction of youth masculinity in urban peripheries through TikTok.

In conclusion, in discussing these cases, I emphasise how the analysis of practical sense emerging from the interaction between autochthonous contexts and digital platforms inevitably leads to the interpretation of doxa (Cristante, 2014) and illusio that guarantee the virality of certain digital trends. This points back to symbolic power as a key element in the perpetuation of dynamics of domination, also concealed within the algorithmic hegemony of platforms.

References

- Bourdieu, P. (1979). *Distinction: A social critique of the judgement of taste*. Cambridge, MA: Harvard University Press.
- Bourdieu, P., Accardo, A., Balazs, G., Beaud, S., Bonvin, F., Bourdieu, E., ... Wacquant, L. (1993). *The weight of the world: Social suffering in contemporary society*. Stanford, CA: Stanford University Press.
- Chamboredon, J.-C. (1982). La diffusion de la chasse et la transformation des usages sociaux de l'espace rural. *Études rurales*, (87–88), 83–106.
- Cohen, R. S., & Wartofsky, M. W. (2013). *Models: Representation and the scientific understanding*. Dordrecht: Springer.
- Cristante, S. (2014). Dalla sociologia dell'opinione pubblica alla mediologia della doxasfera. *Redes.com: Revista de estudios para el desarrollo social de la comunicación*, (10), 161–183.

- Gerli, M. (2021). Tra rifrazione e traduzioni: Bourdieu, Latour e l'autonomia scientifica. *Sociologia e ricerca sociale*, 125(2), 5–26.
- Hakim, C. (2010). Erotic capital. *European Sociological Review*, 26(5), 499–518.
<https://doi.org/10.1093/esr/jcq014>
- Paolucci, G. (2014). *Introduzione a Bourdieu*. Roma Bari: Gius. Laterza & Figli.
- Passeron, J.-C. *La mort d'un ami, disparition d'un penseur*.
(Anno non indicato perché non fornito e non verificabile senza rischio di errore.)
- van Dijck, J., Poell, T., & De Waal, M. (2016). *De platformsamenleving: Strijd om publieke waarden in een online wereld*. Amsterdam: Amsterdam University Press.
- Wacquant, L. (2010). Habitus as topic and tool: Reflections on becoming a prizefighter. *Qualitative Research in Psychology*, 8(1), 81–92.
- Wacquant, L. (2018). *Bourdieu comes to town: On habitus, field, and symbolic power*. Cambridge: Polity Press.

Session 9E - Platform Design and Politics

Day 2 · Tues 14 July · 15:35-17:05, 225, ARC

9E.1 Decoding algorithmic suppression: How digital activists uncover and navigate platform censorship in Finland?

Sanna Malinen (Tampere University)

Introduction

Social media has afforded political influencers and activists an unprecedented level of visibility, thereby enhancing their ability to amplify political messages. Recently, scholars have pointed out how popular social media platforms have shifted from fostering citizen participation to intensifying algorithmic moderation of political content (Elmimouni et al., 2025; Szabo, 2025; Riedl et al., 2024). When platforms such as TikTok and Instagram exert their automated moderation practices globally, they have created new and unforeseen constraints on freedom of speech for political participation in democratic societies. Paradoxically, activists today can protest in the streets while expressing the same views on social media may lead to censorship (Earl et al., 2022). Building on social media's dual role as both a facilitator and a controller of public discourse (Elmimouni et al., 2025), this ongoing study explores political censorship in an atypical context, in a Nordic high-trust society, identifying topics that are considered sensitive to social media algorithms and discussing future implications of censorship for digital activism.

Similar to commercial social media influencers, the success of political content creators depends on the visibility granted by platform algorithms. Consequently, understanding how these algorithms function has become a central aspect of activism on social media (Cotter, 2021). The increasing preoccupation of activists with algorithms and the subsequent adaptation of their communication practices to align with these algorithmic logics has likewise received scholarly criticism. Social media is encouraging self-centred forms of participation and weakening collective action, this way reinforcing individualization, depoliticization, and commodification of messages and engagement (Malafaia & Meriluoto, 2024).

Sometimes activists may operate in hostile contexts, where simply being visible online can become a form of activism, as in the case of queer activism in Hungary (Szabo,

2025). This way, any deliberately targeted activities to silence or limit activists' reach can be viewed as censorship (Earl et al., 2022). In recent years, social media moderation has increasingly shifted to techniques called "reduction of visibility" through machine learning tools, rendering platform governance a highly politicized issue (Gillespie, 2022). Although privately owned social media platforms have the right to control their content, they have no interest in protecting users' freedom of expression and assembly, resulting in preventing, reducing, and controlling social movements and activists (Earl et al., 2022). Prior studies indicate that groups with marginalized identities, namely individuals coming from racial and sexual minorities, are disproportionately suppressed by platforms (Szabo, 2025; Delmonaco et al., 2024). This leads to amplification of the voices that are already overrepresented, thereby deepening social inequities.

Objectives/Research Questions

To understand social media users' perceptions of and reactions to platform censorship, qualitative interviews were conducted with active political content creators in Finland, and the following research questions were formulated:

- (1) How does censorship manifest itself on popular social media platforms?
- (2) What types of content are subject to censorship in the Finnish context?
- (3) What kinds of explanations do they themselves offer for content removal or restrictions on their activity?
- (4) How does censorship shape political activity on social media?

Methods

This work-in-progress study relies on two types of data: observation of political content on social media (Instagram, TikTok, X) and 15 interviews with political content creators who have public social media accounts. First, persons who had talked about censorship on social media were contacted via the researcher's social media account or email, if the email address was public. Then, potential interviewees were identified on Instagram, TikTok, and X using various hashtags and search terms such as "censorship", "activism", "protest", and "elections", in addition to various political topics. A total of 30 social media accounts representing a wide range of political and societal topics, including war in Palestine, sexual and gender minorities, disability rights, environmental protection, climate change, animal rights, health politics, feminism, antiracism, anticapitalism, and contemporary national and geopolitical issues, were contacted and asked if they had observed any censorship. Fifteen of them responded, reporting about experiences of political censorship on social media—either personally or directed at others—and agreed to participate in the study. The

participants used several platforms, with Instagram being the most popular site for activism and the most frequently mentioned in relation to censorship.

Most of the interviewees were female (11), two identified as male, and two as non-binary; their ages varied from 25 to 56 years.

Before starting the data collection, the university's research data experts were consulted to ensure the ethical and legal accuracy of the documents and procedures. All interviews were audio recorded, transcribed, and coded using MAXQDA qualitative coding software. The coding scheme followed the Grounded Theory approach, aiming to construct new theories or explanations in a data-driven manner, rather than starting with hypotheses (Corbin & Strauss, 2014).

Results

Platform censorship manifests as the silent, fast, and unaccountable removal of content or the limited visibility of content and accounts. In this study, censorship was most often targeted at human rights-related content; the most censorship-triggering topics were Palestine and LGBTQ+ rights. Among other censored topics mentioned were anti-vaccine views, domestic violence, sex work, and climate change denialism.

The findings confirm that, like commercial influencers, activists must adjust to platforms' modus operandi to optimize their visibility and avoid censorship. They have adopted visual, influencer-like modes of communication that blur boundaries between platform-friendly and activist content. Activists must continuously invent new communication strategies to maintain their visibility and popularity, as AI tools quickly learn to identify political content. They create various ways of circumventing censorship through collective algorithmic reasoning. For instance, by frequently posting non-political, visually pleasing content to attract an audience, taking breaks to reset the algorithm, mixing political and lifestyle content in the same video, boosting each other's content, and replacing censored words and images with neutral ones.

This study reveals a deep conflict of values between activists and platforms, and a strong distrust towards platforms, which continues to be strengthened by censorship. Activists explain censorship by commercial values that also drive moderation: serious political content and conflicts do not fit the platforms' preference for light, feel-good material. Because digital activism takes place in the intersection of commercial and political discourses, underlining the entertaining and consumeristic lifestyle content, activists often struggle with trying to fit their message. The findings reveal an ongoing trade-off between platform-friendly and serious political content on social media.

Because non-normative and minority identities are more likely to be suppressed, content on queer rights is more heavily suppressed than, e.g., economic topics or environmental activism. These findings underline the volatility inherent in globally

operating, privately owned platforms and questions the sustainability of digital activism that relies upon them. Platforms can continuously alter their policies without transparent communication, and such changes are closely tied to broader political shifts. The activists currently lack realistic alternatives to platforms that are overwhelmingly dominant in terms of visibility and audience reach. Should censorship render participation on these platforms impossible, they would be compelled to revert to traditional forms of protest and more private modes of discussion. To conclude, this study shows that today's digital activism has become increasingly vulnerable to platforms' political and commercial steering.

Future Work

These preliminary findings align with a large body of research on shadowbanning, confirming that certain topics are systematically suppressed on social media, and in this study, particularly on Meta. This study was motivated by limited research on digital repression enacted by privately owned online spaces in democratic contexts. The findings here point out how censorship can create new tensions in different societal contexts; for example, in Finland, the issues censored by Meta can be discussed publicly and are not regulated in public discourse. While prior work has extensively discussed the facilitating role of social media in activism, less is known about its more recent role as a suppressor of participation. As global social media platforms have grown in importance for activism, more research is needed on how they shape political participation in different types of societies. These findings demonstrate how activists navigate the conflicting demands of visibility and avoidance of moderation, while trying to maintain their authenticity and credibility in an environment that prefers non-political content.

Similar to other studies on moderation and particularly shadowbanning, this study reflects on the participants' understanding of what kind of content is suppressed and what gets promoted on social media. To gain a more comprehensive view of today's political censorship, we need more cross-platform research and methods that enable the analysis of large-scale data. Finally, in a qualitative research design like the present study, the sample is small and self-selected, and some topics and political groups were likely left out. A larger sample and longitudinal data would enable a greater variety of political topics and activist groups to be analyzed.

References

- Corbin, J., & Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Los Angeles: Sage.
- Cotter, K. (2021). "Shadowbanning is not a thing": black box gaslighting and the power to independently know and credibly critique algorithms. *Information, Communication & Society*, 26(6), 1226–1243.

- Delmonaco, D., Mayworm, S., Thach, H., Guberman, J., Augusta, A., & Haimson, O. L. (2024). "What are you doing, TikTok?": How Marginalized Social Media Users Perceive, Theorize, and "Prove" Shadowbanning. *Proceedings of the ACM on Human-Computer Interaction*, 8(CSCW1), 1-39.
- Earl, J., Maher, T. V., & Pan, J. (2022). The digital repression of social movements, protest, and activism: A synthetic review. *Science advances*, 8(10), eabl8198.
- Elmimouni, H., Rüller, S., Aal, K., Skop, Y., Abokhodair, N., Wulf, V., & Tolmie, P. (2025). Exploring Algorithmic Resistance: Responses to Social Media Censorship in Activism. *Proceedings of the ACM on Human-Computer Interaction*, 9(2), 1-24.
- Gillespie, T. (2022). Do not recommend? Reduction as a form of content moderation. *Social Media+ Society*, 8(3), 20563051221117552.
- Malafaia, C., & Meriluoto, T. (2024). Making a deal with the devil? Portuguese and Finnish activists' everyday negotiations on the value of social media. *Social Movement Studies*, 23(2), 190-206.
- Riedl, M. J., Martin, Z. C., & Woolley, S. C. (2024). 'I get suppressed:' pro-and anti-abortion activists' folk theories of platform governance and shadowbanning. *Information, Communication & Society*, 27(11), 2153-2170.
- Szabó, H. D. (2025). LGBTQ+ TikTok creators: Strategic visibility negotiation in commercial and activist online spaces. *Journal of Digital Social Research*, 7(2), 62-96.

9E.2 Immigration-Related Community Notes on X: Topic Structure and Evaluation Dynamics

Junyuan Lin, Anatoliy Gruzd (Toronto Metropolitan University)

Introduction

Social media platforms have become central infrastructures for public communication, political debate, and information consumption in the 21st century (Chadwick & Stanyer, 2022; Hossain, 2019; Quan-Haase & Young, 2010). While these platforms enable rapid information exchange, they also facilitate the large-scale diffusion of misinformation. Prior research shows that false or misleading information spreads faster and more widely than accurate content, driven by platform affordances such as algorithmic ranking, engagement-based visibility, and low barriers to content production (Chadwick & Stanyer, 2022; Vosoughi et al., 2018). These dynamics challenge traditional fact-checking approaches, which struggle to operate at the speed and scale of online information flows.

In response, platforms have increasingly adopted participatory and technologically mediated approaches to misinformation correction. One prominent example is Community Notes (CNs; formerly Birdwatch) on X, a crowdsourced fact-checking system that allows approved contributors to add contextual notes to potentially misleading posts (Pröllochs, 2022). Notes are evaluated through a rating process that incorporates feedback from contributors with diverse perspectives, with algorithmic thresholds determining whether a note is publicly displayed (Wojcik et al., 2022). While prior studies have examined contributors' behavior, evaluation patterns, and the accuracy of CNs in the aggregate, there is a lack of research exploring how misinformation is evaluated across particularly contentious policy areas (Allen et al., 2022; Drolsbach & Pröllochs, 2022).

Immigration represents one such highly contested domain. This is because immigration-related misinformation is embedded in political ideology and national identity, and frequently involves complex claims about legal status, asylum procedures, crime, and welfare access (Eberl et al., 2018). As a result, immigration-related misinformation is difficult to evaluate, even when corrective information is available (e.g., professional fact-check).

Objectives/Research Questions

The primary objective of this study is to examine whether and how topical and thematic context is associated with differences in the outcome and timing of evaluations of CNs dealing with immigration-related misinformation on X.

Accordingly, this study addresses three research questions:

RQ1: How is immigration-related misinformation topically structured within Community Notes on X?

RQ2: To what extent do evaluation outcomes (helpful vs. not helpful) differ across immigration-related topics and thematic categories?

RQ3: Do Community Notes addressing different thematic categories of immigration-related misinformation differ in how quickly they receive evaluative attention?

Methods

This study analyzes a dataset of 17,014 publicly available immigration-related CNs (including a note's textual content, creation timestamp, and evaluation status) using a mixed computational and statistical approach that combines semantic topic modeling, evaluation outcome analysis, and temporal measures.

To identify the thematic structure of immigration-related misinformation, we conducted semantic topic modeling using sentence-level text embeddings and density-based clustering. First, notes were analyzed using a custom Python script that used a transformer-based sentence-embedding model to capture semantic similarity beyond keyword overlap. These embeddings were then clustered using HDBSCAN, a density-based clustering algorithm that identifies clusters of varying size while explicitly modeling noise (Campello et al., 2013).

The initial clustering solution identified 51 fine-grained clusters, with approximately 26.7% of notes classified as outliers. To balance structural fidelity with analytical interpretability, these fine-grained clusters were subsequently aggregated into 15 topics based on semantic similarity. Cluster quality was confirmed based on Davies–Bouldin (= 0.764) and Calinski–Harabasz (= 5,945) indexes. For interpretive purposes, we manually reviewed a random sample of notes for each topic to assign descriptive labels. These topics were further grouped into four broader thematic categories based on substantive similarity: (1) Immigration Enforcement and Legal Status; (2) Citizenship, Elections, and Political Authority; (3) Refugees, War, and International Humanitarian Claims; and (4) General Misinformation, Rumors, and Stigmatization. Evaluation outcomes are captured via CNs' publicly available rating statuses: Helpful, Not Helpful, or Needs More Ratings. To assess whether evaluation outcomes differ across topics and thematic categories, we conducted chi-square tests of independence and reported Cramér's V as a measure of effect size.

To examine temporal dynamics, we operationalized evaluation timing as the duration (in hours) between a note's creation timestamp and the moment it first received a rating status other than Needs More Ratings. This measure captures how quickly a note transitions from initial posting to receiving evaluative attention. Differences in duration

were examined using one-way ANOVA at both the topic and thematic-category levels, followed by post-hoc Tukey HSD tests where appropriate.

Results

Semantic topic modeling shows that most immigration-related CNs fall under either the immigration, enforcement and legal status category or the general misinformation category. The largest clusters center on irregular or unauthorized immigration and border enforcement (16.22%) and misinformation claims linking migrants to crime, welfare, or public safety threats (16.13%). These two topics account for over one-third of all immigration-related CNs. Other prominent topics include asylum seekers' legal status and benefits, refugee-related claims in the context of the Israel–Palestine conflict, and U.S. political discourse surrounding immigration leadership and citizenship rights (Table 1). When topics are examined at the level of four broader thematic categories, enforcement- and legality-focused categories and general misinformation-driven themes dominate the immigration-related discourse addressed by CNs, while humanitarian and conflict-related categories form a smaller but distinct category.

Evaluation outcomes differ significantly across both topics and thematic categories. Chi-square tests indicate a systematic association between topic membership and evaluation status ($\chi^2 = 187.86$, $p < 0.001$), though effect sizes are modest (Cramér's $V = 0.087$). At the category level, differences remain statistically significant ($\chi^2 = 91.21$, $p < .001$; Cramér's $V = 0.061$). Notes addressing Refugee, War, and International Humanitarian Claims exhibit the highest proportion of Helpful ratings, but also a relatively higher Not-Helpful ratings. In contrast, notes correcting Misinformation, Rumors, and Stigmatization claims show comparatively high Helpfulness ratios with fewer CNs receiving Not-Helpful ratings. This suggests a more consistent evaluative consensus.

Temporal analyses show that evaluation timing varies primarily at higher levels of thematic abstraction ($F = 4.44$, $p = 0.004$). Specifically, notes related to the Refugees, War, and International Humanitarian Claims thematic category take substantially longer to transition out of the Needs More Ratings stage compared to notes addressing misinformation-driven or enforcement-related content. These results indicate that while topical differences are only weakly associated with evaluation outcomes, thematic category is more strongly associated with the speed at which evaluative attention accumulates.

Future Work

This study extends our understanding of crowdsourced fact-checking by showing that evaluation dynamics within CNs are shaped by thematic context rather than operating uniformly across different topics.

Future research could extend this work by examining non-topical features of notes, such as evidentiary quality, source diversity, and stylistic presentation, to better explain residual variation in evaluation outcomes. Comparative analyses across other highly contested policy domains would further clarify whether the category-level temporal dynamics observed here are specific to immigration-related discourse or reflect more general patterns in crowdsourced fact-checking systems. Finally, cross-national or cross-linguistic studies could help illuminate how political context and information environments condition the functioning of participatory moderation across different sociopolitical settings.

References

- Allen, J., Martel, C., & Rand, D. G. (2022). Birds of a feather don't fact-check each other: Partisanship and the evaluation of news in Twitter's Birdwatch crowdsourced fact-checking program. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22). Association for Computing Machinery. <https://doi.org/10.1145/3491102.3502040>
- Bansak, K., Hainmueller, J., & Hangartner, D. (2016). How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers. *Science*, 354(6309), 217–222. <https://doi.org/10.1126/science.aag2147>
- Campello, R. J. G. B., Moulavi, D., & Sander, J. (2013). Density-based clustering based on hierarchical density estimates. In *Advances in Knowledge Discovery and Data Mining* (pp. 160–172). Springer. https://doi.org/10.1007/978-3-642-37456-2_14
- Chadwick, A., & Stanyer, J. (2022). Deception as a bridging concept in the study of disinformation, misinformation, and misperceptions: Toward a holistic framework. *Communication Theory*, 32(1), 1–24. <https://doi.org/10.1093/ct/qtab019>
- Drolsbach, C., & Pröllochs, N. (2022). Diffusion of community fact-checked misinformation on Twitter. *arXiv*. <https://doi.org/10.48550/arXiv.2205.13673>
- Eberl, J. M., Meltzer, C. E., Heidenreich, T., Herrero, B., Theorin, N., Lind, F., Berganza, R., Schemer, C., & Strömbäck, J. (2018). The European media discourse on immigration and its effects: A literature review. *Annals of the International Communication Association*, 42(3), 207–223. <https://doi.org/10.1080/23808985.2018.1497452>
- Hossain, M. A. (2019). Effects of uses and gratifications on social media use. *PSU Research Review*, 3(1), 16–28. <https://doi.org/10.1108/PRR-07-2018-0023>
- Pröllochs, N. (2022). Community-based fact-checking on Twitter's Birdwatch platform. *Proceedings of the International AAAI Conference on Web and Social Media*, 16(1), 794–805. <https://doi.org/10.1609/icwsm.v16i1.19335>

- Quan-Haase, A., & Young, A. (2010). Uses and gratifications of social media: A comparison of Facebook and instant messaging. *Bulletin of Science, Technology & Society*, 30(5), 350–361.
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>

9E.3 Triadic polarisation in climate information sharing: Insights from Twitter during COP26 and COP27

Ümit Bedrettdin (University of Helsinki), Dayei Oh (University of Strathclyde)

Introduction

The debate over climate change has long been characterized as a polarized contest between supporters of anthropogenic climate change (those advocating for mitigation and adaptation actions) and climate sceptics (those who either deny or downplay anthropogenic climate change or oppose proposed policies to mitigate it). Such binary framings of climate discourse and its polarization, both in mainstream media and digital spaces, have contributed to the perception that climate politics is largely a two-sided conflict. Scholars have widely documented this polarization, highlighting the ways in which misinformation, ideological sorting, and partisan media ecosystems reinforce division between supporters and sceptics (CAAD, 2023; Chinn et al., 2020; Falkenberg et al., 2022). Beyond ideological disagreement, scholars have also examined the affective side of climate discourse polarization, marked by the consistent hostility between climate action supporters and sceptics (Moernaut et al., 2022; Tyagi et al., 2020).

However, such a bimodal conceptualisation of climate discourse polarization is increasingly challenged by the growing chasm within the climate action supporting community itself, namely between the mainstream-moderate institutional actors and more radical climate activists who criticize existing approaches as insufficient (Falkenberg et al., 2022; Oh & Downey, forthcoming). While Falkenberg et al. (2022) continue to frame and evaluate online climate discourse largely through the bimodal lens while acknowledging the growing chasm, Oh and Downey (forthcoming) propose an alternative lens of “triadic polarization,” in which contemporary polarization unfolds across three positions: namely, not just between two extremes, but between two opposing extremes and a contested center. Despite growing recognition of this triadic structure, empirical studies on online climate discourse and online polarization rarely operationalize it in large-scale analysis of digital data (Oh & Dowey, forthcoming).

Objectives/Research Questions

This paper aims to further investigate online climate discourse polarization through the ‘triadic polarization’ lens by focusing on information sharing behaviors during COP26 (2021) and COP27 (2022) on Twitter (now X). While Oh and Downey's (forthcoming) analysis conceptualizes triadic polarization primarily through textual features such as topic polarization, emotion, moralization, and antagonism, we extend this metaphor to the information side of climate discourse polarization.

Specifically, we examine whether the three polarized communities (i.e., mainstream-moderate institutional climate actors, radical climate activists, and radical climate contrarians) consume and circulate similar news and information about climate change and environmental issues, or whether they engage with distinct, segregated information ecosystems in digital media environments. Prior research has explored ICT-induced and information-exposure-based polarization, often treating online information flows as drivers of ideological or affective division (Sunstein, 2009; Leeper, 2014; Xu et al., 2020). In contrast, our approach conceptualizes "information polarization" as a social phenomenon in its own right.

We conceptualize that information sharing practices are not merely antecedents to ideological or affective polarization, but are themselves embedded with polarized patterns of sourcing, thematic focus, and recontextualization into the broader climate politics. Accordingly, we examine three interrelated layers of information polarization in COP26-27 Twitter discourse: (1) polarization at the level of sources, reflected in different citation of media outlets and authoritative institutions; (2) polarization at the level of topics of the cited content, reflected in divergent topical emphases and issue framings; and (3) polarization at the level of recontextualization, reflected in how users linguistically reinterpret and embed cited information in their tweets, in connection with broader climate discourse.

In doing so, this study addresses three guiding questions: RQ1. To what extent do different communities consume and circulate the same or different sources of information about climate change; RQ2. To what extent do they emphasize similar or divergent climate-related issues within the information they share; and RQ3. How do these communities interpret and contextualize shared information in distinct or even contradictory ways. Examining these informational dimensions of polarization provides deeper insight into how climate change discourse and climate politics become polarized in digital media environments. Beyond the climate case, this triadic approach to information polarization offers a conceptual lens that may be applicable to other contested social and political issues characterized by growing distance between institutional or centrist actors and more radical positions.

Methods

The data consists of English-language tweets collected using keywords 'climate crisis,' 'climate change,' 'IPCC,' and 'cop26' and 'cop27' during COP26 (27 October-12 November 2021) and COP27 (5-21 November 2022). The COP26 dataset comprises 8,614,127 tweets posted by 1,836,535 unique users, while the COP27 dataset includes 4,287,835 tweets from 1,284,079 unique users.

To group users into polarized communities, we employed Louvain community detection (Blondel et al., 2008) on retweet networks on Gephi. Retweets were used as a proxy for

ideological alignment and homogeneous information environments, as users are likely to retweet content from accounts they agree with (Cherepnalkoski & Mozetič, 2016; Evkoski et al., 2021). To reduce noise and enhance detection accuracy, we excluded users with fewer than ten retweet interactions. The resulting networks contained 158,937 nodes and 2,699,608 edges (COP26) and 62,385 nodes and 729,785 edges (COP27).

From each network, we retained the eight largest communities by node count for further analysis. The final data totals 11M and 7M tweets for Cop26 and Cop27 summits respectively, from which 4.5M and 1M tweets were sampled for further computational analysis. The data contains original tweet text, computationally assigned communities (mainstream, radical activists, contrarians) and metadata such as embedded URLs. To categorize these communities into the three polarized camps (i.e., mainstream-moderate, radical climate activist, radical contrarians), we conducted a manual validation process. We examined user biographies and randomly sampled tweets from the ten most influential users (based on in-degree centrality) as well as users ranked between 100th and 110th in influence. This comparison enabled us to assess whether communities exhibited consistent positions on climate change and climate action across both highly influential and less prominent users. Based on this manual verification, we confirmed the assignment of Louvain-detected communities into three polarized camps with a high degree of consistency.

For analysing layer (1), the data was augmented by scraping titles from embedded URLs linking to twitter-external sources. This was achieved by creating a custom-built scraper. Media bias labels were attached from an external database maintained by the Media Bias/Fact Check organisation (Media Bias/Fact Check, 2025). To facilitate analysis of layers (2) and (3), the tweet texts and titles of linked sources were dependency parsed, lemmatized, part-of-speech tagged and tokenized with UDPipe (Straka, 2018).

We analysed common and divergent vocabularies across subgroups with TF-IDF and TextRank. Smoothed log-odds ratios of word usage between communities were calculated to identify terms disproportionately favored by each camp. Source-level polarization was assessed by counting cited sources per community and plotting their relative prominence against each other.

Results

Our preliminary findings show that information polarisation spans all three levels to varying degrees.

On the source level, all three camps circulate articles from reputable, legacy media outlets like the Guardian, but to varying degrees. Using the Guardian as a comparison axis, for instance, activists and mainstream camps favour it roughly equally, compared

to the contrarian camp. The three camps also interact with distinct camp-specific outlets: the contrarian camp favours outlets like Daily Mail, Daily Express, and Breitbart, while the activist camp circulates articles from the Guardian and the Independent. These sharing patterns also vary in relation to the geographical location of the summits: during COP26, the contrarian camp frequently references the Glasgow Herald, while during COP27, activists share a disproportionate amount of articles from the Middle East Eye. Overall, three camps tend to not share the same articles, even when links originate from the same media outlets.

The differences in information sharing practices are also visible at the topical level, where we document camp-specific lexical choices and by proxy, distinct discourses linked to each camp. These are reflections of the types of news that tend to get circulated: the contrarian camp predominantly shares articles highlighting the perceived hypocrisy of globalist world leaders, flying by private jets to the summits, while the activists share articles documenting environmental crises including the loss of biodiversity in the face of unending resource extraction.

Our linguistic analysis shows that information polarization extends to the recontextualization level, reflected in syntactic differences within tweets. For instance, when the contrarian camp shares news related to climate change, the syntactic position ‘climate change’ takes is most commonly an oblique (‘What the latest science says about climate change?’), whereas the mainstream and activist subgroups circulate news where climate change predominantly takes the agentive subject position (‘Climate change is threatening/affecting our environment’). The subject position is generally understood to carry discourse centrality, while an oblique functions as an auxiliary or a modifier (Grosz et al. 1995, de Marneffe et al. 2021).

Together, our preliminary results show that polarization in online climate discourse is not confined to ideological or affective dimensions, but is embedded in the sources, topics, and recontextualization practices through which information circulates.

Future Work

While this study focuses on journalistic news and media content, other influential information types (e.g., YouTube videos, cross-platform social media links, and PDF documents including policy reports) remain outside our scope. Future work could extend this analysis beyond journalistic news to include non-journalistic and multimodal information sources to capture a broader spectrum of information sharing practices within polarized digital environments. As we mostly focused on salient linguistic cues and media references, the less common media outlets and more marginal language practices remain potential objects for future studies.

References

- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008.
- CAAD. (2023). Deny, Deceive, Delay (Vol 3): Climate Information Integrity Ahead of COP28. <https://caad.info/analysis/reports/deny-deceive-delay-vol-3-climate-information-integrity-ahead-of-cop28/>
- Cherepnalkoski, D., & Mozetič, I. (2016). Retweet networks of the European Parliament: Evaluation of the community structure. *Applied Network Science*, 1, 1–20.
- Chinn, S., Hart, P. S., & Soroka, S. (2020). Politicization and polarization in climate change news content, 1985-2017. *Science Communication*, 42(1), 112–129.
- Evkoski, B., Mozetič, I., Ljubešić, N., & Kralj Novak, P. (2021). Community evolution in retweet networks. *Plos One*, 16(9), e0256175.
- Falkenberg, M., Galeazzi, A., Torricelli, M., Di Marco, N., Larosa, F., Sas, M., Mekacher, A., Pearce, W., Zollo, F., & Quattrocioni, W. (2022). Growing polarization around climate change on social media. *Nature Climate Change*, 12(12), 1114–1121.
- Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2), 203–225. <https://aclanthology.org/J95-2003/>
- Leeper, T. J. (2014). The informational basis for mass polarization. *Public Opinion Quarterly*, 78(1), 27-46.
- de Marneffe, M.-C., Schuster, S., Manning, C. D., Nivre, J., & Zeman, D. (2021). Universal Dependencies. *Computational Linguistics*, 47(2), 255–308. https://doi.org/10.1162/coli_a_00402
- Media Bias/Fact Check. (Dec, 2025) MBFC's Data API. <https://mediabiasfactcheck.com/mbfcs-data-api/>
- Mihalcea, R., & Tarau, P. (2004). TextRank: Bringing order into text. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing* (pp. 404–411). Association for Computational Linguistics.
- Moernaut, R., Mast, J., Temmerman, M., & Broersma, M. (2022). Hot weather, hot topic. Polarization and sceptical framing in the climate debate on Twitter. *Information, Communication & Society*, 25(8), 1047–1066.
- Oh, D. & Downey, J. (forthcoming). Triadic polarization in climate debates: Climate change and culture wars in English-language tweets in COP26 and 27.

- Straka, M. (2018). UDPipe 2.0 prototype at CoNLL 2018 UD Shared Task. In Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies (pp. 197–207). Association for Computational Linguistics.
<https://doi.org/10.18653/v1/K18-2020>
- Sunstein, C. R. (2001). Republic.com, Princeton, NJ: Princeton University Press.
- Tyagi, A., Uyheng, J., & Carley, K. M. (2020). Affective polarization in online climate change discourse on twitter. 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 443–447.
- Xu, C., Li, J., Abdelzaher, T., Ji, H., Szymanski, B. K., & Dellaverson, J. (2020). The paradox of information access: On modeling social-media-induced polarization. arXiv preprint arXiv:2004.01106.

9E.4 X marks the end: Wartime digital activism, community-led moderation and decentralisation in the “ruins” of social media platforms

Kateryna Kasianenko (Queensland University of Technology), Ariadna Matamoros-Fernández (University College Dublin), Olga Boichak (The University of Sydney), Ashwin Nagappa (Queensland University of Technology)

Introduction

Elon Musk’s takeover of Twitter/X in late 2022 became a watershed moment, prompting many internet researchers and ordinary users to abandon hope in the potential of digitally networked spaces (Tkacz & Gehl, 2025). Some, like O’Sullivan (2025), have positioned these spaces as “ruins”, pointing to the large platforms’ abandonment of content moderation efforts, as well as the unchecked proliferation of spam, scams, and AI slop. However, some of the supporters of Ukraine who had turned to social media, and especially Twitter, in the first days of the full-scale invasion by Russia, chose to maintain their public actions for the sake of Ukraine and Ukrainians despite the increased hurdles presented by the digital platforms. This article examines one notable example of this persistence - the online self-mobilised collective NAFO (The North Atlantic Fella Organization).

NAFO’s members, also known as “fellas”, have been active on Twitter/X since May 2022 by debunking and ridiculing online falsehoods spread by highly visible Russian government accounts and pro-Russian actors, reporting problematic behaviour, as well as fundraising on behalf of Ukraine (Kasianenko & Boichak, 2025). However, since late 2022, these users have been reporting decreased effectiveness of the platform’s response to hateful conduct. NAFO responded to this in two ways. First, they adopted a number of community-led moderation practices (Matias, 2019; Seering et al., 2022) on Twitter/X, which allowed them to maintain and grow their presence on the platform (Authors, 2026). Beyond these community-led moderation practices, NAFO fellas either switched to or created accounts on decentralised platforms with affordances similar to Twitter/X, such as Bluesky and Mastodon. While not all of these platforms are new, they have gained attention from users and scholars following the takeover of Twitter/X by Elon Musk (Roscam Abbing and Gehl, 2024). Some scholars also see these spaces, which are set up, operated, or governed by communities of users, as able to address complex issues in online governance (Authors, 2025), which was not possible on larger, commercial platforms.

In pursuing these multiple tactics to continue their online activism, NAFO serve as an example of surviving in what Tsing (2015) calls “blasted landscapes” or “ruins”. Tsing

(2015) explores this possibility of life in capitalist “ruins” through the case of a matsutake mushroom, which thrives in the landscapes left after industrial destruction of forests, and the diverse communities across the world associated with the fungi through foraging, trade, and other practices. This approach is valuable as a complement to critiques of capitalist production, as it calls our attention to the potential of practices and spaces on the margins of the mainstream to counter its harms and open the pathways for action. In this study, we turn to one group of actors involved in such practices in the context of digitally networked platforms.

Objectives/Research Questions

We investigate how an online community of supporters of Ukraine, NAFO (The North Atlantic Fella Organization), continued to maintain their online activism in the face of the crumbling of Twitter/X - the platform central to the formation and development of the collective (Kasianenko & Boichak, 2024).

We answer three research questions:

- 1) What motivates NAFO members to move away or stay on Twitter/X in the face of the decline in its content moderation?
- 2) What motivates NAFO members to choose a particular decentralised platform?
- 3) How do they achieve community-led moderation on Twitter/X and the decentralised platforms such as Mastodon or Bluesky?

Methods

Our study is conducted in two stages. In the first stage, we drew from 24 semi-structured online video interviews with members of NAFO, a purposive sample of tweets (n=64) which contained mentions of “community notes” extracted from a corpus of over 4 million tweets mentioning the collective between May 2022 and May 2023, and 189 Twitter/X community notes including the keyword “NAFO” collected in October 2024. We utilised a grounded theory approach (Corbin & Strauss, 2014) for the analysis of interviews, community notes and tweets. This allowed us to identify three types of community-led moderation practices relevant to those users who chose to maintain their presence on Twitter/X (see Results section).

The second (ongoing) stage of our study is centered on online video or text-based follow-up interviews conducted from January 2026 with the users we have interviewed during the previous stage. Through these, we primarily explore the motivations NAFO members had for moving away or staying on Twitter/X, especially in the light of the larger-scale departures for users from the platform, such as the move to Bluesky in December 2024 (Boyd, 2024; Radivojevic et al., 2025), as well as their practices for navigating these new spaces and moderating them in a way that supports the collective’s goals. While we are considering complementing the findings from the

interviews with a digital ethnography of relevant servers, feeds, or other spaces within the decentralised platforms, we will not collect digital trace data at scale to stay in line with the user expectations of privacy in these spaces (maloki, 2020; Roscam Abbing and Gehl, 2024).

Results

Community-led moderation, also known as volunteer community moderation (Seering et al., 2022), volunteer moderation (Matias, 2019) or similar terms, refers to moderation performed by unpaid users of online platforms, rather than commercial content moderators employed by online platforms. In a previous iteration of this study (Authors, in press), we identified three types of such moderation practices deployed by NAFO on Twitter/X. These included soft moderation, or reducing the visibility or recontextualising problematic content, hard moderation, or leveraging the platform's reporting mechanisms to remove problematic content or users associated with it, and self-moderation, or practices of caring for the community, maintaining its cohesion and aligning with the collective's values. All of these, but especially self-moderation, demanded a significant amount of emotional labour, highlighting the moral toll that the ongoing care for the community takes on the members engaged in it.

However, as the participants communicated already during the first stage of interviews, the members of the collective were willing to pay that price to stay on Twitter/X rather than move to a different platform with more possibilities for community-led moderation. This was for two reasons. First, they recognised Twitter/X as a space where the online falsehoods by Russia and Russia's supporters were prominent and reached new audiences. One of the objectives of the collective was combatting such propaganda through both soft and hard moderation, so moving away from the problematic platform would defy this goal. Second, for the users who remained on Twitter/X, the move to decentralised platforms had a symbolic meaning of "ceding ground" to problematic actors, like Russia's supporters or the platform company's current owner, Elon Musk. For this reason, while in the follow-up text-based interviews, some participants admitted to creating accounts on Bluesky and Mastodon, they chose to primarily stay in the "ruins" of Twitter/X. Despite this, some fellas recognised the affordances of decentralised platforms as allowing more freedom for users to curate their feeds. Features like curated lists helped users find the collective's members and other supporters of Ukraine more easily. According to our participants, this allowed them to organise and support fundraisers for Ukraine in a more targeted manner.

Future Work

Our initial text-based follow-up interviews indicate that while the members of NAFO value some affordances of the decentralised platforms, they still see them as imperfect tools for activism, noting issues such as the lack of reach. As their presence on these

platforms becomes more established, we will continue investigating their approaches for navigating and moderating these spaces.

References

- Boyd, R. (2024, December 11). From X to Bluesky: Why are people fleeing Elon Musk's 'digital town square'? The Guardian.
<https://www.theguardian.com/media/2024/dec/11/from-x-to-bluesky-why-are-people-abandoning-twitter-digital-town-square>
- Corbin, J., & Strauss, A. (2014). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. SAGE Publications.
- Kasianenko, K., & Boichak, O. (2025). Canonizing online activism: Memetic iconography in the North Atlantic Fella Organization. *Media, War & Conflict*, 18(2), 179–196.
<https://doi.org/10.1177/17506352241279957>
- maloki. (2020, January 27). On Scraping Mastodon. Medium.
<https://maloki.medium.com/on-scraping-mastodon-dffd83139bc1>
- Matias, J. N. (2019). The Civic Labor of Volunteer Moderators Online. *Social Media + Society*, 5(2). <https://doi.org/10.1177/2056305119836778>
- O'Sullivan, J. (2025). The Last Days Of Social Media. <https://www.noemamag.com/the-last-days-of-social-media>
- Radivojevic, K., Adams, D. J., Laszlo, G., Kery, F., & Weninger, T. (2025). User migration in the Twitter diaspora. *EPJ Data Science*, 14(1), 36.
<https://doi.org/10.1140/epjds/s13688-025-00552-y>
- Roscam Abbing, R., & Gehl, R. W. (2024). Shifting your research from X to Mastodon? Here's what you need to know. *Patterns*, 5(1).
<https://doi.org/10.1016/j.patter.2023.100914>
- Seering, J., Kaufman, G., & Chancellor, S. (2022). Metaphors in moderation. *New Media & Society*, 24(3), 621–640. <https://doi.org/10.1177/1461444820964968>
- Tkacz, N., & Gehl, R. W. (2025). After Twitter: Fragmentation, Platform Politics and Protective Sociality. *Social Media + Society*, 11(3).
<https://doi.org/10.1177/20563051251366907>

Session 9F - **PANEL: Conditional Belonging, Regulated Visibility: Women's Platform Participation in China**

Day 2 · Tues 14 July · 15:35–17:05 · Seminar Suite C, ARC

Yue Qin (London College of Communication), Yijia (Yika) Gu (Deakin University), Xintong Jia (University of Essex), Ruoqian Li (University of Sheffield), Nan Ding (London School of Economics)

This panel brings together five empirical studies examining the gendered conditions and costs of women's digital visibility on Chinese social media. Each paper centres on a distinct domain: feminist content creation and discussion, bisexual women's identity negotiation, and everyday social media engagement. The research explores a wide spectrum of platforms, ranging from short-video platforms to podcasts, and from lifestyle social media to consumption apps. Across these different empirical focuses and platforms, the papers collectively forge a shared argument: visibility for women on Chinese social media is not a given but a conditional and precarious achievement, shaped by the intersection of state censorship, platform governance, and gender norms. Women must continuously decode rules, anticipate risks, calibrate their expression, manage emotional costs, and perform cognitive work to be seen, or even to remain present. Yet these demands are rarely recognised as structural burdens; instead, they are naturalised as what women do, or celebrated as individual resilience.

The panel's central concern is conditional and responsibilised platform participation, asking not only what it takes for women to be seen, but what they bear in return. Organised along a continuum from the most explicitly political to the most depoliticised, the panel's structure is itself part of the argument: it reveals that the same mechanism of gendered responsibilities operates across domains conventionally studied in isolation. While grounded in Chinese digital platforms, this panel contributes to advancing conceptualisations of digital visibilities by foregrounding the gendered costs, conditions, responsibilities, and dilemmas that women bear but remain largely invisible in existing scholarship. The panel also includes a moderated discussion that encourages audience participation, inviting attendees to draw connections across the papers and to reflect on how the dynamics identified here may resonate with other platform contexts.

Visibility on digital platforms has become increasingly significant for women, shaping opportunities for their self-expression, economic gain, the circulation of everyday gender inequalities, and cultural participation beyond the gatekeeping of traditional

media. Relevant concepts have been developed and applied to facilitate academic and social discussion about this trend, such as visibility labour (Abidin, 2016), gendered (in)visibility (Wilhelm, 2020), and the entanglement of popular feminism with economics of visibility (Banet-Weiser, 2018).

Various factors, including commercial logics, social norms, and technological affordances, influence how women achieve and negotiate their visibility on platforms (Gillespie, 2018; Thompson, 2005). These factors do not operate evenly; they replicate offline gender power structures, ultimately leading to the selective amplification of women's expressions that align most closely with platform capitalism (Duffy & Hund, 2019), while feminist activities, deviant women and non-conforming identities are subjected to regulation and marginalisation (Liao, 2024; Sobieraj, 2018).

Given this backdrop, women's negotiation and contestation of digital visibility are highly culturally contextualised. China provides a particularly compelling case, as its platform ecology further complicates the conditions for women's visibility, belonging, and even just participation. Under what Yang (2021) terms 'state-sponsored platformisation', party-state influence operates not through direct command-and-control but through an entanglement of regulatory pressure and commercial incentive (De Kloet et al., 2019). Within this structure, platforms seeking to balance profit with compliance make governance boundaries nominally visible, while treating compliance as a task users must learn and adapt to on their own (Shen, 2016; Ye et al., 2025). The conditions of visibility are thereby responsibilised and internalised as demands of self-management: women must continuously decode rules, anticipate risks, and calibrate their expression and audience engagement to maintain legitimacy and, first and foremost, platform survival.

Existing research has documented Chinese female users' adaptive work within these constraints, from feminist activists' censorship navigation (Chang, 2025; Lin, 2023), to building female-oriented safe spaces (Gu & Heemsbergen, 2023; Wang & Cong, 2024), to sexual minority women's creative identity expression under hostile moderation regimes (Ai et al., 2023; Wei et al., 2025). Yet as Duffy and Meisner (2023) caution, it is too easy to read circumvention as creative achievement without attending to the precarious and ongoing costs that sustain it. When individual efforts for visibility are celebrated in isolation, a familiar displacement occurs: structural constraints are reframed as personal challenges (Gill, 2016), while the unequal conditions that force women to employ these strategies become naturalised, and the costs behind their visibility are made invisible and privatised (Banet-Weiser, 2018).

This panel thus turns attention to the other side: asking not only what women do to be 'seen', but what they bear in return. We pose two interlinked questions: (1) what conditions must Chinese women navigate to be present and seen on platforms, and what work does this demand? and (2) how are these conditions converted into

individual responsibilities that women are expected to bear, rendering the costs of participation invisible and privatised?

Five empirical studies included in this panel span feminist content creation, cultural reception, identity negotiation and everyday social media engagement, and range from short-video platforms to lifestyle social media and to audio podcasting. Organised from the most explicitly political to the most thoroughly depoliticised, this progression reveals a shared continuum of gendered platform participation: regardless of the content's nature, female users are universally compelled to perform continuous, invisible work, emotional, cognitive, or strategic, to bridge the gap between platform governance and their desire to be seen. While grounded in the Chinese context, this panel contributes to advancing conceptualisations of digital visibilities by foregrounding the gendered costs, conditions, responsibilities, and dilemmas that women bear but remain largely invisible in existing scholarship.

These five papers are presented as follows:

Paper 1: A Platform-Adaptive Feminism: Affordances and the Rearticulation of Feminist Discourse on Douyin (Yue Qin)

Aligning with the panel's most explicitly political dimension, or to say the domain most directly subject to platform moderation (Han, 2018; Liao, 2024), this paper examines how creators maintain 'feminism' visible on Douyin (TikTok's Chinese counterpart), currently China's most influential social media platform, and poses the core question: how do Douyin's platform affordances influence and reshape these feminist articulations? Using multimodal critical discourse analysis informed by feminist critical discourse analysis as a theoretical lens, the study analyses 200 public Douyin videos posted by 23 feminist content creators. The analysis conceptualises 'platform-adaptive feminism' to explain practices whereby creators treat platform adaptation as a means, visibility as an aim, and the preservation of feminist critique as a baseline — while being reshaped by the very rules they navigate. Three overarching themes are identified: (1) Mass-oriented feminist pedagogy, in which creators need to translate specialist knowledge into vernacular micro-explainers, framing feminism as public, easy-understanding knowledge work; (2) The entrepreneurial self, whereby creators have to couple feminist advocacy with promotional labour, integrating sponsorships and cultivating a marketable personal brand to secure visibility and monetisation; And (3) post-production for imagined censorship, in which creators relocate risk to the video editing layer, deploying caption homophones, emoji, and blurring to pre-empt sanctions under opaque algorithmic monitoring. Taken together, Douyin's affordance regime organises a strategic re-articulation of feminist expression, revealing the multiple layers of work that women must perform to make feminist discourse legible and visible on the platform.

Paper 2: Building Relational Agency through Chinese Feminist Podcasting (Yijia Gu)

Shifting from video to audio platforms, this paper turns attention from the visible performance of feminist content to the often-invisible relational work that sustains feminist media production. Feminist podcasts have emerged since 2020 as significant sites of resistance within China's increasingly restrictive media environment, yet audio platforms offer neither guaranteed belonging nor unconditional space for feminist discourse. Drawing on semi-structured interviews with 13 Chinese feminist podcasters at different stages of development, Gu adopts relational approaches to agency (Mackenzie & Stoljar, 2000) to examine how feminist media practice is developed through and sustained by relationships of care, connection, and mutual support. The findings reveal that podcasting arises from and through relationships: friendships, intergenerational ties, and shared experiences of gendered marginalisation provide both the motivation and infrastructure for production. Sustaining this practice, however, demands substantial yet largely invisible work: managing financial constraints without reliable revenue, navigating opaque and unevenly enforced platform moderation, and carefully calibrating between audience intimacy and protective distance. Notably, access to platform negotiation is itself stratified, as prominent podcasters can contest moderation decisions through accumulated relational capital, while newer producers face arbitrary enforcement with limited resources. This paper therefore argues that the responsibility for sustaining feminist media falls not on platforms or institutions but on individual women and their personal relationships, a dynamic that speaks directly to the panel's concern with how platform governance is converted into gendered, invisible labour.

Paper 3: Trending Feminism in China: Audience Responses to the Feminist Film Her Story (2024) on RedNote (Xintong Jia)

Pivoting from production to reception, this paper examines how audiences engage in interpretive and discursive work when encountering feminist cultural texts on social media. Along with the growing visibility of feminist discussion in Chinese popular media culture, the study examines audience reception of the Chinese feminist film *Her Story* (2024) on RedNote (Xiaohongshu). As one of the pioneering feminist films produced and released in China, *Her Story* has sparked extensive discussion of feminist issues on the platform. The study collects posts and comments via hashtag searches and analyses them using critical discourse analysis. The initial findings indicate a narrative shift away from (post) feminist self-empowerment and stress on women's agency toward greater visibility, acceptance, and embrace of vulnerability and ordinariness on RedNote, alongside a rejection of confidence-centred narratives and positive, affirmative affect. Building on Western scholarship on popular feminism (Banet-Weiser, 2018) and postfeminism (Gill, 2007), this paper interrogates the (in)commensurability of applying these frameworks to Chinese social media contexts. Within the panel's framework, it

demonstrates that audience reception of platformised media is not passive consumption, but active meaning-making. It raises the question of who bears the interpretive work of making feminism legible under platform-specific conditions of visibility.

Paper 4: Queer but Not 'Queer Enough': Chinese Bisexual Women's Online Practices and the Politics of Belonging on RedNote (Ruoqian Li and Tong Meng)

Moving from feminist discourse construction and discussion to identity politics, this paper investigates how Chinese bisexual women engage with RedNote, and how the platform environment shapes their navigation of sexual identity, visibility, and belonging. Drawing on semi-structured interviews with 10 Chinese bisexual women and a qualitative analysis of 30 posts and 150 associated comments, the study is guided by the concepts of conditional belonging (Yodovich, 2021) and emotional labour (Hochschild, 2002). The preliminary findings suggest that bisexual identities are largely invisible on RedNote and are often stigmatised. Participants described a form of double marginalisation: critiqued as not 'normal' within heterosexual contexts, while being treated as 'not queer enough' within queer communities. Importantly, this marginalisation was further reinforced by RedNote's moderation regime and algorithmic logics, such as search and categorisation. In response, many participants adopted strategies of self-silencing or identity substitution, presenting themselves as 'lala' (拉拉, lesbian) to access queer communities, while others spent extra time and emotional effort posting and commenting to contest bisexual stigma, a process they repeatedly described as exhausting and frustrating. The findings complicate celebratory accounts of digital queer inclusivity and reveal how platform structures convert the structural problem of bisexual erasure into a personal burden, as it is individual women who must manage, perform, and defend their identities to belong.

Paper 5: From 'Scrolling' to Social Reproduction: Platformised Cognitive Labour and Middle-Class Women's Leisure in Urban China (Ding Nan and Tianling Yang)

At the far end of the panel's spectrum, this paper examines forms of gendered work so thoroughly embedded in everyday life as to be rarely recognised as such. Introducing an intersectional perspective that foregrounds the interplay of class and gender, the study draws on the concept of cognitive labour (Damingier, 2019) to examine how socio-cultural expectations are translated into everyday practices of information search, evaluation, and decision-making. Focusing on urban middle-class women in the domain of leisure, the research shows that women continue to shoulder a disproportionate share of consumption-related cognitive labour within the household. This structural positioning renders them more likely to engage intensively with platformised information and to participate as prosumers. Crucially, this dynamic is not merely a matter of individual preference or digital habit. Instead, it is rooted in intertwined socio-cultural formations: middle-class identity and aspirational

imaginaries of the ‘good life’. For instance, the expectation that a responsible mother should efficiently curate enriching weekend activities for her children drives women to resort to platformised knowledge bases. At the same time, these platforms reciprocally shape and intensify such aspirations, demanding ever greater cognitive labour in the form of continuous information gathering, comparison, and optimisation. By demonstrating how contemporary middle-class femininity is translated into cognitive labour and how this labour, in turn, consolidates gendered divisions within the household, the study highlights the subtle ways in which platform cultures both mobilise and normalise women’s responsibility work.

Session 10A - **PANEL: Cloud-C: Studying climate communication on social media through a "many researchers, one dataset" approach**

Day 3 · Wed 15 July · 10:40–12:10 · Seminar Suite A, ARC

Dayei Oh (University of Strathclyde), Axel Bruns (Digital Media Research Centre at Queensland University of Technology), Barbara Pfetsch (Freie Universität Berlin), Risto Kunelius, Anton Berg (University of Helsinki)

This panel offers highlights of an international research network Cloud-C launched in 2024 to study the intersections of climate communication and social media through a ‘many researchers, one shared social media dataset’ model. Cloud-C brings together researchers and institutions across the globe and across disciplines.

The panel presents both empirical findings and methodological lessons from analysing 20 million tweets from two UN climate summits: COP26 (2021) and COP27 (2022). We also reflect on practical insights and challenges in building transnational data-sharing and collaborative research practices driven by shared dataset instead of more common models of shared research questions.

This shared-data approach speaks to two core contributions. First, we test and show how applying different methods to the same shared dataset allows scholars to leverage different affordances of social media data. Sharing the same data also builds more robust opportunities for assessing reliability, validity and limits of social media data.

Second, our data set represents a pivotal moment in global climate politics. The 2021-22 summits can be seen as the end of an era: since then, we have moved from science-based international climate governance to increasing national competition over remaining resources, from national voluntary pledges to ditching support for transnational accountability measures, and from the sustainability-centred narratives to aggressive claims that AI will “solve” the climate crisis. Examining social media climate debate at this turning point offers important lessons.

Our panel begins with an introduction to Cloud-C, followed by four empirical papers examining online climate discourse with diverse conceptual and methodological interests including triadic relational dynamics in climate polarization, user interactions/practice mapping as an innovative method to study digital trace data, discursive use of geography in transnational climate publics, and the role technology talk in COP-interactions.

Together, we aim to 1) provide conceptual, methodological, and empirical contribution to climate communication research; and 2) highlight data-sharing-centred transnational cooperative practices.

Anthropogenic climate change has increasingly become one of the focal points of global politics in diverse ways including growing climate skepticism, ideological and affective polarization, unequal implications between Global North and South, geopolitics, discourses of retribution and accountability, and radicalisation of climate activism, just to name a few. How climate communication unfolds in digital environments is particularly significant as an increasing amount of political communication, information sharing, and collective mobilisation is mediated by social media platforms. Global summits like COP offer an exceptional and fruitful case, as these focusing events provide a specific time window where climate change becomes an especially salient and relevant topic across political spectrums and geographies. This panel provides timely and highly relevant insights and evidence about contemporary climate communications by different actors, across different regions, mediated by platform affordances (Twitter/X in this case).

Beyond empirical contributions, this panel advances an innovative model of scientific collaboration based on the ‘many researchers, one shared dataset’ approach in contrast to more conventional research networks practices that are organised around shared research questions and objectives. In doing so, this shared-data centred network offers diverse innovations.

First, it enhances methodological reproducibility and controlled comparability. Shared data infrastructure also reduces duplications of data collection efforts, partially mitigating the global inequalities in the access to digital trace data, especially in the post-API digital research landscapes.

Second, this approach enables meta methodological reflections by systematically examining how different conceptual constructs and methods reveal diverse insights about global climate communications from the same datasets.

Third, this approach extends the range of transnational and multidisciplinary collaboration and data sharing practices, as the network is bound by the shared data, instead of shared epistemic or disciplinary, or even regional interests. Our network comprises scholars from five different continents (Africa, Asia, Australia, Europe, and North America) and different disciplines (from media and communication studies to computational social science to anthropology), generating a polysemic interpretation of global climate communications.

Lastly, our collaborative efforts will be developed into a journal special issue and future grant co-applications, highlighting the sustained possibilities of transnational collaborative networks.

Here is the list of abstracts of the papers presented in the panel:

1. Triadic relational dynamics in climate debate polarization: Culture wars and discursive convergence in English-language tweets in COP26 and 27 (Dayei Oh & John Downey)

This paper examines polarization in climate debates on Twitter during COP26 (2021) and COP27 (2022) using large-scale digital data. We articulate triadic relational dynamics of polarization as an empirically observed analytical construct describing growing disagreements and hostility between a mainstream-moderate center and two opposing poles, which are not fully captured by a binary believer-skeptic axis. Using retweet-based community detection, we identify three macro-communities: (1) mainstream-moderate center, (2) radical climate activists, and (3) radical climate contrarians. Text mining shows that the interactions across these three macro-communities increasingly organize climate discourse as a three-party conflict, with both radical poles orienting discursively against the mainstream-moderate center. We also observe a growing entanglement of culture wars, connecting climate issues to broader ideological conflicts with heightened emotional, moral, and antagonistic language. While contrarians remain key drivers of culture-war rhetoric, climate activists also increasingly adopt similar affective and moralized strategies, indicating discursive convergence across radical groups.

2. COP26/27 Engagement on Twitter: An Analysis Using the Practice Mapping Approach (Axel Bruns, Ehsan Dehghan, Michelle Riedlinger, Carly Lubicz-Zaorski, Laura Vodden, Daniel Angus, and Samantha Vilkins)

Major international gatherings such as the COP26 and COP27 climate summits usually attracted considerable global engagement on social media, especially on Twitter (before the considerable downturn in active use following its acquisition by Elon Musk). Conventional analysis of the hashtag or keyword datasets gathered for the public discussion of such events often produces only relatively limited and predictable insights, however: much of the public debate tends to be centered around a handful of major, widely @mentioned and retweeted accounts (including official event, host, and UN accounts), uses the generic event hashtags, and represents well-known groups of activists and lobbyists. None of this is surprising.

In this paper, we advance our analysis beyond such predictable outcomes by applying a novel analytical approach called practice mapping, which instead of building on an interaction network analysis approach investigates similarities and disjunctures between the distinct posting practices of participants in the datasets for the two events. This still takes into account patterns in whom accounts @mention and retweet, and how they are addressed in return, but also adds further activity patterns that relate to their distinct uses of hashtags, URLs, and language choices. Drawing on this

expanded range of activity patterns, the practice mapping approach detects a broader range of distinctive participation practices within the COP26 and COP27 debates on Twitter, enabling us to identify and interpret a more diverse set of participant groups within the datasets.

3. *A Global Problem? The Discursive Geography of Climate Change on Twitter during COP26 and COP27 (Daniela Stoltenberg, Ellen Linnert, and Barbara Pfetsch)*

Climate change is a planetary issue, yet its causes and impacts are unevenly distributed across the globe. This imbalance is also reflected in public communication. In this study, we explore where social media users spatially locate climate change and what issues they discuss in relation to different regions. Building on theories of transnational public spheres and issue spaces, we map the Twitter discourse surrounding climate change. We develop computational methods to analyze multilingual tweets from two COP summits, because we suppose that social media users leverage the global visibility of the summits to promote local climate agendas. Our results indicate that, similar to mass media coverage, communication is structured by regional interests and emphasizes climate mitigation over adaptation and the concerns of vulnerable regions. Thus, online communication about climate change is skewed and biased toward the issues and locations of major greenhouse gas emitters, such as advanced and emerging economies.

4. *Tech Talk at a Climate Turning Point. Analyzing the solution-risk language in COP21 and COP22 Twitter data (Risto Kunelius, Matt Tegelberg, Anton Berg, Juho Pääkkönen and Matti Pohjonen)*

In climate politics, belief and doubts about technology play an important role. Given the changing political landscape of recent years, its importance is growing. For instance, the discussion of about geoengineering climate is gaining strength (both as way of avoiding catastrophic tipping points and playing for more time to business as usual). At the same time, the benevolent sustainability image of Big Tech has given way to aggressive belief that rapid growth of AI is – despite its climate burden – is necessary risk that needs to be taken. In these narrowing horizons, understanding the role of technology discourse is crucial.

In this paper, we conduct an analysis on how key actors in the 2021-2022 COP-context defined the role of technology as part of the global effort to address the climate crisis. The research design comprises a novel, iterative, LLM-enhanced (ChatGPT-4o) human-validation analytical process. We aimed to capture the solution-vs-risk dynamics of COP-Twitter interaction while addressing two distinct challenges. First, to qualitatively compare the hopes and worries across different technology sectors—ranging from AI and communication technology to human mobility and geoengineering—and to map the role of "climate tech talk" in 2021 and 2022 (a period when global discursive

horizons reached a turning point), we required a method capable of handling a large dataset of over 20 million tweets. Second, due to concerns regarding validity, reliability, and known risks such as model biases and hallucinations, we determined that the analytical task could not be delegated to AI alone. Instead, an approach was needed that combines LLM annotation with human validation. This is also crucial because, as mixed-methods researchers, we must develop new future protocols to ensure AI-assisted research is conducted in the most reliable, reproducible, and transparent manner possible. Our analysis looks at how key actors in the 2021-2022 COP-context defined the role of technology as part of the global effort to address the climate crisis. We compare the hopes and concerns in different sectors of technology (from AI and communication technology to human mobility and geoengineering), and map out what role the “climate tech talk” played in 2021 and 2022 during a moment when the global discursive horizons were coming to a turning point.

Session 10B - **PANEL: Algorithmic Childhood, Intimacies, and Literacies: Networked & synthetic selves and relationships**

Day 3 · Wed 15 July · 10:40–12:10 · Seminar Suite B, ARC

Jessica Ringrose, Chiara Fehr (UCL Institute of Education), Ayşe Aslı Bozdağ (Bahçeşehir University), Devina Sarwatay (City St George's, University of London), Sandra El Gemayel (Digital Futures for Children Centre, LSE)

This panel brings together leading experts to discuss new and innovative research on how AI and algorithms are reshaping children and young people's identities, relationships and education. The participants represent established academics, early career scholars and a doctoral student, bridging differing levels of career, aiming to spark dialogue across generations. The panel intends to generate discussion and debate around critical issues such as: What are the experiences of young people in navigating social media and AI technologies like LLMs, chatbots and EdTech? How should social media and AI be addressed in school contexts? What policies and practices could support young people amidst rapid technological shifts and change? The overall objectives of the session are to offer cutting edge insights into this emergent area of research on how young people are navigating algorithmic life and to offer recommendations around how educational policy and practice can keep pace with technology and better support youth.

To this end, the panel titled, "Algorithmic Childhood, Intimacies and Literacies: Networked & synthetic selves and relationships," will have the following presentations:

1. *Devina Sarwatay, The 'Synthetic Self' and Algorithmic Childhood: Navigating AI Literacy and Digital Intimacies*
2. *Chiara Fehr, Algorithmicized Girlhood: Exploring how Algorithmic social media Influences the Sexual Self-Perception and Wellbeing of Teenage Girls*
3. *Jessica Ringrose, Ayşe Aslı Bozdağ and Marc Tibber, Understanding young people's everyday use of AI: From AI Dependencies and Vulnerabilities to AI Literacies*
4. *Kim R. Sylwander, Sonia Livingstone and Sandra El Gemayel, Algorithmic childhood in the classroom: Personalisation, synthetic selves and children's right to education*

Chair, Jessica Ringrose

The papers in the session critically explore children and young people's experiences of social media platforms, LLMs and chatbots and EdTech. This topic is extremely timely what with the rise in AI chatbots, debates over social media bans and Smartphone banning in school settings, and controversies over whether EdTech is harmful or helpful for youth. The papers in the session offer a range of empirical research and theoretically framed inquiry each offering a unique perspective on issues of algorithmized childhood and adolescence.

Sarwatay's paper contributes important key concepts into this field of debate including newly formed synthetic selves which create new forms of AI intimacy, and therefore the need for enhanced AI literacy. Fehr's paper delves into how the algorithmic architecture of social media platforms are reconfiguring platformed girlhood, paying special attention how TikTok promotes consumption based heterosexulised ideals of femininity. Ringrose, Bodzag and Tibber's paper explores young people's everyday use of AI. They argue use of AI like ChatGPT at school normalises adoption, promoting attachment, dependency and cognitive offloading. They also, however, share participant's critical insights into how AI is impacting their futures, relationships and learning. Finally, Sylwander, Livingstone and El Gemayel examine personalised educational algorithms in EdTech, critically analysing distinctions drawn between "good" educational algorithms and "harmful" social media algorithms, finding both are extractive. They call for a rights-based governance framework that reframes educational benefit through children's rights.

The 'Synthetic Self' and Algorithmic Childhood: Navigating AI Literacy and Digital Intimacies - Devina Sarwatay

As generative AI systems increasingly permeate young people's everyday lives, the boundaries between self-expression, algorithmic influence, and mediated identity are becoming blurred. I interrogate how young people negotiate AI-mediated intimacies and construct synthetic selves within networked environments, drawing on insights from my prior research on young people's digital cultures and my conceptualisation of the 'Synthetic Self.' While AI literacy initiatives often foreground technical competencies, such as understanding algorithms or coding skills, they rarely address the affective and relational dimensions of algorithmic systems: how these systems shape belonging, vulnerability, and identity formation in ways that are deeply embodied and emotional.

Through a critical synthesis of existing studies and emerging literature on algorithmic childhood, I explore three interlinked dynamics that illuminate the complexity of youth engagement with AI:

'Synthetic Self'-making: Young people increasingly curate identities through generative AI tools, social media filters, and algorithmically driven recommendation systems. This

process involves oscillating between authenticity and algorithmic optimisation, where self-presentation becomes a negotiation between personal expression and platform logics. The ‘Synthetic Self’ is not merely a digital avatar, but a composite identity shaped by machine learning predictions, cultural aesthetics, and peer validation.

AI Intimacies: Beyond technical interaction, young people experience emotional labour when engaging with recommendation engines, chatbots, and digital archives. These encounters often blur the line between human and machine intimacy, raising questions about trust, care, and mental health. For instance, reliance on AI companions or therapeutic bots introduces new forms of mediated vulnerability, while algorithmic curation of memories and histories can evoke affective responses that challenge traditional notions of privacy and autonomy.

Critical AI Literacy: Current frameworks for AI literacy tend to prioritise functional skills, such as understanding data privacy or algorithmic bias, while neglecting ethical, affective, and socio-political dimensions. I argue for a reconceptualisation of AI literacy that foregrounds empathy, relationality, and critical awareness of power structures embedded in AI systems. This approach moves beyond instrumentalist views of technology to consider how young people can cultivate agency and resilience in environments saturated with algorithmic influence.

By situating these dynamics within broader debates on digital empathy, platform governance, and youth agency, I call for a paradigm shift in how scholars, educators, and policymakers conceptualise AI literacy. Rather than treating literacy as a set of discrete technical skills, we must recognise it as a holistic practice that encompasses emotional intelligence, ethical reasoning, and socio-cultural critique. This intervention contributes to ongoing conversations about the futures of social media scholarship, highlighting the need for interdisciplinary approaches that bridge technology studies, childhood studies, and critical media theory.

Algorithmicized Girlhood: Exploring how Algorithmic social media Influences the Sexual Self-Perception and Wellbeing of Teenage Girls - Chiara Fehr

In past years social media has shifted to increasingly incorporate algorithms, fostering what Bhandari and Bimo (2022) have termed the ‘Algorithmized Self.’ While many users of algorithmic apps such as TikTok, YouTube, or features such as Instagram Reels enjoy the hyper individualized, curated content (Steen et al., 2021), the extent to which the algorithm holds influence over young women and girls’ contemplation of femininity and self-articulation must be problematized. As Ringrose et al. (2024) raises; young people’s postdigital algorithmic subjectivity must be considered as an affective phenomenon, as offline ‘material embodiment’ become deeply entangled with online experiences. Deeply rooted in neoliberal capitalism, TikTok’s algorithmic architecture entangles teenage girls’ techno-sexual assemblage surveillance based, consumerist

social order. This research consisted of a series of creative workshops with 16-18 year olds, including collaging (Coleman, 2009), a utopic mapping exercise (Ringrose & Renold, 2019) and a 'photo voice' inspired 'TikTok diaries' task using screen recordings.

What emerged was a complex image of girlhood where algorithmically facilitated self-articulation was deeply entangled with gendered capitalist logics of consumption through complex affective relationships to social media use. Central to their lived experiences were consumption-based rituals of embodiment, backlash to post-feminist doctrine of self-sexualization, and anxiety around algorithmically platformed misogyny and its influence on opinions and expectations held by their male peers.

Furthermore, TikTok's algorithmic environment specifically platformed heteronormative hyper-sexualisations of female bodies for the male gaze, thereby facilitating an environment in which young women are disempowered in their sexuality yet feel pressured to conform to pornographic sexual standard. These pressures were ultimately post digitally re-enforced through male peers influenced by the TikTok algorithmic platforming of toxic masculinity cultures. Particularly concerning was the pervasive circulation of content and products promoting toxic behaviour around diets and exercise. While TikTok has taken actions like banning #SkinnyTok, weight loss products remain at the centre of the app's revenue stream. However, is important to note that TikTok created a space similar to previous platforms such as Tumblr for teenage girls to explore desire. This specifically relates to diverging from heteronormative scripts thorough the sharing of 'thirst edits', trends like the hear me out cake, and sharing of romance fiction.

This project provides insight into the dynamics amongst teenage groups in a contemporary context where young people are increasingly looking to social media to inform their gendered and sexual learning as their digital needs are not met by standard RSHE (Ringrose et al., 2024). This research seeks to contribute to ongoing discourses around pedagogy to provide young people with tools to navigate themes of selfhood, consumerism, sex and sexuality on social media as part of official RSHE. In allowing young people to raise their concerns and priorities, this project furthermore aims to contribute to education and policy decisions that values teenager's needs, instead of projecting moral panics onto education and regulations.

Understanding young people's everyday use of AI: From AI Dependencies and Vulnerabilities to AI Literacies - Jessica Ringrose, Ayşe Aslı Bozdağ and Marc Tibber

Artificial intelligence (AI) is an increasingly present part of children's online lives, shaping how they search, speak, and interpret information. Generative AI (GenAI) enters the school ecology as chatbots, translators and creative tools (Sengar, 2025). Popular commentary often frames AI's educational impact through narrow concerns over cognitive offloading (Gerlich, 2025). Educational environments are highly

contradictory, with phone free policies promoting fears and concerns over social media and AI on the one hand yet implementing and promoting LLMs and EdTech to support schoolwork on the other. What rarely enters these debates is the complexity of how AI settles into young people's daily negotiations, becoming part of how they manage interactions from schoolwork to peer and family relations (Chirayath et al., 2025). Very few qualitative studies have explored youth experiences, voices, feelings and concerns around everyday use of AI in and around school contexts.

To address this gap, this paper draws on a mixed methods study on Smartphone use and restrictions in school, focusing on one strand which charted young people's everyday use of AI. We draw on focus group data with 40 participants aged 15 to 18 from three schools, exploring teens' experiences and understandings of GenAI, including perceptions of associated risks, politics, and implications for their education and futures. We argue that school-based uptake of chatbots like Chat GPT normalises AI into everyday routines and we show how contradictory and uneven school responses and tensions around AI, detract from meaningful AI literacies. We found that the perceived anonymity, reliability, and conversational ease of AI chatbots fosters attachment and weakened judgement. Whereas helpfulness and enjoyment were common, young people worried about over-dependence and reduced autonomy. Participants also expressed anxieties and future-oriented fears around how AI would fundamentally shift society, work trajectories and reduce the value of formal education. Finally, we document moments of critical engagement including self-recognition, self-regulation and explicit sociopolitical critique.

Overall, our analyses show how AI weaves into adolescence across infrastructural and developmental terrains, with AI intimacy distributed across an everyday digital ecosystem rather than located in a single app; binding cognition, affect, and behaviour with privacy and data vulnerability. Our conclusions offer insights into how educational policy and school practice can develop more supportive AI literacies capable of fostering more reflexivity and transparency around youth engagement with AI.

Algorithmic childhood in the classroom: Personalisation, synthetic selves and children's right to education - Kim R. Sylwander, Sonia Livingstone and Sandra El Gemayel

Personalised algorithms are reshaping childhood across domains of education, communication, entertainment and play. In schools, adaptive learning systems, AI tutors and analytics dashboards promise tailored instruction, inclusion and improved attainment. In social media environments, similar algorithmic logics are widely criticised for datafication, behavioural manipulation and the shaping of networked identities for commercial gain. This paper examines whether personalised educational algorithms cultivate child-rights-respecting literacies and relationships, or whether they

reproduce the same extractive logics that largely structure children's online social lives.

The paper is situated within a child rights framework grounded in the UN Convention on the Rights of the Child and General comment No. 25, focusing on children's rights to education, privacy and the primacy of their best interests. First, the paper draws on a scoping review of reviews, including 88 systematic reviews and meta-analyses published between 2023 and 2025 examining causal evidence of EdTech's impact on children's measurable learning outcomes. Second, qualitative workshops with over 450 primary and secondary school children and interviews with teachers across the United Kingdom, exploring everyday experiences of education technologies deployed by schools, exploring everything from school platforms, homework apps, gamified systems, AI tools and school procurement practices.

The scoping review shows that many personalised and adaptive systems are associated with statistically significant positive effects on performance-based outcomes, particularly in mathematics, literacy and STEM. However, the evidence base is narrowly focused on attainment metrics and rarely considers how algorithmic systems shape children's agency, identities, relationships or education and data privacy rights. In other words, what counts as "benefit" remains largely instrumental. At the same time, most personalised systems depend on continuous data capture, profiling and automated recommendations that construct children as measurable, optimisable learners.

Our qualitative findings illuminate that teachers value granular performance data and automated task setting yet describe a fragmented and commercially driven procurement landscape. Children articulate more ambivalence where some appreciate tailored feedback and others report rigid homework timers, opaque scoring mechanisms and AI-generated tasks that prioritise speed, compliance and point accumulation over understanding and child-teacher interaction.

Gamified leaderboards and competitive metrics shape peer relations and self-perception, producing new forms of comparison and visibility among students. These practices blur boundaries between educational and social algorithmic environments, contributing to networked and synthetic selves that are simultaneously learners, data subjects and quantified performers.

We argue that the distinction between "good" educational algorithms and "harmful" social media algorithms is not technical but normative and institutional. Personalised systems become child-rights-respecting when pedagogical purpose clearly overrides commercial imperatives, when data processing is necessary and proportionate, when teachers retain meaningful interpretive authority, and when systems support inclusive literacies that foster critical understanding of how algorithms work. Without these

conditions, classrooms risk importing the same logics of optimisation, surveillance and behavioural nudging that characterise commercial platforms. We therefore call for a rights-based governance framework, that reframes educational benefit through children's rights, particularly their right to education, and offers a principled way to evaluate whether algorithmic systems support children's flourishing as learners and as developing subjects within networked societies.

Session 10C - **PANEL: Algorithmic Sociality: Platforms, Connection, and the Reconfiguration of Digital Relations**

Day 3 · Wed 15 July · 10:40–12:10 · Seminar Suite C, ARC

Ellenrose Firth (Sapienza University of Rome), Andreas Schellewald (Goldsmiths University London), Hui Lin (King's College London), Yingwen Wang, Wenxin Guo (London College of Communication)

This panel advances the concept of algorithmic sociality to examine how social media users build, manage, and bound their social relationships on recommendation-driven platforms. Bringing together five qualitative studies across TikTok, Instagram, and Douyin, the panel asks a shared question: what does it mean to form and sustain social ties when algorithms mediate who becomes visible to whom, and when platforms operationally define what counts as connection?

We argue that algorithmic sociality produces distinctive tensions that existing frameworks of platformed connection do not fully address. These tensions structure the panel's three objectives. First, to trace how relations with algorithmic systems become conditions for relations with others, from users who develop affective bonds with "their" algorithm to those who strategically engage algorithmic systems to manage social ties. Second, to examine the boundary work through which users decide where platform relationships belong: some distribute ties across multiple platforms, others deliberately contain relationships within a single platform. Third, to foreground inequality by asking who possesses the capacity for such strategic navigation and whose participation is structurally constrained before it begins.

By centring algorithmic sociality as an analytic lens, the panel addresses a gap between platform studies, which emphasise infrastructural and economic logics, and research on digital social life, which often treats algorithms as background conditions rather than active mediators of connection. The five papers provide empirical grounding for how algorithmic sociality operates across different populations, platforms, and contexts, revealing that in recommendation-driven environments, the distinction between human-algorithm relations and relations with other people is not a binary but a continuum.

The panel brings together researchers across four institutions with expertise spanning Western and Chinese social media platforms. All presentations draw on qualitative methods including interviews and ethnography, contributing to methodological conversations about studying human-algorithm relations empirically.

Recommendation-driven platforms such as TikTok and Douyin have become central reference points for rethinking what sociality looks like under contemporary platform conditions. A recurring argument in recent scholarship is that these environments organise participation less through a user's declared network and more through algorithmic selection (Bhandari and Bimo, 2022). Bhandari and Bimo capture this shift as a movement from the "networked self" to the "algorithmized self," in which self-making, recognition, and encounter are shaped by personalised recommendation rather than by prior connections. Schellewald (2025) makes a complementary case, showing that content-centric design displaces networking as the primary organising logic and positions the recommendation system as a central point of user orientation. The question that motivates this panel follows directly: what does it mean to form and sustain social ties when algorithms mediate who becomes visible to whom, and when platforms operationally define what counts as connection?

To address this question, the panel develops the concept of algorithmic sociality. We use this term to describe the conditions of social connection that emerge when recommendation systems, interface design, metrics, and governance arrangements jointly structure the possibilities of encounter, interaction, and relational boundary-making. The concept does not claim that platforms have replaced social relations with algorithms. Rather, it names a situation in which relations with algorithmic systems are constitutive of relations with other people, and in which everyday decisions about whether to engage, extend, contain, or refuse ties are shaped by platform-specific logics of visibility and evaluation.

The concept draws on scholarship that insists the "social" should not be taken for granted in social media research. Couldry and van Dijck (2015) argue that the social must be taken seriously as "a site of necessary, and necessarily contested, representation," particularly when platforms have become sites of new economic value, a context in which, as Mejiias (2013) puts it, the term "social media" functions not as description but as an appropriation of the social. Sujon and Dyer (2020) similarly calls for researchers to critically interrogate what it means to be social under contemporary digital conditions. Algorithmic sociality takes up these challenges by foregrounding the socio-technical conditions under which connection and belonging become possible.

To clarify what has shifted, it is useful to contrast recommendation-driven platforms with earlier modes of platformed sociality. In classic accounts, on social network sites (e.g., Facebook), connections were organised around the "social graph," -- the global network of linkages between all individuals within a system (cf. Fitzpatrick and Recordon, 2007). Sociality was built through pre-existing connections: users followed accounts they already knew, curated feeds based on their network, and encountered content primarily from people they had chosen to connect with (Ellison & boyd, 2013).

Under such contexts, social media sites function as relatively neutral stages, or tools for self-presentation and social connection (Bhandari and Bimo, 2022). Thus, these platforms were supplement of the existing connections, but not changing the sociality (Papacharissi, 2013).

What is distinctive about the platforms foregrounded in this panel is not that algorithms newly exist, but that content-centric recommendation becomes the primary organising experience of participation. As Bhandari and Bimo (2022) argue, TikTok's sociality differs from that offered by other platforms: interaction shifts from connections with one's social network to engagement with algorithmically curated content from strangers. Schellewald (2025) observes that TikTok's design disrupts the centrality of social networking with a form of content-centricity, where users' primary interaction partner becomes the recommendation system itself. This shift from the "networked self" to the "algorithmized self" (Bhandari and Bimo, 2022) transforms the conditions under which social ties form: relationships may emerge as consequences of algorithmic encounter rather than preconditions of content flow.

Platform studies remind us, however, that these dynamics are not reducible to algorithms acting on their own. Algorithms are components within complex platform ecosystems shaped by corporate strategy, interface design, policy, labour, and political economy. In Bucher's terms, this is what "programmed sociality" makes visible: software and computational infrastructures shape the conditions under which social relations become organised and actionable (Bucher, 2018). Schroeder (2018) comes at the problem from everyday sociability. His notion of "tethered togetherness" captures the routine practices of checking, monitoring, and responding that tether people to one another and to information, often asynchronously, as part of ordinary life. A longer historical view is offered by van Dijck's (2013) account of a culture of connectivity, which tracks how platforms engineer sociality by coding human connectedness into automated connectivity, and by treating that connectivity as a resource to be managed and monetised.

The panel's contribution lies in extending these conversations to the specific conditions of algorithmic-driven social media. We argue that algorithmic sociality produces distinctive tensions that existing frameworks do not fully address: tensions between personalised addressability and collective belonging, between frictionless connection and deliberate boundary work, and between visibility as opportunity and visibility as risk. The five papers provide empirical grounding for these tensions across different platforms, populations, and contexts.

The five papers trace a progression from human-algorithm relations to human-human connections, examining how algorithmic sociality operates at each point along this continuum.

The first paper *Affectively domesticating TikTok: audience relationships with the platform algorithm* written by Ellenrose Firth. Through 30 in-depth interviews with TikTok users aged 18-34, Firth theorises the process of affective domestication of an algorithmic entertainment app through interaction with platform affordances interpreted through the establishment of folk theories. The algorithm, personified throughout the interviews, becomes the focus of the affective relationship users develop when using the platform. Users referred to the idea that their algorithm, in some way, had gotten to know them; in turn, users felt they had to care for their algorithm. Crucially, the primary relationship that was described was not with other users, nor with content creators. Firth's contribution establishes that algorithmic sociality begins with this algorithm-facing orientation, a foundation the subsequent papers will complicate.

The second paper, *Thinking about you: playful algorithm interaction on personalised social media apps* written by Andreas Schellewald. Drawing on ethnographic research on TikTok (2020-2021) and Instagram (2023-2024), Schellewald showcases how ordinary people playfully interact with algorithmic systems to manage social relationships online. Rather than conceptualising people as "lonely" consumers of content, he argues for attention to the interpretative mechanisms and cultural practices through which people actively connect 'the personal' and 'the social' online today, doing so with algorithms not as a distraction from such connection, but rather as a tool and resource that helps to achieve social connection in new and meaningful ways. Where Firth shows users relating to algorithms, Schellewald shows users working with algorithms to reach other people, shifting the panel toward human-human connection.

Third, *Friending across platforms: how users navigate algorithmic sociality* written by Hui Lin. Lin examines multi-platform friending across Chinese platform ecosystems, conceptualising it as a form of strategic relational labour. Drawing on interviews and digital ethnography with young users across WeChat, RedNote, Bilibili, and Douyin, she identifies three logics: contextual segmentation, assigning different social roles to different platforms; algorithmic calibration, adding friends strategically to influence recommendation systems; and anticipatory connection, pre-emptively extending connections across platforms to safeguard future access. Lin's contribution reveals that connection becomes modular, strategic, and anticipatory, as users creatively inhabit, adapt to, and subtly resist the infrastructures that organise contemporary digital life. Her paper documents users who extend relationships across platforms; Wang's paper examines those who deliberately contain them.

Fourth, *Douyou but not friends: strategic invisibility and platform-bounded sociality among older content creators* written by Yingwen Wang. Drawing on 23 interviews with older Douyin creators aged 60-80 and ethnographic fieldwork, Wang develops the

concept of platform-bounded sociality. Participants formed connections they termed Douyou (Douyin friends), involving regular interaction, emotional support, and genuine affection. Yet despite these intimate ties, participants actively maintained their relationships within platform boundaries. They declined offline meetings, refused to add contacts on WeChat, and blocked unwanted followers. Wang shows these creators deployed strategic invisibility across specific dimensions, most notably offline space and cross-platform contexts. Her contribution argues that platforms function not only as infrastructures of connection but as resources for the deliberate organisation of social distance. Both Lin and Wang demonstrate strategic agency; then, Guo's paper reveals its limits.

Last paper, *Watching without speaking: restrained participation among migrants in China's urban villages on Douyin* written by Wenxin Guo, drawing on ethnographic fieldwork in a Wuhan urban village, Guo shows how migrants engage with Douyin through viewing-oriented engagement: primarily watching and liking content, rarely producing or publicly discussing images related to their own communities. This does not indicate an absence of digital participation. Rather, it represents a cautious choice shaped by the combined effects of state regulation, internalized stigma, and platform logic. Guo conceptualizes this as a structured form of digital participation functioning as a protective mechanism within a digital environment characterized by high visibility and asymmetrical risk. Where previous papers show users navigating algorithmic sociality with varying degrees of agency, Guo's paper foregrounds structural constraint, revealing whose sociality is circumscribed from the start.

Together, the five papers trace a continuum from human-algorithm relations to human-human connections, revealing three key findings. First, algorithmic mediation does not simply distribute content; it produces relational orientations. Firth shows users developing affective bonds with recommendation systems; Schellewald demonstrates how these systems become tools for achieving human connections. Second, algorithmic sociality involves deliberate boundary work. Lin documents users who strategically extend relationships across platforms; Wang reveals users who deliberately contain them within a single platform. These are not contradictory findings but complementary responses to what van Dijck's (2013) culture of connectivity, a regime in which users must actively manage how relationships are distributed across proprietary infrastructures. Third, the capacity for such strategic navigation is unevenly distributed. Guo's paper demonstrates that for marginalized populations, the same visibility that enables connections for some constitutes risk for others. This finding responds directly to Sujon and Dyer's (2020) call to examine not only what "the social" means under digital conditions but also whose sociality is enabled and whose is foreclosed.

The panel's comparative design strengthens these contributions. By examining cases across Western and Chinese platforms (TikTok/Instagram versus Douyin/Xiaohongshu), age groups (18-34 versus 60-80), and structural positions (urban users versus migrant workers), we demonstrate that the forms and stakes of algorithmic sociality vary systematically with social location. This comparative approach moves beyond single-platform or single-population studies to reveal how algorithm interact with existing social hierarchies.

Methodologically, the panel demonstrates how qualitative research can access the experience of algorithmic mediation and sociality. Interviews surface users' folk theories of recommendation; ethnography captures playful and strategic engagement with platform systems; multi-sited observation follows relationships across platform infrastructures; place-based fieldwork connects platform participation to lived vulnerability. Together, these approaches show how qualitative methods can make algorithmic sociality empirically accessible while keeping sociality as the central object of inquiry.

In sum, this panel offers algorithmic sociality as a concept for understanding how recommendation systems, platform architectures, and governance arrangements shape contemporary social connection. It brings five qualitative studies into dialogue across platforms and populations, with a focus on how algorithmic mediation becomes a condition for encounter, how users perform boundary work to organise connection, and how these processes unfold unevenly under different social conditions.

Session 10D - **PANEL: From Creative Class to Popular Class: TikTokization, AI, and the Standardization of Art, Media, and Platform Culture**

Day 3 · Wed 15 July · 10:40–12:10 · Seminar Suite C, ARC

Adam Arvidsson, Roberto Graziano (University of Naples Federico II), Maitrayee Deka (University of Essex), Sabrina Bellafronte (University of Naples Federico II), Alessandro Gandini (University of Milan), Guo Jianpeng, Vincenzo Laezza (University of Naples Federico II)

This panel seeks to embrace a situated perspective that explicitly reflects the geographical and cultural diversity of its contributors. The panel brings together presenters working on empirical cases from different parts of the world, including China, India, Sweden, Spain, Greece, and both Northern and Southern Italy. This transnational composition is not incidental, but central to the panel's analytical ambition. The processes of TikTokization and AI-mediated creativity do not unfold uniformly across the globe; rather, they are deeply shaped by local economies, infrastructures, regulatory environments, and cultural repertoires (Miller et al., 2021; Ardévol & Lanzeni, 2022). By juxtaposing research from heterogeneous contexts, the panel highlights how platform logics are continuously reworked through situated practices.

The panel addresses one of the most pressing and accelerated transformations of our time: the shift from the creative class to the popular class as the symbolic, cultural, and productive engine of the digital society. The global spread of TikTok and the increasingly popular use of artificial intelligence are redefining how culture is created, circulated, and legitimized, reversing traditional hierarchies of distinction and symbolic authority that characterized the age of elite media and cultural industries. The significance of this topic lies in its ability to interconnect social, economic, and cultural dimensions: on the one hand, the algorithmic standardization and platformization of creativity; on the other, the emergence of popular aesthetics and bottom-up practices that blur the boundaries between production and consumption, symbolic capital and social visibility.

This panel thus offers a multidisciplinary and cross-cultural framework to examine how TikTokization, popular digital culture, and the widespread adoption of AI are transforming the moral economy of platforms, creative labor, and the production of taste, value, and social distinction. The topic directly contributes to international

debates on digitalization, media sociology, and the reconfiguration of power, technology, and class in the global cultural economy.

Panel Contributions

Contribution 1, Scaling Gender Ideology through Algorithmic Advertising: An Analysis of TikTok Ads - Dr. Silvia Semenzin (AI Forensics)

Dr. Silvia Semenzin is an Italian author, scholar, and activist.

Short-form video platforms such as TikTok, YouTube, and Instagram have become central sites for the circulation of political meaning and gender imaginaries among young people, shaping gendered norms and forms of social belonging. Through influencer economies and targeted advertising, these platforms operate as hybrid media environments in which cultural expression, commercial promotion, and political narratives converge, fostering polarized gendered ecosystems in digital spaces (Baker et al., 2024). Alongside the manosphere, encompassing alpha male influencers, incel communities, and pick-up artist cultures articulating hypermasculine and anti-feminist worldviews (Ging, 2019), the so-called femosphere (Kay, 2025) has emerged as a relational counterpart, characterized by hyperfeminine lifestyles and anti-feminist narratives through which gendered power is reproduced via commodified performances of femininity. The paper examines advertising as a socio-technical mechanism through which these gendered narratives are rendered visible, normalized, and scaled within platform ecosystems. Focusing on TikTok, the study draws on a digital research infrastructure developed by AI Forensics to collect and analyse advertising content on TikTok Ads, enabling the identification of recurrent narrative frames and visual strategies. By showing how gender ideologies are amplified and monetized through opaque advertising and recommendation systems, this research contributes to sociological and feminist debates on platform power, online misogyny, and algorithmic governance, extending analyses of the manosphere beyond subcultural spaces to the economic logics of digital platforms.

Contribution 2, Performing or Performative Care? The Symbolic Market of Digital Motherhood and the Reproduction of the Popular Habitus on TikTok -Georgios K. Bougiouris (University of Athens) & Era Domi (University of Crete)

Drawing on Pierre Bourdieu's theoretical framework, this case study analyzes the digital field of motherhood by examining how maternal care is transformed into a performative practice aimed at economic and symbolic gain. Based on a qualitative content analysis of short-form videos from Greece and Italy, the paper shows how TikTok operates as a competitive symbolic market in which symbolic violence is reproduced through the platformization of care labor. The analysis focuses on micro-influencers embodying a popular habitus and highlights how the "honest mom" narrative functions as a

mechanism of misrecognition that masks the economic conversion of maternal identity into capital while reinforcing gender inequalities.

Contribution 3, Gentrified Labor: The Resignification and Platformisation of Manual Work - Alessandro Gandini (University of Milan)

Drawing on the Marxist labor process tradition, the paper introduces the concept of gentrified labor to explain how certain forms of manual and material work are resignified as culturally valuable in response to the declining promise of meaningful white-collar employment. The paper argues that this process involves both the culturalization of manual labor and its platformization through digital infrastructures, transforming manual work into an experiential consumer good. In doing so, gentrified labor becomes a strategy of middle-class reproduction, whereby middle-class actors appropriate the material and symbolic repertoires historically associated with working-class life in order to maintain cultural privilege and social status.

Contribution 4, AI-Generated Music: Platform Ecologies, Model Design, and Emerging Creative Practices - Vincenzo Laezza (University of Naples “Federico II”)

This paper examines the generative AI music platform Suno as a case study to explore how AI technologies operate as technical forms that shape cultural and social practices within contemporary platform ecologies. Building on critical app studies, the analysis emphasizes how model architectures, interface design, and platform logics actively frame what kinds of musical expression become possible, valuable, and shareable. The paper shows how generative AI systems are embedded within platform environments characterized by remixing, circulation, and algorithmic recommendation, raising broader questions about creativity, authorship, labor, and ownership in AI-mediated cultural production.

Contribution 5, From Compliance to Branding: Regulatory Adaptation of China’s Digital Platforms - Guo Jianpeng (Nanyang Technological University, Singapore) & Marica Castaldi (Università degli Studi di Napoli Federico II)

The paper introduces the COPE model of branded compliance, consisting of Compliance, Orchestration, Performance, and Externalisation, to analyze how Chinese digital platforms adapt to strict and fluctuating regulatory regimes. Through a longitudinal analysis of short-video platforms and key regulatory events, the study shows how compliance evolves from a survival-oriented practice into a symbolic and economic resource that becomes integrated into platform growth strategies. The paper highlights the specificity of China’s regulatory adaptation model and its implications for global debates on platform governance and digital monopolies.

Session 11A - **Workshop: Would a Poem by Any Other Platform Sound As Sweet?: Examining Social Media Metrics through Sentiment Analysis**

Day 3 · Wed 15 July · 13:30–14:50 · Seminar Suite A, ARC (Please note – there is a coffee and comfort break at 15:00)

JuEunhae Knox (University of Stirling)

This workshop is structured around three key objectives, moving from conceptual framing through methodological comparison to practical implementation.

1) My first aim is demonstrating how platform infrastructures differ in how they regulate the dynamics of digital cultural objects, including audience commentary itself. In dialogue with actor-network perspectives, I seek to position platforms not as neutral environments but as active forces that condition participation and visibility, and reception. By exploring the ways audience responses are divergently prioritised, nested, or hidden on TikTok and Instagram, the workshop enables participants to test how audience responses become embedded within digital-cultural objects themselves and study the impact that distinct platform architectures exert upon supposedly independent comments.

2) The workshop then models a parallel convergent approach in which qualitative and quantitative analyses are used alongside one another rather than in opposition. Through controlled comparative case studies of cross-platform poetry, attendees move between critical discourse analysis and aggregate examination of commentary patterns. By exploring how each mode of inquiry challenges and enriches the other, participants will examine whether the effects of algorithmic curation on engagement dynamics become more visible through mixed methods.

3) A final core objective is equipping attendees with hands-on experience with accessible tools for ethically gathering and analysing platform commentary, including an introduction to responsible data extraction practices through Python-based notebooks. Working with sentiment scoring and thematic analysis, participants contrast how different models generate disparate readings of the same dataset. This process highlights where automated classification diverges from human evaluation, critically reflecting how our methodological choices shape interpretation of platform-mediated engagement.

Session 11B - Workshop: Collaging the Datafied Self: Mapping Postdigital Subjectivity through Creative Methods

Day 3 · Wed 15 July · 13:30–14:50 · Seminar Suite B, ARC (Please note – there is a coffee and comfort break at 15:00)

Liv Owens (City St George's, London), Chiara Fehr (UCL Institute of Education)

This workshop aims to demonstrate the value of creative methodologies in researching the affective lived experience of post-digital, datafied subjects. We argue that communal arts-based practices of disassembling and re-assembling generate meaningfully reflexivity around the fragmentation of self which users are subjected to on social media. Ultimately this workshop is proposed as part of a methodology for those researching datafication amongst a variety of contexts.

As digital platforms increasingly mediate everyday life, users experience themselves as collections of data points, behavioural predictions, and algorithmic profiles. These fragmented representations generate diverse affective realities and contradictions that are difficult to access through conventional, screen-based methods alone.

The primary objective of the session is to demonstrate how creative, participatory, craft-based research methods can surface tacit forms of knowledge that are often overlooked in digital research. By engaging participants in hands-on collage-making, the workshop will show how material practices can render postdigital datafication processes tangible and discussable.

A second objective is to facilitate participants' exploration of their own experiences as datafied bodies, foregrounding how digital systems actively shape feelings, behaviours, and self-understandings. Through a guided creative practice, participants will be encouraged to reflect on the tension between their active online self-representation and the processes of data capture, profiling, and algorithmic inference that occur covertly, beyond their control or visibility.

Finally, the workshop seeks to equip participants with transferable methodological tools that can be applied in their own research and teaching. Participants will leave with a critical understanding of how creative methods can be used to investigate postdigital subjectivity. Additionally, they will gain practical experience in facilitating participatory sessions that centre communal and creative reflexivity.

Session 11C - Workshop: Resurrecting platform pasts: A workshop for research in a changing platform landscape

Day 3 · Wed 15 July · 13:30–14:50 · Seminar Suite C, ARC (Please note – there is a coffee and comfort break at 15:00)

Frances Corry (University of Pittsburgh), Chelsea Butkowski (American University's School of Communication)

Scholars have been writing about social media platforms (as we know them) for nearly thirty years. Recently, the social media ecosystem has undergone a series of rapid changes, as major platforms restructure, de-centralized platforms grow, and new platforms emerge. This workshop will allow participants to reflect on the shifting nature of social media research in this time as echoed in the shifting platform landscape itself.

In this workshop, we will ask participants to reflect on the legacies of salient social media scholarship examining platforms (like Vine, Fotolog, MySpace, Livejournal, GeoCities, FriendFeed, Orkut and iWiW) that are today significantly changed or defunct, to address what we can learn when we revisit social media research insights after the object of study has receded or no longer exists. Which research insights are specific to a platform being studied, and which observations—about community, self-presentation, privacy, publicity, virality, and influence—continue, or are intensified in new environments, and in what ways?

At the end of our session, we hope to have built momentum for a collaborative effort that can address major intellectual threads in social media scholarship across time and platforms.

During our workshop, we aim to 1) create an intellectual foundation for studying both consistency and change across the platform environment. This will be accomplished through activities that ask participants to reflect on how social media research has changed as platforms themselves have changed. We hope to accomplish this in part by 2) connecting scholars inter- and intra-generationally, especially connecting early career social media scholars with more established scholars who may have direct experience studying platforms that are now considered defunct. This establishes the basis for the tangible outcome of this workshop, which is to 3) develop a collaborative scholar network and, in the future, an edited volume on “resurrecting platform pasts.”

Session 11D - **Workshop: Beyond the Hairball: An Applied Approach to Social Network Analysis with Communalytic**

Day 3 · Wed 15 July · 13:30–14:50 · Studio 2, ARC (Please note – there is a coffee and comfort break at 15:00)

Anatoliy Gruzdt (Ted Rogers School of Management), Philip Mai (Toronto Metropolitan University)

The primary objective of this tutorial is to demystify Social Network Analysis (SNA) for interdisciplinary scholars and equip them with the methodological tools to apply relational thinking to social media data. As social media platforms are inherently networked structures, standard statistical methods often fail to capture the complex interdependencies between users, content, and interactions.

This tutorial has three core learning objectives. First, we aim to ground participants in the theoretical foundations of SNA, enabling them to formulate research questions that move beyond attribute-based analysis (e.g., "who posted this?") to relational analysis (e.g., "how did this information diffuse?"). We will cover key concepts such as centrality, homophily, and structural roles within the context of platforms like Bluesky and Reddit.

Second, the session will provide hands-on technical training on Communalytic (<https://communalytic.org/>), a research platform for collecting and analyzing social media data. Participants will learn how to use this web-based tool to collect data, visualize communication patterns, and prepare datasets for advanced analysis; all without requiring prior coding knowledge.

Third, we aim to foster critical visual literacy. Network visualizations are often aesthetically pleasing but analytically opaque. We will show participants how to use Communalytic's built-in Network Analyzer Module to identify latent community structures and influential actors.

Session 12 – Poster Presentations

Day 2 · Tues 14 July · 18:00–22:00 Dining Hall and Drawing Room, Glasgow University Union

12.1 Dynamics of Perceived Control in Finnish Young Adults' Social Media Use

Pinja Akkanen (University of Turku)

Introduction

The study explores how Finnish young adults (aged 18–35) perceive control over their social media use. The data consists of 572 responses to an online survey that included both close- and open-ended questions.

Over the last decade young adults have lived through social media's shift from enjoyment to management. In 2012, users described social media as fun, leisure-oriented, and a way to stay connected. By 2022, social media use was increasingly associated with challenges. These include limiting screen time, avoiding endless scrolling, and protecting themselves from negative effects like stress and anxiety. Social media use has become something that requires users to manage and regulate it. (Jungselius & Weilenmann, 2025).

Simultaneously, both scholarly and media discourse increasingly frame heavy social media use in terms of addiction, a framing that may lead users to view their own social media use as addictive. Prior research suggests that such labeling can have adverse consequences, including diminished perceived control and increased self-blame for overuse. (Anderson & Wood, 2025). Moreover, measures of social media addiction often rely on scales ranging from “not at all” to “extremely” addicted, which provide limited scope for capturing individuals' perceptions of control (Lee & Hancock, 2024). Against this backdrop, the present study examines young adults' perceived control, defined here as individuals' subjective sense of the extent to which they manage and regulate their use.

In Finland, 76% of young adults spend on average at least three hours a day on social media, and 37% spend more than five hours a day (E2 Tutkimus, 2026), indicating that social media plays a substantial role in their everyday lives. However, it is also important to look beyond screen time and consider how individuals perceive and make sense of the time they spend on social media (Lee et al., 2021). To address this gap, the present study adopts a mixed-methods approach that combines quantitative and

qualitative data. It draws on the concept of an agency social media mindset, understood as the core assumptions individuals hold about how much agency they have over their social media use (Lee & Hancock, 2024). This perspective shifts attention away from usage metrics toward individuals' lived experiences of social media use (Gladkaya et al., 2025).

Objectives/Research Questions

The study examines whether young adults perceive themselves as being in control of their social media use and how they define what having control means. To address these aims, the study is guided by the following research questions:

RQ1: What do young adults perceive as having control over their social media use?

RQ2: Do young adults perceive themselves as being in control of their social media use?

Methods

The study employed a cross-sectional mixed-methods design. Data were collected through an online survey conducted in Finland during April and May 2025. A total of 572 participants were recruited through targeted communication via social media platforms and a university mailing list. To be eligible, participants had to be between 18 and 35 years of age and able to complete the survey in Finnish. 85,5% of participants were women.

Participants were asked to reflect on their perceived agency over social media use and the valence of social media's effects on their lives. These assumptions were measured using the Social Media Mindset Scale (Lee & Hancock, 2024), which includes 12 statements related to social media.

First, the results were examined by calculating the mean scores of all participants for the closed-ended survey items, such as screen time and phone pickups. Based on the Social Media Mindset Scale (the SMMS) mean scores were computed for both the agency and valence mindsets, as well as for each individual statement.

Participants were then categorized into three groups according to their agency social media mindset scores: low agency mindset (1.0–2.4), moderate agency mindset (2.5–3.5), and high agency mindset (3.6–5.0). Subsequently, other survey results were compared across these groups.

To explore young adults' perceived control over their social media use, open-ended responses to the question "Do you feel that you are in control of your social media use? Why or why not?" were analyzed using qualitative content analysis. The results were examined in three sections according to agency mindset groups. The analysis focused on how participants explained why they perceived their social media use as either controlled or uncontrolled, as well as how strongly this sense of control was expressed.

In addition, the analysis identified descriptions of digital disconnection strategies and references to factors perceived to influence control over social media use.

Results

The quantitative findings of this study support earlier research showing that people who feel more in control of their social media use tend to use it less (Parry & Coetzee, 2025; Lee & Hancock, 2024). However, the open-ended responses indicate that many young adults perceive their social media use as high, or even excessive, while still feeling in control of their use. Rather than being determined solely by the amount of use, perceived control appears to be shaped by how social media use is experienced in everyday life.

The findings indicate that young adults who perceive a high level of control over their social media use describe their behavior as intentional and clearly bounded. Their use is less likely to interfere with other areas of life, and they report being able to abstain from social media for short or long periods when they choose to do so. In contrast, young adults who perceive lower control over their social media use tend to describe it as excessive and difficult to disengage from. Their use is often experienced as unintentional and characterized by aimless scrolling. They also perceive social media as interfering with other aspects of their lives.

Although the findings reveal shared patterns in how young adults describe varying levels of perceived control over social media use, experiences of control are shaped by individual standards. What is perceived as controlled use varies between individuals. The findings further suggest that maintaining control over social media use requires ongoing effort. This was reflected in participants' narratives of various digital disconnection strategies used to control their social media use.

Overall, the findings suggest that while approximately half of young adults perceive themselves as being in control of their social media use, the qualitative responses reveal a more nuanced picture. Perceived control should not be understood as a fixed or stable state, as many young adults experience control over their social media use as situational, variable and partial.

Earlier research on social media addiction often conceptualizes use along a continuum from non-addictive to addictive forms (Lee & Hancock, 2024). Within this framework, controlled and healthy social media use is typically positioned as the opposite of addiction. In contrast, this article argues that perceived control over social media use should itself be understood as a continuum that reflects the dynamic and situational nature of everyday use. The continuum ranges from uncontrolled use through slipping, variable, and mostly controlled use, to controlled use. This perspective recognizes that fluctuations in use and occasional lapses are common and do not necessarily indicate problematic behavior.

Future Work

The study constitutes the first article of my doctoral dissertation. Subsequent research will examine what kinds of support young adults hope for in controlling their social media use. Building on these insights, I will develop an intervention aimed at supporting young adults in controlling their social media use. The two week intervention consists of several support methods that are introduced to participants gradually.

References

- Anderson, I.A. & Wood, W. (2025). Overestimates of social media addiction are common but costly. *Sci Rep* 15, 39388. <https://doi.org/10.1038/s41598-025-27053-2>
- E2 Tutkimus: Mistä lääkkeit someajan pahoinvointiin? Tutkimus työikäisten nuorten keinoista hillitä älylaitteiden ja sosiaalisen median kielteisiä vaikutuksia, 2026. <https://www.e2.fi/julkaisut/julkaisut/e2-tutkimus-elo-ja-yths-selvittivat-enemmisto-nuorista-aikuisista-yrittäa-vahentaa-alylaitteiden-kayttoa-yli-80-prosenttia-arvioi-kayttavansa-laitteita-enemman-kuin-haluaisi.html>
- Gladkaya, M., Baum, K., Logemann, H. & Krasnova, H. (2025). "“I Am in Control!” An Intensive Longitudinal Study of the Role of Agentic Mindsets in User Well-Being". ICIS2025 Proceedings. 7. https://aisel.aisnet.org/icis2025/is_media/is_media/7
- Jungselius, B., & Weilenmann, A. (2025). Tracing Change in Social Media Use: A Qualitative Longitudinal Study. In CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan. ACM, New York, NY, USA. <https://doi.org/10.1145/3706598.3713813>
- Lee, A. Y., Katz R. & Hancock J. (2021). The Role of Subjective Construals on Reporting and Reasoning about Social Media Use. *Social Media + Society*, July-September 2021: 1–13. <https://doi.org/10.1177/20563051211035350>
- Lee A. Y., Ellison N. B., & Hancock J. T. (2024). Social media mindsets: a new approach to understanding social media use and psychological well-being, *Journal of Computer-Mediated Communication*, 29(1). <https://doi.org/10.1093/jcmc/zmad048>
- Parry, D. A. & Coetzee, B. (2025). Do mindsets really matter? A second look at how perceptions of social media experiences relate to well-being, *Journal of Computer-Mediated Communication*, 30(4). <https://doi.org/10.1093/jcmc/zmaf011>

12.2 Perceived Influence of Health Misinformation on Social Media: A Qualitative Study through the Lens of Third-Person Effect Theory

Wed Alfotawi (Bournemouth University)

Introduction

Health misinformation on social media significantly shapes users' beliefs and behaviours, particularly during major health crises such as the COVID-19 pandemic (Gwon et al. 2024). Exposure to such content can influence disease perceptions, risk assessments, and public health compliance (Rocha et al. 2021; Gupta et al. 2022). For instance, COVID-related misinformation and conspiracy theories have led some to believe the virus was man-made and have contributed to rejection of health measures like mask-wearing and distancing (Roozenbeek et al. 2020; Atehortua and Patino 2021; Kim and Tandoc Jr 2022). Conspiracy beliefs may also undermine trust in credible health institutions (Jennings et al. 2021; Vaal et al. 2023).

These influences often operate outside conscious awareness, and individuals differ in how susceptible they believe they are to such misinformation. Some consider themselves immune, while viewing others as more vulnerable, based on different assumptions about media literacy, knowledge, or social roles. This dynamic is especially relevant in Saudi Arabia, where people rely heavily on social media and government sources for health information (Alsulaiman 2018). High trust in Ministry of Health guidance, along with strong social ties and cultural norms, further complicates how influence is perceived (Almutairi et al. 2020). Simultaneously, the presence of widespread misinformation on Saudi platforms creates a paradox that merits deeper investigation (Alotaibi and Alhammad 2022).

This study draws on Third-Person Effect Theory, which posits that individuals tend to believe others are more influenced by media content, especially harmful or misleading content, than themselves (Davison 1983; Perloff 1993; Gunther and Storey 2006). Applying this lens, the study explores how Saudi social media users perceive the impact of health misinformation on both themselves and others. It examines how users justify these perceptions, how they differentiate between self and other in terms of susceptibility, and how these judgments are shaped by trust, digital habits, social norms, and perceived expertise.

By analysing users' interpretations and reasoning, the study aims to understand how ideas of personal immunity and others' vulnerability are constructed and sustained within everyday social meaning-making. This offers insights into the cultural and

psychological dimensions of health misinformation perception in Saudi Arabia and contributes to broader theoretical discussions on media effects and public trust.

Objectives/Research Questions

Research Objectives:

1. To examine how social media users in Saudi Arabia perceive the influence of health misinformation on themselves and on others.
2. To identify and analyse the explanations participants provide for perceived influence on the self versus others, including beliefs about personal immunity, vulnerability and credibility judgements.
3. To examine the role of social relationships and cultural norms, e.g., family and peer ties, in shaping perceptions of being affected by health misinformation.

Research Questions:

1. How do social media users in Saudi Arabia perceive the influence of health misinformation on their own beliefs, emotions, and health-related behaviours?
2. How do they perceive the influence of health misinformation on others?
3. What reasons and interpretive frames do participants use to explain perceived influence on the self-versus others?

Methods

This study adopts face-to-face semi-structured interviews as its primary method, given their suitability for exploring participants' nuanced perceptions, attitudes, and experiences regarding health misinformation (Shah et al. 2012; Kallio et al. 2016). This approach allowed for open, in-depth discussion and enabled the researcher to observe non-verbal cues, such as tone, facial expressions, and physical tension, which added depth to the analysis (Barriball & While 1994; Adeoye-Olatunde & Olenik 2021). The study investigates how Saudi social media users perceive their own and others' susceptibility to health misinformation, using Third-Person Effect (TPE) theory as a guiding framework.

Participants reflected on issues such as trust in sources, content credibility, and the psychological states that may shape both belief and sharing behaviours, such as anxiety, stress, or overconfidence. Semi-structured interviews offered a space to articulate these dimensions clearly, especially in relation to personal immunity and perceived vulnerability of others (Wang & Fussell 2020; Metzger et al. 2021). The method enabled participants to explain their justifications for perceived influence and to describe how social and cultural norms inform these beliefs.

A purposive sampling strategy was employed to ensure alignment with the study's objectives and to maximise the relevance of participant insights (Ames et al. 2019; Andrade 2021). While this sampling method has limitations in terms of generalisability (Rai & Thapa 2015), it was appropriate for this qualitative study's focus on depth and context (Tashakkori & Teddlie 2021). The inclusion criteria were: (1) Saudi nationals residing in Saudi Arabia at the time of interview, (2) frequent social media users (defined as ≥ 2 hours/day), and (3) following at least one pre-identified account on Platform X.

Participants were recruited via direct private messages and interviews were conducted in person with 32 individuals. Data were analysed thematically using an inductive approach (Fereday & Muir-Cochrane 2006; Guest et al. 2011), allowing themes to emerge from the data without being confined by a pre-set coding framework. Although theory informed the study design, the analysis remained grounded in participants' own language and meaning-making processes (Denzin & Lincoln 2011; Gray 2021), ensuring a bottom-up understanding of how users perceive and interpret the influence of health misinformation in their digital environments.

Results

1. Perceived Immunity to Health Misinformation

Most participants described themselves as immune to health misinformation, despite acknowledging exposure. This self-image was justified by their confidence in personal judgement, education, and critical thinking. Experiences during the COVID-19 pandemic were framed as evidence of their resistance, reinforcing a contrast between self and more gullible others.

2. Conditional Susceptibility

Some participants acknowledged rare moments of influence under specific conditions, such as illness, stress, or pain, when they were more willing to try unverified remedies. In a few cases, trust in familiar or reputable sources reduced scrutiny. Though described as exceptions, these cases revealed subtle and situational vulnerability.

3. Discrepancy with Evidence

Participants' strong claims of immunity contrast with research showing that misinformation can influence beliefs and behaviours, even unconsciously. They may not recognise gradual influence, especially when no immediate harm occurs. Influence may also manifest through sharing misinformation, not personal belief, displacing impact onto others.

4. Emotion Centred on Others

Emotional responses to misinformation were primarily directed at concern for others, particularly older relatives or those with health issues. Participants reported frustration,

anxiety, or sadness when loved ones were exposed to or acted upon false claims. The issue was framed as a social threat, not a personal one.

5. Perceived Vulnerability of Others

Participants viewed other Saudi users as easily misled, quick to believe, and generally lacking critical awareness. These descriptions closely mirror the third-person effect, where individuals perceive media as more influential on others than themselves. Influence was not denied, but consistently located outside the self.

6. Explanations for Others' Susceptibility

Others' vulnerability was explained through perceived deficits in health literacy, age-related digital inexperience (especially older adults), and gendered assumptions, often portraying women as more susceptible to appearance-related or family content. Additional factors included emotional distress, confusion from conflicting information, or excessive trust in authority.

Future Work

Future research could extend the findings of this study by examining perceived susceptibility to health misinformation across a broader demographic spectrum, including younger adolescents and older adults, or by conducting comparative studies across cultural contexts. Quantitative approaches could complement the qualitative findings by measuring the prevalence of third-person perceptions and their behavioural consequences. Experimental designs could also assess how message framing or source credibility affects susceptibility to misinformation. Further work could explore the role of platform-specific features “e.g., algorithmic exposure on TikTok or Instagram” in shaping perception and sharing behaviour. Finally, applied research is needed to test interventions aimed at strengthening critical media literacy and reducing the circulation of health misinformation in social networks.

References

- Adeoye-Olatunde, O. A., & Olenik, N. L. (2021). Research and scholarly methods: Semi-structured interviews. *Journal of the American College of Clinical Pharmacy*, 4(10), 1358-1367.
- Almutairi, A. F., BaniMustafa, A. a., Alessa, Y. M., Almutairi, S. B., & Almaleh, Y. (2020). Public trust and compliance with the precautionary measures against COVID-19 employed by authorities in Saudi Arabia. *Risk Management and Healthcare Policy*, 753-760.
- Alotaibi, F. L., & Alhammad, M. M. (2022). Using a rule-based model to detect arabic fake news propagation during Covid-19. *International Journal of Advanced Computer Science and Applications*, 13(1).

- Alsulaiman, S. A. (2018). Health crisis in the kingdom of Saudi Arabia: A study of Saudis' knowledge of Coronavirus, attitudes toward the ministry of health's coronavirus preventive campaigns, and trust in coronavirus messages in the media [Bowling Green State University].
- Ames, H., Glenton, C., & Lewin, S. (2019). Purposive sampling in a qualitative evidence synthesis: A worked example from a synthesis on parental perceptions of vaccination communication. *BMC medical research methodology*, 19, 1-9.
- Andrade, C. (2021). The inconvenient truth about convenience and purposive samples. *Indian journal of psychological medicine*, 43(1), 86-88.
- Atehortua, N. A., & Patino, S. (2021). COVID-19, a tale of two pandemics: novel coronavirus and fake news messaging. *Health Promotion International*, 36(2), 524-534.
- Barriball, K. L., & While, A. (1994). Collecting data using a semi-structured interview: a discussion paper. *Journal of Advanced Nursing-Institutional Subscription*, 19(2), 328-335.
- Davison, W. P. (1983). The third-person effect in communication. *Public opinion quarterly*, 47(1), 1-15.
- Denzin, N. K., & Lincoln, Y. S. (2011). *The Sage handbook of qualitative research*. sage.
- Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International journal of qualitative methods*, 5(1), 80-92.
- Gray, D. E. (2021). Doing research in the real world. *Doing research in the real world*, 1-100.
- Guest, G., MacQueen, K. M., & Namey, E. E. (2011). *Applied thematic analysis*. sage publications.
- Gunther, A. C., & Storey, J. D. (2006). The Influence of Presumed Influence. *Journal of Communication*, 53(2), 199-215. <https://doi.org/10.1111/j.1460-2466.2003.tb02586.x>
- Gupta, A., Li, H., Farnoush, A., & Jiang, W. (2022). Understanding patterns of COVID infodemic: A systematic and pragmatic approach to curb fake news. *Journal of business research*, 140, 670-683.
- Gwon, N., Jeong, W., Kim, J. H., Oh, K. H., & Jun, J. K. (2024). Effects of Intervention Timing on Health-Related Fake News: Simulation Study. *JMIR Formative Research*, 8(1), e48284.

- Jennings, W., Stoker, G., Bunting, H., Valgarðsson, V. O., Gaskell, J., Devine, D., McKay, L., & Mills, M. C. (2021). Lack of trust, conspiracy beliefs, and social media use predict COVID-19 vaccine hesitancy. *Vaccines*, 9(6), 593.
- Kallio, H., Pietilä, A. M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. *Journal of advanced nursing*, 72(12), 2954-2965.
- Kim, H. K., & Tandoc Jr, E. C. (2022). Consequences of online misinformation on COVID-19: two potential pathways and disparity by eHealth literacy. *Frontiers in psychology*, 13, 783909.
- Metzger, M. J., Flanagin, A. J., Mena, P., Jiang, S., & Wilson, C. (2021). From dark to light: The many shades of sharing misinformation online. *Media and Communication*, 9(1), 134-143.
- Minhat, H. (2015). An overview on the methods of interviews in qualitative research. *International Journal of Public Health and Clinical Sciences*, 2(1), 210-214.
- Nela, A., & Parruca, E. (2023). Impact of Social Media Disinformation and of Fake News Overexposure on the Actual Capacities and the Psychological Wellbeing During the COVID-19 Pandemic: A Systemic Literature Review. *Glob. Psychother*, 3, 126-132.
- Nyimbili, F., & Nyimbili, L. (2024). Types of Purposive Sampling Techniques with Their Examples and Application in Qualitative Research Studies. *British Journal of Multidisciplinary and Advanced Studies*, 5(1), 90-99.
- Perloff, R. M. (1993). Third-person effect research 1983–1992: A review and synthesis. *International Journal of Public Opinion Research*, 5(2), 167-184.
- Rai, N., & Thapa, B. (2015). A study on purposive sampling method in research. *Kathmandu: Kathmandu School of Law*, 5(1), 8-15.
- Rocha, Y. M., De Moura, G. A., Desidério, G. A., De Oliveira, C. H., Lourenço, F. D., & de Figueiredo Nicolete, L. D. (2021). The impact of fake news on social media and its influence on health during the COVID-19 pandemic: A systematic review. *Journal of Public Health*, 1-10.
- Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L., Recchia, G., Van Der Bles, A. M., & Van Der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society open science*, 7(10), 201199.
- Shah, S. K., Arora, S., Skipper, B., Kalishman, S., Timm, T. C., & Smith, A. Y. (2012). Randomized evaluation of a web based interview process for urology resident selection. *The Journal of urology*, 187(4), 1380-1384.

Tashakkori, A., & Teddlie, C. (2021). Sage handbook of mixed methods in social & behavioral research. SAGE publications.

Vaal, S., Schofield, M. B., Baker, I. S., & Roberts, B. L. (2023). Narcissism, national narcissism, COVID-19 conspiracy belief, and social media use as predictors of compliance with COVID-19 public health guidelines. *Current Psychology*, 42(30), 26868-26875.

Wang, L., & Fussell, S. R. (2020). More than a click: exploring college students' decision-making processes in online news sharing. *Proceedings of the ACM on Human-Computer Interaction*, 4(GROUP), 1-20.

Zanatta, E. T., Wanderley, G. P. d. M., Branco, I. K., Pereira, D., Kato, L. H., & Maluf, E. M. C. P. (2021). Fake news: the impact of the internet on population health. *Revista da Associação Médica Brasileira*, 67(7), 926-930.

12.3 Sensation seeking, online risks and digital skills: Untangling the pathways in adolescence.

*Hansika Ambahelagedara, Hansika Ambahelagedara, Hansika Ambahelagedara,
Florenxia Sorthaix (University of Helsinki)*

Introduction

Adolescence is a critical developmental period characterised by heightened sensation seeking and immature self-regulation (Steinberg et al., 2018). Defined as the pursuit of novel, intense, and varied experiences alongside a willingness to take risks (Zuckerman, 1994), sensation seeking peaks during middle and late adolescence (Steinberg et al., 2018) and is positively associated with both offline and online risk-taking behaviour (Livingstone & Görzig, 2014; Siraj et al., 2021). Specifically, higher sensation seeking increases the likelihood of exposure to harmful online content, including unintended exposure (Tercova & Smahel, 2025).

Digital skills represent another key factor shaping adolescents' online experiences. While greater digital competence is linked to more time online and consequently greater exposure to online harm (Cabello-Hutt et al., 2018; Machackova et al., 2024), it has also been argued to buffer against the negative consequences of such exposure (Sonck & de Haan, 2013). Yet how sensation seeking and digital skills jointly evolve and interact over time remains poorly understood. The present study is an attempt to address this gap and to explore the dynamic relationships among sensation seeking, digital skills and exposure to online risks.

Objectives/Research Questions

RQ1: How does sensation seeking relate to intended and unintended exposure to online risks across adolescence, both within individuals over time and between individuals?

RQ2: How do digital skills relate to intended and unintended exposure to online risks across adolescence, both within individuals over time and between individuals?

RQ3: How are sensation seeking and digital skills associated with one another across adolescence, both within individuals over time and between individuals?

Methods

The data was derived from a longitudinal project (ySKILLS) that took place from 2021 to 2023 annually across 6 European countries (Estonia, Finland, Germany, Italy, Poland and Portugal). (Machackova et al., 2024). Data was collected from 10821 participants across the six countries (Mean age = 15.32; girls = 48.30%; boys = 48.83%; NA = 2.89%).

Sensation seeking was measured using the brief sensation seeking scale (4 items) and digital skills were measured by the youth digital skills indicator that was designed as a part of the ySKILLS project. The Digital skills have 24 items across 4 domains; technical and operational skills, information navigation and processing skills, communications and interaction skills and content creation and production skills. Intended exposure to online risks was measured by asking whether they have been exposed to cyberhate, harmful content, sexting (receiving) and sexual content when they intended to see such content and unintended exposure to online risks was measured by asking whether it was accidental that they were exposed to the above-mentioned risk types.

Random-Intercept Cross-Lagged Panel Model (RI-CLPM) was estimated to explore both stable between-person differences from within-person dynamics over time. Random intercepts were specified for sensation seeking and digital skills and exposure to online risks, with intended and unintended risk exposure loading onto a single shared random intercept, given their strong correlation at the between person level while at the within-person level, all four constructs were modelled separately to capture their distinct temporal dynamics. The analysis was carried out in R.

Results

Sensation seeking demonstrated significant carry-over effects from Wave 1 to Wave 2 ($\beta = .180$) and Wave 2 to Wave 3 ($\beta = .283$), as did digital skills (W1→W2: $\beta = .272$; W2→W3: $\beta = .311$). Both intended and unintended online risk exposure were similarly stable within persons across waves (intended: $\beta = .352$ and $.408$; unintended: $\beta = .388$ and $.395$). Within individuals, sensation seeking at Wave 1 significantly predicted higher intended online risk exposure at Wave 2 ($\beta = .093$), though this effect was not significant from Wave 2 to Wave 3. Sensation seeking did not predict unintended risk exposure at either transition. In the reverse direction, neither intended nor unintended online risk exposure significantly predicted within-person changes in sensation seeking at any wave. Digital skills

did not significantly predict within-person changes in either intended or unintended online risk exposure at any wave, nor did online risk exposure predict changes in digital skills. Within individuals, digital skills at Wave 2 significantly predicted higher sensation seeking at Wave 3 ($\beta = .093$, $p = .014$), suggesting that adolescents who develop stronger digital skills may subsequently report elevated sensation seeking. However, sensation seeking did not predict within-person changes in digital skills at either transition.

At the trait level, the random intercept correlations revealed meaningful stable associations across all three constructs. Adolescents who were higher in sensation seeking on average also tended to score higher on digital skills ($r = .197$) and to exhibit greater online risk exposure ($r = .424$). Similarly, adolescents with persistently higher digital skills showed somewhat higher stable levels of online risk exposure ($r = .183$). These findings suggest that, across adolescence, sensation seeking and digital skills are both positively linked to chronic tendencies toward online risk exposure at the between-person level.

Future Work

The present study examined the longitudinal associations between sensation seeking, digital skills, and both intended and unintended online risk exposure of young people. Findings reveal that stable between-person differences play a substantial role during adolescence. At the between-person level, higher sensation seeking and digital skills were associated with greater online risk exposure overall. Within persons, sensation seeking predicted subsequent intended online risk exposure, and digital skills predicted increases in sensation seeking over time, suggesting a dynamic interplay between these constructs across adolescence.

The limited number of significant cross-lagged effects may reflect the relatively long intervals between waves that is incapable of capturing the temporal dynamics in young people's everyday lives and these preliminary findings warrant further investigation. Understanding how sensation seeking and digital competence jointly shape online risk trajectories is critical for developing timely, targeted digital literacy interventions that account for individual differences rather than treating adolescents as a homogeneous group.

References

Cabello-Hutt, T., Cabello, P., & Claro, M. (2018). Online opportunities and risks for children and adolescents: The role of digital skills, age, gender and

- parental mediation in Brazil. *New Media & Society*, 20(7), 2411–2431.
<https://doi.org/10.1177/1461444817724168>
- Livingstone, S., & Görzig, A. (2014). When adolescents receive sexual messages on the internet: Explaining experiences of risk and harm. *Computers in Human Behavior*, 33, 8–15. <https://doi.org/10.1016/j.chb.2013.12.021>
- Machackova, H., Bedrosova, M., Muzik, M., Zlamal, R., Fikrlova, J., Literova, A., Dufkova, E., Smahel, D., Boomgaarden, H., Song, H., Tolochko, P., d’Haenens, L., Joris, W., Kalmus, V., Tikerperi, M.-L., Opermann, S., Napp, M., Soidla, I., Uibos, A., ... Helsper, E. J. (2024). Digital skills among youth: A dataset from a three-wave longitudinal survey in six European countries. *Data in Brief*, 54, 110396. <https://doi.org/10.1016/j.dib.2024.110396>
- Siraj, R., Najam, B., & Ghazal, S. (2021). Sensation Seeking, Peer Influence, and Risk-Taking Behavior in Adolescents. *Education Research International*, 2021(1), 8403024. <https://doi.org/10.1155/2021/8403024>
- Sonck, N., & de Haan, J. (2013). How The Internet Skills Of European 11- To 16-Year-Olds Mediate Between Online Risk And Harm. *Journal of Children and Media*, 7(1), 79–95. <https://doi.org/10.1080/17482798.2012.739783>
- Steinberg, L., Icenogle, G., Shulman, E. P., Breiner, K., Chein, J., Bacchini, D., Chang, L., Chaudhary, N., Giunta, L. D., Dodge, K. A., Fanti, K. A., Lansford, J. E., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Tirado, L. M. U., ... Takash, H. M. S. (2018). Around the world, adolescence is a time of heightened sensation seeking and immature self-regulation. *Developmental Science*, 21(2), e12532. <https://doi.org/10.1111/desc.12532>
- Tercova, N., & Smahel, D. (2025). Digital Skills’ Role in Intended and Unintended Exposure to Harmful Online Content Among European Adolescents. *Media and Communication*, 13, 8963. <https://doi.org/10.17645/mac.8963>
- Zuckerman, M. (1994). *Behavioral Expressions and Biosocial Bases of Sensation Seeking*. Cambridge University Press.

12.4 Shooting the Messenger: Audience Responses to Travel Vloggers.

Gracey Nana Amoakoa-Amoako (Indiana University Bloomington)

Introduction

Social media has transformed patterns of information creation and organization (Hays et al., 2013), influencing how users interact with and respond to content (McGregor, 2018). The advent of social media influencers and their communities of followers (Riedl et al., 2023) has led to the emergence of travel influencers, often referred to as travel vloggers, who visit different countries and create travel vlogs to share their experiences, the lifestyle, and the happenings of the countries they visit with online audiences. Some travel vloggers focus on political happenings, discussing issues such as conflict, governance, or security situations they encounter in the countries they visit, while others focus on cultural aspects, including religion, belief systems, food, and everyday social life. This has created a distinction between political travel vloggers and cultural travel vloggers, with audience engaging with these travel vlogs for different reasons depending on the type of experiences being presented. The audience responds to these vlogs in different ways, offering support and encouragement (prosocial) (Bao et al., 2021), or engaging in trolling, or antagonistic behaviors (antisocial) (Hardaker, 2010). Existing research on travel vlogs has primarily focused on how vloggers curate their content, achieve virality, or influence destination marketing and tourism promotion, leaving broader audience interaction underexplored.

To fill this gap, I adopted the notion of the mutual shaping process (Fichman et al., 2015), to examine how technologies such as social media platforms (sociotechnical features), and the social contexts of countries that travel vloggers present online, with users' responses to these presentations on YouTube are mutually constitutive. Comment sections on YouTube bring together a diverse range of audience with varying motivations to watch these travel vloggers hence, this study further employs social informatics perspective on collapsing context (Fichman & Amidu, 2024) to understand the nature of audience interactions and reactions with the different genres of travel vlog (political versus cultural). From this perspective, information and communication technologies acquire meaning and influence through how audience receive and react to these travel vlogs curated by travel vloggers, while those contexts are themselves reshaped through technological mediation.

Objectives/Research Questions

The objective of this study is to examine political and cultural travel vloggers on YouTube from a social informatics mutual shaping and collapsing context perspective.

The focus is on understanding the mutual shaping process with audience engagement, as well as the effects of collapsed contexts on political and cultural travel vlogs narratives online. YouTube's platform features, such as recommendation algorithms, visibility metrics, and interactive tools like likes, shares, and comments, shape how these videos circulate, while audience engagement simultaneously influences which travel vlogs gain prominence. Travel vlogs on YouTube are usually long-form videos which attract a geographically and socially diverse audience, with viewers following for different reasons, including entertainment, informational purposes: interest in the political or cultural aspects of a country. The convergence of multiple audience types within the same comment space illustrates a collapsed context, where diverse interpretations of the same travel vlog coexist and interact. This study examines how travel vlogs are shared online, platform features, and diverse audience engagement jointly shape the circulation, reception, and interpretation of information, particularly when these vlogs present information about a specific country or geopolitical region.

The study is guided by the following research questions:

1. What are the differences in audience prosocial/non-trolling and antisocial/trolling to YouTube political versus cultural travel vloggers?
2. How do audience prosocial and antisocial behaviors differ in semantic tone towards cultural versus political travel vloggers?

Methods

The study employed qualitative comparative content analysis to examine comments directed at cultural travel vloggers versus political travel vloggers on YouTube. Travel vloggers were identified through keyword searches including “travel vlog,” “travel experience,” “vlog,” and “visit.” Six travel vloggers (Table 1) were purposively selected from the search results based on whether their content focused on political or cultural themes, with three vloggers representing each category. Channel descriptions and overall content focus were used to verify the classification of vloggers as political or cultural travel creators. To construct the video sample, twenty videos were selected chronologically from each vlogger's channel, beginning with videos published in October 2025, which marked the date of data collection, and moving backward until the required number of videos was reached. This process resulted in a dataset of 120 videos. Comments associated with the selected videos were collected using YouTube Data API v3. Video URLs were compiled, and the corresponding video IDs were extracted and used in a Python script to retrieve publicly available comments for each video. For each video, 100 comments were collected. To reduce sampling bias and maintain a manageable dataset, Excel's randomization function was used to randomly select 30 comments per video, producing a final dataset of 3,600 comments.

Data analysis was conducted in two stages corresponding to the two research questions. First, a thematic analysis was conducted using a codebook derived from prior scholarship, which included prosocial (Bao et al., 2021), antisocial (Knustad & Johansson, 2021), and an additional category for comments that did not fall under either prosocial or antisocial behavior. Comments were coded according to these to answer the first research question. Second, drawing on the results of the thematic analysis, a stratified random sample of 100 comments was selected from each resulting category, namely cultural prosocial, cultural antisocial, political prosocial, and political antisocial. This produced a subset of 400 comments that were used for preliminary semantic analysis. Codes from previous studies were applied, including Semantic Prosocial Codes (Supportive Encouragement and Community Building) (Mukkamala et al., 2023) and Semantic Antisocial Codes (Hostility and Ideological Attack) (Potts et al., 2025). The semantic analysis goes beyond simple categorization to identify how meanings, sentiments, and interactional cues are expressed in prosocial and antisocial comments. The preliminary analysis provides an initial exploration prior to applying semantic coding to the full dataset. To ensure coding consistency, 10% of the comments were coded by a second coder for both the thematic and semantic analyses. This resulted in an interrater agreement of 85% for the thematic analysis and 88% for the semantic analysis.

Results

The thematic analysis (Figure 1) revealed prosocial scores were almost similar towards cultural versus political travel vloggers, however political travel vloggers received slightly high prosocial comments (26.2%) than cultural travel vloggers (25.3%). Antisocial responses were more frequent toward political travel vloggers (28.2%) compared to cultural travel vloggers (14.4%). The semantic analysis (Figure 2) showed that supportive and encouraging comments dominated prosocial engagement for both cultural and political travel vlogs, comprising most prosocial comments (Cultural: 87%, Political: 82%). Community-building frames were rare, suggesting that positive engagement mainly reflects immediate affirmation of the travel vlogger. Antisocial frames varied: cultural antisocial comments were more likely to express generalized hostility, whereas political antisocial comments were primarily ideological attacks (55%). A chi-square test of independence yielded a statistically significant result ($\chi^2 = 22.41$, $df = 3$, $p < 0.001$), indicating that the result is not random but significant with distribution of semantic tone and behavioral responses towards travel vloggers. The high antisocial comments directed at political travel vloggers, compared to cultural travel vloggers, are consistent with literature that political content online is associated with hostility and polarization in user interactions (Rathje et al., 2021). Political topics tend to activate group-based identities, which are linked to more antagonistic engagement patterns. In contrast, comments on cultural travel vloggers reflect lower levels of antisocial behavior, suggesting differences in how audiences respond to

political versus non-political travel content. The predominance of prosocial behavior across both categories aligns with influencer communication research indicating that audience engagement commonly takes the form of supportive and affirming expressions (Hudders et al., 2021). Overall, the findings indicate that content orientation and platform affordances jointly shape patterns of prosocial and antisocial behavior, illustrating mutual shaping, whereby the political or cultural orientation of the content interacts with YouTube's affordances to structure engagement. The convergence of diverse audiences within a single comment space produces divergent forms of engagement, reflecting elements of collapsed context.

Future Work

In future work within this project, it would be valuable to examine how patterns of prosocial and antisocial responses toward travel vloggers manifest across other social media platforms such as TikTok or Instagram, as each platform's sociotechnical features shape audience engagement in different ways. However, this study contributes to understanding how different types of travel vlogger content inspire distinct patterns of audience responses, revealing how thematic framing, whether cultural or political, shapes the emergence of prosocial or antisocial behaviors and the tone of interactions within digital environments.

References

- Bao, J., Wu, J., Zhang, Y., Chandrasekharan, E., & Jurgens, D. (2021). Conversations Gone Alright: Quantifying and Predicting Prosocial Outcomes in Online Conversations. *Proceedings of the Web Conference 2021*, 1134–1145. <https://doi.org/10.1145/3442381.3450122>
- Fichman, P., & Amidu, G. (2024). The roles of collapsing contexts and TikTok's features in reciprocal trolling. *Information Communication and Society*. <https://doi.org/10.1080/1369118X.2024.2391820>
- Fichman, P., Sanfilippo, M. R., & Rosenbaum, H. (2015). *Social Informatics Evolving*. Springer International Publishing.
- Hardaker, C. (2010). Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. *Journal of politeness research*, 6(2), 215-242.
- Hays, S., Page, S. J., & Buhalis, D. (2013). Social media as a destination marketing tool: its use by national tourism organisations. *Current Issues in Tourism*, 16(3), 211–239. <https://doi.org/10.1080/13683500.2012.662215>
- Knustad, M., & Johansson, C. (2021). Anonymity and Inhibition in Newspaper Comments. *Information*, 12(3), 106. <https://doi.org/10.3390/info12030106>

- McGregor, S. C. (2018). Personalization, social media, and voting: Effects of candidate self-personalization on vote intention. *New Media & Society*, 20(3), 1139–1160.
<https://doi.org/10.1177/1461444816686103>
- Mukkamala, R., Kauffman, R. J., & Henriksen, H. Z. (2023). Champions for social good: How can we discover social sentiment and attitude-driven patterns in prosocial communication?. *Journal of the Association for Information Systems*, 24(6), 1562-1593.
- Potts, R., Riddle, J., Hollander, J. B., & Hartt, M. (2025). The online conversion framework: Understanding antagonism, planning theory, and social media. *Planning Theory*, 24(2), 115-140.
- Riedl, M. J., Lukito, J., & Woolley, S. C. (2023). Political influencers on social media: An introduction. *Social Media+ Society*, 9(2), 20563051231177938

12.5 The impact of Verbal, Non-Verbal, and Vocal Cues in Political Posts on TikTok on Users' Verbal Responses

Gracey Nana Amoakoa-Amoako (Indiana University Bloomington)

Introduction

Political communication entails the strategies, techniques and diverse methods of communication employed by political actors to convey their messages to influence public opinions and behaviors (Bennett & Iyengar, 2008; Serapşah & Sevgihan, 2023). Political communication was dominated by conventional news media, but the proliferation of social media, especially on contemporary platforms like TikTok, shifted the communicative power toward political influencers (Riedl et al., 2023), who communicate to their often-large followers directly on the social media platforms (Perloff, 2021). TikTok influencers utilize platform features to deploy verbal, non-verbal and vocal communication cues, transcending beyond the limits of verbal communication online. Existing studies on online political communication were confined to textual communication, on platforms such as X (formerly Twitter), and scholars have mainly analyzed the flow of verbal political communication on digital platforms (Lipschultz, 2023). However, TikTok, which is a short-video form platform, enables the construction of layered communication that extends beyond only verbal delivery.

Mehrabian's (1971) communication model posits that only 7% of communication is conveyed through verbal cues, with 38% vocal, and 55% non-verbal cues. This model originally was developed to examine the impact of verbal, vocal, and non-verbal cues in face-to-face emotional interactions of face-to-face interpersonal communication. While Mehrabian's model is not literally interpreted for online content, the model provides a framework to examine how meaning is conveyed through three distinct modes of communication: verbal, vocal, and non-verbal cues. In this study the Mehrabian model is applied here conceptually to highlight the multimodal nature of online political communication on TikTok. TikTok's platform features, including video editing tools, captions, background audio, filters, and visual overlays, allow creators to integrate these cues in ways that extend beyond spoken words alone. These features enable political influencers to emphasize verbal arguments, reinforce messages through gestures, facial expressions, and symbolic visuals, and create additional meaning using tone, intonation, or background sounds, reflecting the multimodal components highlighted by Mehrabian's model.

While prior research has examined non-verbal communication, such as gestures and facial expressions, in online communication, few studies have analyzed how the three

cues: verbal, non-verbal, and vocal cues interact together in digital environments. Grounded in Mehrabian's conceptual emphasis on the interplay of verbal, vocal, and non-verbal cues, this study adopts a social informatics approach to investigate the sociotechnical features of TikTok that enable political influencers to integrate verbal, non-verbal, and vocal cues in their political messaging and to explore how these multimodal cues shape audience response.

Objectives/Research Questions

Political influencers on social media utilize sociotechnical affordances of social media platforms like to create multimodal political communication for their political agenda. TikTok for instance has unique affordances such as short form video, editing tools, captions, background audio, stitch, and overlays that allow users to create political content characterized by the interplay of multiple communication techniques. These platform features enable creators to blend the three communication cues emphasized in Mehrabian's communication model: verbal cues such as spoken words and captions, non-verbal cues such as gestures, facial expressions, and visual symbols, and vocal cues such as tone, pitch, and intonation. This study examines how TikTok political influencers affiliated with the Democratic and Republican U.S. political parties used TikTok during the lead up to the 2024 U.S. presidential election to share political messages with multimodal communication. The study seeks to examine the use of multimodal communication by political influencers on TikTok and to understand how audiences engage with and respond to verbal, non-verbal, and vocal cues in the comment sections.

The study is guided by the following research questions:

1. How do TikTok Democrat and Republican political influencers use verbal, non-verbal and vocal cues in their posts on TikTok?
2. How do TikTok users react to the verbal, non-verbal and vocal cues in TikTok video posts by political influencers?

Methods

In January 2025, we sampled 100 video posts from TikTok using the purposive sampling method. Using keywords such as "MAGA," "Harris 2024," "Trump 2024," "I am voting for Kamala," "I am voting for Trump," and "US elections," we identified 50 prominent political influencers split into 25 Democrats and 25 Republican supporters who were active on TikTok in the month leading to the 2024 US presidential election. We then selected their top 2 videos with the largest counts of viewers and likes and downloaded the videos into NVivo15, a software for qualitative analysis (Mortelmans, 2019). Using ExportComments.com in April 2025, we extracted the comments associated with these

videos, and with three exceptions, we randomly sampled 20 comments per video, for a total of 1,939 comments.

We conducted thematic analysis on both videos and comments using an inductively developed structured codebook. For the non-verbal category, we included gestures, facial expressions, eye movement, posture, and symbolic visuals, drawing on research showing these behaviors as examples of non-verbal communication (Dumitrescu, 2016). For the verbal category, we focused on spoken political topics consistent with research on integrating verbal and non-verbal messages (Jones & LeBaron, 2006). For the vocal category, we considered tone and intonation as part of paralanguage that modulates meaning beyond words (Knapp et al., 2013). The codebook was pilot tested on a subset of videos and comments to refine definitions and ensure clarity.

For the videos, 14 codes were applied across communicative cue categories. Verbal codes examined themes such as Economic Policies, Election & Politicians, Foreign Policy & National Security, and Rights. Non-verbal codes analyzed Amateurs, Professional, Symbols-Flags, TikTok Features, Eye Movement, Facial Expression, Hand Gesture, and Posture. Vocal codes focused on Background Audio and Intonation (Highfield & Leaver, 2016). Videos were systematically watched, and timestamps were used to code specific moments that aligned with the coding categories. To analyze audience responses, comments were coded using a codebook with three categories: Verbal, Non-Verbal, and Vocal communication cues. Verbal codes captured perceptions of trolling (e.g., provocative, rhetorical, reverse) and constructive engagement (e.g., requests, dialogue, recognition, repetition, tagging). Non-verbal codes identified audience references to the creator's body language, symbolic visuals like flags or political paraphernalia, and audiovisual overlays. Vocal code included mentions of tone, accent, and platform audio features such as background sounds or sound props. Each comment was treated as a unit of analysis. To ensure coding consistency, a second coder independently coded 20% of videos and comments, achieving an inter-rater reliability of 85% simple agreement. After coding all 100 posts and 1,939 comments, we aim to perform a code co-occurrence on NVivo's matrix query.

Results

We found that all influencers utilized the platform features, and both Republican and Democratic influencers relied more heavily on non-verbal cues, such as visual symbols and body language, than on vocal elements, with Republicans tending toward more professional video setups and Democrats favoring amateur settings. Verbal content largely focused on election-related topics and candidate representation, consistent with previous literature on political communication during election cycles (Boulianne et al., 2020). Analysis of comment reactions reveals that although influencers use by far more non-verbal cues in their videos compare to non-verbal or vocal cues, audience

responses predominantly reacted to the verbal cues, followed by reaction to vocal cues with little reactions to the non-verbal cues in the videos

Future Work

While this study focused on verbal, non-verbal, and vocal cues in video content and user comments during a U.S. electoral period, future work could explore longitudinal patterns across multiple election cycles or compare engagement across different social media platforms, such as Instagram or X, to understand whether the prominence of verbal cues in audience responses persists. Additionally, experimental or mixed-method designs could investigate causal relationships between specific combinations of verbal, non-verbal, and vocal cues and audience perception, engagement, or persuasion. Finally, future research could examine the role of multimodal cues in more diverse political contexts, including local or international political campaigns, to assess how cultural, linguistic, or platform-specific affordances shape the production and reception of political messages online.

References

- Bennett, W. L., & Iyengar, S. (2008). A New Era of Minimal Effects? The Changing Foundations of Political Communication. *Journal of Communication*, 58(4), 707–731. <https://doi.org/10.1111/j.1460-2466.2008.00410.x>
- Boulianne, S., Koc-Michalska, K., & Bimber, B. (2020). Mobilizing media: comparing TV and social media effects on protest mobilization. *Information, Communication & Society*, 23(5), 642–664. <https://doi.org/10.1080/1369118X.2020.1713847>
- Dumitrescu, D. (2016). Nonverbal communication in politics: A review of research developments, 2005–2015. *American Behavioral Scientist*, 60(14), 1656–1675. <https://doi.org/10.1177/0002764216678280>
- Highfield, T., & Leaver, T. (2016). Instagrammatics and digital methods: studying visual social media, from selfies and GIFs to memes and emoji. *Communication Research and Practice*, 2(1), 47–62. <https://doi.org/10.1080/22041451.2016.1155332>
- Jones, S. E., & LeBaron, C. D. (2006). Research on the relationship between verbal and nonverbal communication: Emerging integrations. *Journal of Communication*, 52(3), 499–521
- Kjeldsen, J. E., & Gelang, M. (2024). Nonverbal communication as argumentation: the case of political television debates. *Argumentation and Advocacy*, 60(1), 18–37. <https://doi.org/10.1080/10511431.2023.2294236>
- Knapp, M. L., Hall, J. A., & Horgan, T. G. (1972). Nonverbal communication in human interaction (p. 512). Thomson Wadsworth.

Lipschultz, J. H. (2023). Social Media Communication. Routledge.

<https://doi.org/10.4324/9781003281924>

Mehrabian, A. (1971). Silent messages: Vol. Vol. 8. Belmont, CA: Wadsworth.

Mortelmans, D. (2019). Analyzing Qualitative Data Using NVivo. In The Palgrave Handbook of Methods for Media Policy Research (pp. 435–450). Springer International Publishing. https://doi.org/10.1007/978-3-030-16065-4_25

Perloff, R. M. (2021). The Dynamics of Political Communication. Routledge.

<https://doi.org/10.4324/9780429298851>

Riedl, M. J., Lukito, J., & Woolley, S. C. (2023). Political influencers on social media: An introduction. Social Media+ Society, 9(2), 20563051231177938

Serapşah, G. H., & Sevgihan, M. E. (2023). The Influence of Media on Political Communication: A Review of Literature. American Journal of Law and Political Science, 2, 42–54. <https://doi.org/10.58425/ajlps.v2i3.214>

12.6 Everyday Digilantism: Platformed Neighbourhoods and the Informal Governance of Urban Life

Sophie Anson (Birkbeck, University of London)

Introduction

Digital vigilantism (or digilantism) is broadly defined as the organisation and enactment of vigilante practices in digital spaces (Trottier 2017). Existing scholarship on digital vigilantism has focused predominantly on extreme and episodic acts of violence, as well as highly visible digital shaming campaigns such as those conducted by ‘Paedophile Hunters’ (Purhouse 2020, Hadjimatheou 2021, Tippet 2022). While analytically powerful, this focus on extreme cases risks obscuring the more mundane, routinised practices of surveillance and shaming that permeate everyday digital life.

This project shifts the empirical focus to the ordinary and ongoing ways local residents “take law into their own hands” through neighbourhood-based social media. It examines digilantism through the case study of a local Facebook group organised around the issue of bike theft in a UK city. Common posts found within the group include images of abandoned bikes, photos of individuals deemed to be ‘suspicious’, and calls for members to identify the names and addresses of those accused of theft. At its peak in May 2023, the group amassed a membership of over twelve thousand Facebook users, a vast network described by the group’s administrator as ‘eyes and ears across the city’.

While the group presents itself as a practical forum for information sharing and crime prevention, I argue it also functions as a digitally mediated site of informal legal ordering. Through repeated acts of monitoring, reporting, and public exposure, members collectively negotiate who belongs, who poses threat, and what constitutes responsible citizenship. Rather than treating digilantism as a moment of breakdown or digital spectacle, this project conceptualises it as an everyday mechanism of distributed governance, emerging from the intersection of neighbourhood life and social media infrastructures.

Objectives/Research Questions

- How is digilantism enacted in everyday neighbourhood social media spaces?
- How do participants justify and legitimise practices of monitoring, exposure and public shaming?
- How do platform affordances shape the visibility, circulation and moderation of digilante practices?

By foregrounding mundane digital surveillance rather than spectacular violence, the study contributes to debates about platformed publics, participatory surveillance, and the shifting boundaries between state authority and citizen enforcement in digitally mediated urban life

Methods

This study draws on digital ethnography conducted between December 2022 and June 2023, combining:

- Digital content analysis of posts and comment threads from the Facebook group, including retrospective collection dating back to the group's establishment in 2019 and real-time observation between 2022–2023.
- Semi-structured interviews with twenty-two group members and administrators, exploring motivations, perceptions of policing, and interpretations of civic duty.

Real-time collection was necessary as moderators frequently deleted posts that raised legal, ethical, or proportionality concerns. Observations and interviews together allowed for analysis of both practices of surveillance and the moral judgement and the meanings participants attach to them.

Results

Findings revealed that digilantism in this context functions as a platform-mediated, spatially embedded form of informal governance, structured around two interrelated processes:

- The construction of a “criminal other”: Surveillance is directed toward highly visible and socially marginalised populations, including teenagers, the Gypsy, Roma, and Traveller (GRT) community, and people experiencing homelessness. Visibility becomes a proxy for suspicion: those who are present in public or recognisable within social hierarchies are routinely targeted. Posts often relied on aesthetic or behavioural markers (clothing, face coverings, perceived demeanor) rather than evidence of wrongdoing, producing a “visual grammar of criminality.” The group's ‘Rogues Gallery’ exemplifies this, listing individuals identified through crowd-sourced information regardless of conviction, and justifying public exposure. Criticism of these practices was systematically suppressed. Members who questioned racialised or age-based targeting were removed or forced to apologise, consolidating group cohesion around an “us versus them” moral framing.
- Producing the “responsible citizen”: In contrast to the criminal other, group members cultivate a sense of responsible citizenship through vigilance, information-sharing, and moral engagement. Participation conveys alignment with community norms and positions members as moral actors compensating for perceived deficiencies in formal policing. Administrators are celebrated as exemplary citizens, modelling civic virtue

and reinforcing hierarchies of respectability grounded in class, property ownership, and civic participation.

This dual construction produces a disciplinary effect: exposure to shaming content reshapes participants' own behaviour, fostering internalised surveillance, risk-avoidance, and compliance with group-defined norms. Digital platforms intensify these effects by making information persistent, shareable, and widely visible, embedding moral and legal judgement into everyday urban navigation.

Ultimately, the project found that digilantism can be understood not only as a form of digital spectacle, but as a constitutive mechanism of local spatial governance. By embedding surveillance, suspicion, and moral judgment into the everyday practice of digital neighbourhood life, digilantism plays a significant role in shaping how law, order, and belonging are sensed and negotiated. For scholars of social media and society, attending to these mundane forms of informal governance is essential to understanding how digital practices shift the boundaries of legality, citizenship, and belonging in contemporary urban life.

Future Work

My PhD research will build on this project, and its focus on the 'everyday', by examining neighbourhood and location-based social media (e.g. street WhatsApp and Facebook groups, Nextdoor) more broadly as a site of local boundary-making and surveillance. In addition to this broader empirical context, the PhD will extend this research in two key ways:

- A stronger focus on processes of racialisation: A prominent theme emerging from this study was the role of digital neighbourhood spaces in facilitating the othering and racialised governance of Gypsy, Roma, and Traveller (GRT) communities. Drawing on phenomenological approaches to race, the PhD will examine how racialised bodies become hyper-visible as presumed threats, and how entitlement to legal authority and protection is unevenly distributed along racial lines.
- Applying insights from digital sociology and platform governance: The PhD will also engage more deeply with digital sociology and platform governance to examine how affordances (such as anonymity, algorithmic visibility, and post moderation) enable, amplify or constrain digilante practices. This will situate everyday digilantism within broader debates about distributed governance and platform responsibility.

By analysing neighbourhood social media as a site where legality, surveillance, and belonging are co-produced, this research contributes to ongoing discussions about informal governance, digital citizenship, and the transformation of urban social order in platform societies.

References

- Hadjimatheou, K. (2021). Citizen-led digital policing and democratic norms: The case of self-styled paedophile hunters. *Criminology & Criminal Justice*, 21(4), 447-565. <https://doi.org/10.1177/1748895819880956>
- Purshouse, J. (2020). 'Paedophile hunters', criminal procedure and fundamental human rights. *Journal of Law and Society*, 47(3), 384-411. <https://doi.org/10.1111/jols.12235>
- Tippett, A. (2022). The rise of paedophile hunters: To what extent are cyber-vigilante groups a productive form of policing, retribution and justice? *Criminology and Criminal Justice*, 24(4): 711-732. <https://doi.org/10.1177/174889582211368>
- Trottier, D. (2017). Digital vigilantism as weaponisation of visibility. *Philosophy & Technology*, 30(1), 55-72. <https://doi.org/10.1007/s13347-016-0216-4>

12.7 “They Can Say It, We Get Banned”: Folk Theories of Platform Governance in Conservative Reddit Communities

Morgan Badurak (Louisiana State University)

Introduction

Social media platforms have become central infrastructure for political communication, functioning as primary spaces through which users encounter news, express political identity, and engage in public debate. As a result, decisions about content moderation, namely what speech is removed, restricted, or amplified, have become contested political issues. Scholars have increasingly documented how platforms govern speech through formal policies, discretionary enforcement, and algorithmic systems that shape visibility and participation (Gillespie, 2018; Gorwa, 2019). Despite the growing attention to governance as policy and power, less research has examined how platform moderation is experienced by users, interpreted through political identities. These dynamics are particularly visible on Reddit, where platform governance operates through a hybrid system of centralized platform rules and decentralized volunteer moderators who exercise significant authority over individual communities. This structure creates multiple layers of governance that are often opaque to users and can produce conflicts over how rules are interpreted, enforced, and justified.

This gap in the literature is particularly consequential in the contemporary U.S. political context. Since the 2016 election and especially following the deplatforming of President Donald Trump and other high-profile conservative figures on social networking sites like X (Twitter), platform moderation has been reframed by many conservative users as evidence of ideological censorship rather than a neutral enforcement tool. Claims that platforms are silencing conservatives have become prevalent in right-wing media ecosystems. These disputes reflect a broader transformation in how free speech is understood in digital environments, where private companies now govern much of the public sphere while simultaneously denying their role as political actors (Gillespie, 2018).

In this way, platform governance should be understood not only as a technical or institutional process, but as a form of experienced political authority. Users do not primarily encounter governance through policy documents or transparency reports. Instead, they experience it through moments of moderation friction in their own user experience through content removal, alleged shadow bans, and enforcement actions that appear opaque, inconsistent, and seemingly politically motivated. These encounters generate folk theories of governance, or everyday explanations that users

develop to make sense of moderation (Myers West, 2018; Riedl et al., 2024). These folk theories are especially salient in highly polarized environments where moderation decisions are interpreted through a partisan lens and embedded within existing narratives of institutional distrust and elite manipulation (Myers West, 2018; Schulze et al., 2022).

Objectives/Research Questions

The objective of this study is to identify and analyze the folk theories of platform governance that users construct when explaining experiences with moderation, censorship, and bans on Reddit. This study examines discussions from conservative-oriented Reddit communities and broader platform conversations to identify the explanatory narratives users employ when describing bans and perceived ideological bias. Understanding these interpretive frameworks provides insight into how users perceive the legitimacy and fairness of platform governance systems. Based on this central issue and the proposed theoretical framework, I pose the following research question:

RQ1: What explanations do Reddit users provide when describing or interpreting their experiences with bans or moderation decisions?

RQ2: What recurring folk theories of platform governance emerge in these discussions?

RQ3: What actors or mechanisms do users attribute responsibility to when explaining moderation outcomes?

Methods

Data Collection

To identify the explanations users provide for these experiences, I collected posts and comments from Reddit discussions where moderation decisions were being explicitly debated or contested. Posts and comment threads were collected from conservative-oriented Reddit communities. Within these communities, I searched for posts containing keywords associated with moderation experiences. To consolidate the corpus, I only included threads with more than 50 comments and that were posted within the last year. The selected posts and comments were collected through using a Python-based scraping script that accessed Reddit threads through publicly available JSON endpoints. The script retrieved original posts, comments, and nested replies, while recording information including the text of each comment, the username of the author, comment score (upvotes), number of replies, and thread URL. The initial dataset contained 9,140 Reddit comments. After preprocessing, 8,692 comments were included in the final analysis.

Data Analysis

The analysis used a mixed-methods approach combining qualitative content analysis with computational text analysis to sort top-level codes. The goal of this approach was to identify both the major themes discussed in moderation-related conversations and the explanations users employ when interpreting moderation decisions. The first stage of analysis involved inductive qualitative content analysis of Reddit posts and comments. During data collection, I read through every thread and notated themes that occurred across the content. Rather than focusing only on the presence of moderation events themselves, the analysis focused on the ways users attempted to explain why those events occurred. The categories identified during qualitative coding were then used to guide computational analysis of the dataset. Because of the dataset size, I applied multiple topic modeling approaches to identify recurring discourse patterns at scale to compare and consolidate top-level themes. Three topic modeling approaches were used: BERTopic, Latent Dirichlet Allocation (LDA), and Non-negative Matrix Factorization (NMF). Using these models in combination allowed for triangulation across methods and increased confidence that the themes identified in the analysis reflected consistent discourse patterns rather than artifacts of a particular modeling approach.

Results

Across the dataset, seven top-level explanatory themes consistently appeared. Each is explained below.

(1) Ideological Censorship and Free Speech Suppression

One of the most common explanations for moderation actions framed bans and comment removals as ideological censorship targeting conservative viewpoints. Users frequently interpreted moderation as a deliberate suppression of political speech that diverges from the perceived ideological orientation of the platform.

(2) Echo Chamber Enforcement and Political Brigading

Many users believed that moderation decisions are used to maintain ideological echo chambers. In this interpretation, bans and comment removals are seen as tools for maintaining political homogeneity within communities by excluding dissenting viewpoints. Some users also attributed moderation actions to political brigading, arguing that users coordinate to report or downvote posts they disagree with to trigger moderation responses.

(3) Moderator Power and Ideological Gatekeeping

Another recurring explanation focused on the role of moderators as powerful actors who shape discourse within online communities. Moderators are portrayed as ideological gatekeepers who selectively enforce rules in ways that align with their political preferences.

(4) Algorithmic Censorship

Some users interpreted moderation outcomes as the result of automated moderation systems or algorithmic filtering tools rather than direct human intervention. In these explanations, automated moderation infrastructure is believed to suppress certain, mostly conservative, viewpoints.

(5) Unequal Enforcement and Double Standards

Another recurring explanation emphasized perceived inconsistencies in how moderation rules are applied. Users frequently argued that platform rules are enforced unevenly depending on political affiliation, with conservatives punished for behavior that other users are allowed to engage in without consequences.

(6) Coordinated Narrative Control

Some discussions framed moderation as part of a broader effort to control political narratives online. In this explanation, moderation decisions are interpreted as coordinated attempts to shape public discourse by limiting the visibility of certain viewpoints.

(7) Shadowbanning and Visibility Suppression.

Finally, some users described moderation through the concept of shadowbanning, where posts are not removed but are allegedly hidden from other users. In this explanation, platforms suppress viewpoints by reducing their visibility rather than explicitly banning them. These narratives reflect broader concerns about opaque platform governance systems.

Free speech censorship is the most prominent explanation (mean ≈ 0.298 ; Table 2), followed by echo chamber enforcement (≈ 0.288) and moderator bias (≈ 0.285). Technological explanations appear less frequently (≈ 0.256), coordinated narrative control and shadowbanning remain present but less common.

Future Work

Across these findings, users frequently interpreted moderation decisions through broader narratives about political power, technological control, and institutional bias. These narratives function as folk theories of platform governance, allowing users to construct explanations for moderation systems that are largely opaque. Through these explanations, users attempt to make sense of the distribution of authority among moderators, algorithms, and platforms, as well as the perceived ideological dynamics shaping it. Future research should further investigate how these folk theories evolve across different platforms and political contexts, particularly as policies, algorithmic systems, and governance structures change over time. Comparative work examining how users interpret governance across platforms such as Reddit, X, TikTok, or

alternative platforms like Gab or Truth Social could provide insight into how platform design and governance shape these interpretations.

Governance is accepted or rejected based on perceived fairness by users, not through understanding policy texts or transparency statements created by platform companies. Future studies should examine the relationship between governance perceptions and behavioral outcomes more directly, including how perceived censorship or bias influences platform migration, political engagement, and trust in digital institutions. For practitioners, conversations about how to define political influencers and political content, as well as determining the appropriate level of platform intervention, are already taking place among civil society groups and policymakers. This work, therefore, contributes to ongoing debates about moderation, user expectations, and the development of appeals and reconciliation mechanisms. Research on what users' experience is crucial for creating more meaningful policy conversations and institutional change. Additionally, many users in these discussions cite a lack of transparency and limited pathways to contest moderation decisions. Future work should examine how increased transparency, clearer moderation explanations, or improved appeals processes may affect user perceptions of fairness and legitimacy. Developing governance systems that users perceive as fair and understandable may help reduce distrust in platform institutions.

Finally, these findings suggest that improving communication between platforms and users may have broader consequences for online political discourse. If users have clearer mechanisms to understand or contest moderation decisions on mainstream platforms, they may be less likely to migrate to fringe spaces that promise fewer restrictions but often host more extreme or radicalizing content. Future research should therefore continue to develop these concepts by examining how governance transparency, appeals systems, and user feedback mechanisms shape both perceptions of legitimacy and broader patterns of political communication across digital platforms.

References

- Gillespie, T. (2018). *Custodians of the Internet*. Yale University Press.
- Gorwa, R. (2019). What is platform governance? *Information, Communication & Society*, 22(6), 854–871. <https://doi.org/10.1080/1369118X.2019.1573914>
- Myers West, S. (2018). Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. *New Media & Society*, 20(11), 4366-4383. <https://doi.org/10.1177/1461444818773059>

Riedl, M. J., Martin, Z. C., & Woolley, S. C. (2024). 'I get suppressed:' pro-and anti-abortion activists' folk theories of platform governance and shadowbanning. *Information, Communication & Society*, 27(11), 2153-2170

Schulze, H., Hohner, J., Greipl, S., Girgnhuber, M., Desta, I., & Rieger, D. (2022). Far-right conspiracy groups on fringe platforms: a longitudinal analysis of radicalization dynamics on Telegram. *Convergence: The International Journal of Research into New Media Technologies*, 28(4), 1103-1126.
<https://doi.org/10.1177/13548565221104977>

12.8 Building a National Archive of State Social Media

Jake Bickford, Claire Newing (The National Archives, UK)

Introduction

The National Archives (UK) is the official archive and publisher for the UK Government. The Public Records Act (1958) empowers us to collect and preserve the public record in any format or medium, and make it as widely available as possible. As government embraced online spaces from the 1990s, The National Archives recognised the pressing need to preserve this new medium of government communication and citizen-state interaction. Starting in 2004 (though with backfilled content dating from 1996), the UK Government Web Archive (UKGWA) began collecting government websites for permanent preservation via remote web crawling, and today holds some nine billion digital objects in a freely available online archive.

The emergence of social media posed new challenges to the completeness of this online public record, from both a technical perspective (since social media platforms are considerably more complex than traditional websites) and in terms of the policy and ethical challenges presented by the medium (as they contain material that forms part of the public record published on platforms that are outside government ownership, and that is frequently presented alongside non-government content). It was also recognised that content published on social media was even more fragile than other forms of digital record, being by its nature ephemeral and subject to deletion or rapid changes to the host platform. This poster will tell the story of The National Archives and our partners' innovative work to preserve this vital but vulnerable component of the public record, from our initial experiments to the development of production processes. It will highlight our recent efforts to navigate the increasingly restrictive environment for the preservation of social media, and explore how researchers can make use of our growing archive.

Objectives/Research Questions

The objective of this work has been to create a reliable, accurate archive of government social media content for permanent preservation, and to provide access to it to our users on the same terms as we would with other records (for example archived websites, or digital or paper documents). The key questions that this work addresses are:

- What technical processes are required for capturing social media, and to what extent can they follow established web archiving and digital preservation practices?

- What is an appropriate scope for a government social media archive, and how should we navigate the ethical and legal implications of archiving this medium?
- How should archived social media be preserved and how should we provide access to the collection?
- How can archives respond to an increasingly complex and restrictive social media landscape?

Methods

The UKGWA began systematically capturing social media in 2014. Initial experiments focused on archiving Twitter using the same remote crawling system as for websites. This proved technically complex given the highly interactive and dynamic nature of social media sites, while the capture method made separating out in-scope government content from other material very challenging. It was also felt that this type of capture inevitably privileged one form of social media presentation (i.e. on the platform website as it appeared at the time of capture) as opposed to reflecting the diversity of forms of social media consumption then current.

It was therefore decided to focus on only capturing content itself (for example, the text of a tweet, along with any associated media and metadata) rather than the look and feel of specific platforms. This mitigated the risk of capturing content that was outside our legal remit and opened alternative approaches to capture via platform APIs. It also maintained our existing model of archiving public content without requiring engagement from account holders. Working with our then partners, Internet Memory Foundation, we began regular harvesting of Twitter and YouTube, making Twitter content available as archived JSON and XML feeds as well as text. We also developed a process for resolving and crawling embedded links in tweets.

From 2017, working with a new supplier, MirrorWeb, we increased capture frequency to a daily check of the APIs, increased the number of channels captured, and began providing access to Tweets via a new playback platform, removing the JSON and XML feeds. In 2019, in part prompted by Flickr's announcement that it would limit free account holders to 1000 images, we began capturing Flickr content, backfilling over a decade's worth of images from government accounts. At this time, in response to user research, we also introduced what we believe to be the first full-text search across any social media archive.

This approach served us well for YouTube, Twitter and Flickr, but archiving other platforms proved more problematic. We were, however, able to implement an in-house capture process for Instagram using Instaloader. This precedent proved helpful when changes to Twitter's API rules made our model impossible to sustain from mid-2023. In the run-up to the 2024 General Election, we began in-house capture of selected Twitter

and Facebook accounts using similar client-side tools. We have also re-assessed our preference for a remote capture approach and concluded that for high-priority accounts (for example Public Inquiries) we need to engage with account holders and ask them to use the platform's data portability features to extract social media content for deposit. We have piloted this approach with Twitter and have begun receiving our first accessions from account holders. Content captured in-house or submitted by account holders is currently in storage while we work to add it to the public archive.

Results

The results of the UKGWA's social media archiving programme reflect on the one hand the sustained efforts of The National Archives and the wider web archiving community to collect and preserve these valuable records, and on the other the complex and unstable landscape of social media. Following an initial period of experimentation, we were able to construct reliable and sustainable processes for some platforms but not for others. During this period, we accrued and published some 2.5 million Tweets, 126,000 YouTube videos, and 262,000 Flickr images. Our coverage of the government's response to significant national events such as Brexit and the COVID-19 pandemic forms an invaluable component of our digital public record. The ongoing collection, and permanent preservation, of this content in a formal archive both supports public accountability, and potentially adds a temporal dimension that may be absent from 'one off' collections of research data.

However, other platforms used by the government such as LinkedIn remain absent from our collection, and while we continue to accrue YouTube and Flickr content according to our established processes, our Twitter archive has become fragmentary and uneven following the API changes (down from around 850 accounts to around 100 regularly captured and a handful of direct deposits). Whilst this has prompted fruitful experimentation with new tools and approaches, it serves as reminder that digital preservation requires flexibility as well as ongoing commitment. Introducing the model of user-deposited content, and working to integrate this with the main archive, provides us with an alternative route that can potentially be adapted to other platforms as new social media services emerge, or as APIs become more restrictive, but it is difficult to see this matching the scale of previous API-driven archiving.

Future Work

We continue to work to adapt our operating model to the more challenging social media archiving environment that confronts us, and we will shortly be publishing our first user-deposited Twitter content to the main archive. We will then focus on integrating content that has been captured in-house, focusing on Twitter, Instagram and Facebook. We have carried out experimental captures of TikTok and BlueSky, and we will continue to monitor these platforms, with a view to eventually archiving them as available

resources permit. We are also working with our partners to better integrate social media content with the wider web archive and have recently implemented a script that links embedded YouTube videos in the web archive to the corresponding video in the social media archive, allowing users to click through and view the video. This remains work in progress, and more sophisticated forms of linkage between the collections are possible future developments. Finally, we believe that the social media archive is an underused research resource, and are keen to support researchers from all disciplines who would like to explore the official record of government social media.

12.9 Feeling Safe or Being Safe? Social Media Privacy Strategies Among Women of Color

Alexa Bolanos-Carpio (University of Costa Rica) Fanny Ramirez (Louisiana State University)

Introduction

Social media platforms have become integral to everyday life, providing spaces for social interaction, communication, and creative self-expression. But social media applications also come with privacy risks such as data tracking, unwanted visibility, identity theft, online harassment, and the long-term persistence of personal content (Ahmed & Rashid, 2024; Trepte, 2021). Users who wish to manage their online privacy have a wide range of tools and strategies at their disposal. Research on this matter found that users attempt to protect themselves from perceived threats through both technical and behavioral privacy strategies. Technical strategies include the use of tools such as ad blockers, encryption, or privacy settings, while behavioral strategies involve practices like self-censorship, audience segregation, or “whitewalling” (boyd, 2010; Choon, 2018). Studies show that behavioral tactics such as limiting self-disclosure or tailoring posts to imagined audiences are among the most frequently used, in part because many users lack confidence in their ability to navigate technical protections (Boerman et al., 2018). Privacy strategies can also be proactive, preventing disclosures in advance through selective sharing or adjusting settings, or reactive, such as deleting posts, untagging photos, or withdrawing from platforms (Dienlin & Trepte, 2015). These strategies often vary in effectiveness: while clearing cookies or browsing in private mode may reduce local traces, such tactics do little to mitigate institutional tracking by corporations (Boerman et al., 2018).

The process of managing one’s online privacy can be daunting and tiring. For example, research shows that seven-in-ten Americans are overwhelmed by the number of passwords they have to remember, and that about half are unsure whether their passwords are strong and secure enough (Faverio, 2023). Additionally, repeated data breaches have contributed to growing privacy fatigue, that is, a sense of hopelessness and futility about maintaining online privacy that discourages people from engaging in proactive privacy practices (Choi et al., 2018; Sarikakis & Winter, 2017; Trepte, 2021). This raises important questions about how individuals respond: What tools and strategies do people use to proactively protect their social media privacy? And are personal privacy mitigation approaches indicative of people’s broader online social media privacy behavior? This is where our study intervenes.

Objectives/Research Questions

Drawing on data from a mixed method survey of 390 women living in the U.S. who self-identity as either Black or Latinx, we examine personal privacy strategies on social media. We focus our study on women of color for several reasons. First, a recent study by the Pew Research Center found that women are more likely than men to use social media platforms, and that within racial and ethnic groups, Black and Latinx users engage with social media at much higher rates than their white counterparts (Gottfried, 2024). Second, because systems of power (racism, sexism, transphobia, classism, etc.) are reproduced online (Sannon & Forte, 2022), marginalized individuals often experience heightened risks related to digital privacy, including harassment, job loss, policing, or doxxing. As such, research on digital inequality suggests that marginalized individuals often have to do more labor around digital privacy, such as curating audiences or avoiding certain topics, to protect themselves from stigma, harassment, or discrimination (Marwick et al., 2017; Wang & Metzger, 2023). Altogether, this makes Black and Latinx women a particularly relevant population for a study on social media privacy strategies.

There are two parts to our study. First, through a thematic analysis of short-answer responses, we qualitatively examine how women of color describe managing their social media privacy. We ask (RQ1): What types of personal privacy strategies do women of color report using to manage their privacy on social media?

Next, to contextualize women's personal privacy strategies within their broader privacy behaviors, we quantitatively assess whether the use or non-use of various personal strategies shapes women's overall score on the social media privacy behavior scale (Sindermann et al., 2021) using a series of two-sample t-tests. We ask (RQ2): How is the use of different personal privacy strategies associated with participants' scores on the social media privacy behavior scale? We hypothesize that (H1), regardless of the specific strategy, participants who report using a privacy strategy will have higher scores on the social media privacy behavior scale than those who do not report using that strategy.

Methods

Data Collection

In 2022, we used the survey platform Prolific to recruit 415 women currently living in the U.S. who self-identified as Black/African American or Latina/Hispanic. Prolific is widely used in academic research and is recognized for ethical participant treatment and higher data quality than platforms such as MTurk (Peer et al., 2022). The survey included closed-ended questions (e.g., demographics, social media use, online privacy behavior) and an open-ended questions about participants' personal digital privacy strategies: "Thinking about your social media use, what steps do you take to protect

your privacy on social media? What privacy settings and personal strategies do you use? Please explain in a few sentences.” Participation was anonymous, the average completion time was 14 minutes (Min=4, Max=48), and participants received \$3.80.

Participants

After removing timeouts and incomplete responses, the final sample included 390 women. Of these, 48% identified as Latina/Hispanic and 52% as Black/African American. Ages ranged from 18 to 75 ($M=33.48$, $SD=11.61$). Political partisanship was measured on a 7-point scale ranging from strong Republican (1) to strong Democrat (7); overall, the sample leaned Democratic ($M=5.25$, $SD=1.49$).

Data Analysis

Content Analysis of Short Answers. Short responses about privacy strategies were analyzed using content analysis (Drisko & Maschi, 2015). In phase one, the first author read all responses line by line, identifying recurring strategies and developing a codebook that included strategy names, definitions, and example quotes. The codebook was refined as analysis progressed.

In phase two, these categories were added as variables to the survey dataset. The first author coded all participant responses for the presence (1) or absence (0) of each strategy. Strategies were not mutually exclusive; participants often described multiple approaches. For example, a response referencing two-factor authentication and limiting shared information was coded for both authentication tools and selective posting. In total, 761 instances of privacy strategies were coded across the five themes, indicating that participants used about two strategies on average ($M = 1.95$).

Comparing Use and Non-Use of Privacy Strategies. To examine whether the use of a particular strategy impacted participants' score on the privacy behavior scale, we conducted a series of one-tailed, two-sample t-tests. The social media privacy behavior scale is a 7-item variation of Sindermann et al's (2021) online privacy behavior scale. Participants were asked to rate on a 5-point Likert scale how often they engaged in certain online privacy behaviors like adjusting ad preferences or disabling tags ($M=17.26$, $SD=5.22$, $Min=0$, $Max=28$), with higher scores indicating stronger online privacy behaviors. The internal consistency of the scale was assessed using Cronbach's alpha, which yielded a value of 0.71. This indicates an acceptable level of reliability.

Results

Part 1: Personal Privacy Strategies

Using thematic analysis, we identified five privacy strategies based on women's answers to the open-ended question: platform settings ($n=259$), selective posting ($n=170$), selective friending ($n=102$), authentication tools and other cybersecurity

practices (n=98), and identity masking (n=80). We also coded minimal privacy engagement (n=40).

Platform Settings.

Participants described using built-in privacy controls to manage visibility and protect personal information. These included restricting audiences (e.g., “friends only”) and disabling features such as location tagging or real-time location sharing. One participant noted, “I make sure my accounts are private so only my friends can see them.”

Selective Posting.

Many participants reported protecting their privacy by carefully managing what they share online. This included self-censoring and avoiding the disclosure of personal details. A participant stated, “I don’t overshare. My life is private.”

Selective Friending.

Participants also managed privacy through intentional decisions about whom to connect with on social media, typically accepting requests only from people they knew offline. One participant explained, “I’ve only accepted friend requests from people I’ve met in real life.”

Authentication Tools and Cybersecurity Practices.

Some participants reported using strong passwords, two-factor authentication, and biometric login options. Some also noted practicing cautious browsing behaviors such as avoiding suspicious links or clearing browser data, such as “I try not to download anything that is unsafe or fall for any scams.”

Identity Masking.

Participants also described protecting their offline identities by using pseudonyms, fake names, or nicknames and employing tools such as virtual private networks (VPNs).

Minimal Privacy Engagement.

A smaller group reported taking few or no steps to protect their online privacy. These participants often described themselves as infrequent social media users and/or indicated they had not adjusted default privacy settings.

Part 2: Connecting Personal Strategies to Privacy Behavior Scale Score

Next, we examined how the use of these personal privacy strategies was associated with participants’ scores on the privacy behavior scale. One-tailed, two-sample t-tests showed that only selective posting, use of authentication tools, and identity masking were associated with significantly higher privacy behavior scores, whereas minimal

privacy engagement was associated with significantly lower scores. Although platform settings (n=259) and selective friending (n=102) were primary strategies in participants' narratives, the use of these strategies was not associated with a higher score on the privacy behavior scale, suggesting that while these strategies may give a sense of safety and proactivity, they are not indicative of overall strong social media privacy behavior. See tables 1–6 in attached PDF.

Future Work

Only three of the five strategies participants identified as their primary privacy protection measures were significantly associated with higher scores on the privacy behavior scale. Among these, selective posting stood out as both one of the most commonly reported strategies and one of the few linked to stronger overall privacy behavior. Its popularity may stem from the control it offers over self-presentation and information disclosure across audiences. Unlike other strategies, selective posting is an ongoing, active process that occurs each time users share content. This user-driven and context-sensitive approach to privacy may feel more tangible and empowering than platform settings buried within menus or selective friending decisions made long ago. Because selective posting allows users to implement privacy “in the moment,” adapting to shifting contexts and audiences, it may be more closely aligned with broader privacy-protective behaviors.

By contrast, selective friending and the use of platform privacy settings, although commonly reported, were not associated with higher privacy behavior scores. One possible explanation is that these strategies are often implemented early in a user's engagement with a platform and are rarely revisited afterward. Once users curate their networks or adjust their privacy settings, these actions can become “set-and-forget” measures that provide a sense of protection without requiring ongoing engagement. Indeed, participants who reported using selective friending as a strategy had privacy behavior scores ($M = 17.74$) similar to those who did not use this strategy ($M = 17.10$), suggesting that selective friending may signal perceived control rather than meaningful differences in everyday privacy practices.

Future research should further examine the distinction between static privacy strategies, such as initial network curation or default settings, and dynamic strategies that require continual user decision-making. Understanding how these different forms of privacy management operate over time may clarify why some strategies correspond more strongly with broader privacy behaviors than others. Longitudinal or platform-specific studies could also explore whether users periodically revisit “set-and-forget” strategies or whether these practices remain largely unchanged after initial configuration.

References

- Ahmed, L. K., & Rashid, R. D. (2024). Advanced security: The Application Social Media and their security. ITM Web of Conferences, 64, 1-13.
<https://doi.org/10.1051/itmconf/20246401006>
- Boerman, S. C., Kruikemeier, S., & Zuiderveen Borgesius, F. J. (2018). Exploring Motivations for Online Privacy Protection Behavior: Insights From Panel Data. Communication Research, 48(7), 953-977.
<https://doi.org/10.1177/0093650218800915>
- boyd, d. (2010). Risk reduction strategies on Facebook.
<http://www.zephoria.org/thoughts/archives/2010/11/08/risk-reduction-strategies-on-facebook.html>
- Choi, H., Park, J., & Jung, Y. (2018). The role of privacy fatigue in online privacy behavior. Computers in Human Behavior, 81, 42–51.
<https://doi.org/10.1016/j.chb.2017.12.001>
- Choon, M. J. K. (2018). Revisiting the Privacy Paradox on Social Media: An Analysis of Privacy Practices Associated with Facebook and Twitter. Canadian Journal of Communication, 43(2), 339–358. <https://doi.org/10.22230/cjc.2018v43n2a3267>
- Dienlin, T., & Trepte, S. (2014). Is the privacy paradox a relic of the past? An in-depth analysis of privacy attitudes and privacy behaviors. European Journal of Social Psychology, 45, 285–297. <https://doi.org/10.1002/ejsp.2049>
- Faverio, M. (2023). Key findings about Americans and data privacy. Pew Research Center. <https://www.pewresearch.org/short-reads/2023/10/18/key-findings-about-americans-and-data-privacy/>
- Gottfried, J. (2024). Americans' social media use. Pew Research Center.
<https://www.pewresearch.org/internet/2024/01/31/americans-social-media-use/>
- Marwick, A., Fontaine, C., & boyd, d. (2017). “Nobody Sees It, Nobody Gets Mad”: Social Media, Privacy, and Personal Responsibility Among Low-SES Youth. Social Media + Society, 3(2), 1–14. <https://doi.org/10.1177/2056305117710455>
- Peer, E., Rothschild, D., Gordon, A. et al. Data quality of platforms and panels for online behavioral research. Behavior Research Method, 54, 1643–1662 (2022).
<https://doi.org/10.3758/s13428-021-01694-3>
- Sarikakis, K., & Winter, L. (2017). Social Media Users' Legal Consciousness About Privacy. Social Media + Society, 3(1).
<https://doi.org/10.1177/2056305117695325>

Sannon, S., & Forte, A. (2022). Privacy Research with Marginalized Groups: What We Know, What's Needed, and What's Next. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), Article 455, 1–33.

<https://doi.org/10.1145/3555556>

Sindermann, C., Schmitt, H. S., Kargl, F., Herbert, C., & Montag, C. (2021). Online privacy literacy and online privacy behavior—the role of crystallized intelligence and personality. *International Journal of Human–Computer Interaction*, 37(15), 1455–1466. <https://doi.org/10.1080/10447318.2021.1894799>

Trepte, S. (2021). Social media privacy model: Privacy and communication in the light of social media affordances. *Communication Theory* (1050-3293), 31(4), 549–570.

<https://doi.org/10.1093/ct/qtz035>

Wang, L. H., & Metzger, M. J. (2024). The Online Privacy Divide: Testing Resource and Identity Explanations for Racial/Ethnic Differences in Privacy Concerns and Privacy Management Behaviors on Social Media. *Communication Research*, 1–37. <https://doi.org/10.1177/00936502241273157>

12.10 What is ‘acceptable’ versus ‘problematic’ smartphone use for young adults? A qualitative study

Magaly Brodeur, Natalia Muñoz Gómez, Andrée-Anne Légaré, Ariane Girard, Catherine Hudon, Adèle Morvannou (Université de Sherbrooke)

Introduction

The rise of smartphone use has made mobile technology more accessible than it has ever been. Smartphone penetration has reached around 75% or above in most countries in North America, Europe and Asia (Poushter et al., 2024). As smartphone use becomes ubiquitous, especially among youth and young adults (Asselin et al., 2024), concern for the harms associated to problematic smartphone use (such as psychological distress, isolation, poor sleep quality, and impaired functioning) among this population have arisen in research (Candussi et al., 2023; Giansanti, 2025). An interesting approach to better understand behaviors in the field of prevention and harm reduction is to focus on the social norms, or the “unwritten rules” surrounding certain behaviors and practices among a specific group of individuals (González-Ponce et al., 2022; Hong et al., 2021). While few studies have begun exploring young adults’ perspectives of their own smartphone use and motivations for use (Alotaibi et al., 2022; Mostyn Sullivan et al., 2025; Yang et al., 2021), to our knowledge none have done so considering young adults’ normative perceptions. Given emerging adulthood is a period in which peer acceptance and belonging is key in guiding certain behaviors (Ciranka & Van Den Bos, 2019), it remains important to understand young adults’ perceptions of what is acceptable versus problematic regarding the time spent on their smartphones, and how this may influence their smartphone use behavior.

Objectives/Research Questions

This study aims to better understand the smartphone use habits of young adults with a potential problematic smartphone use, their perceptions of what constitutes an “acceptable” daily screen time, and what represents problematic smartphone use for them.

Methods

The present study reports on the results of phase one of a larger, two-phase qualitative research project. This study employed a qualitative descriptive design and used purposive sampling aimed at maximum variation. To be eligible to participate, potential participants had to be aged 18 to 25 years, present a potential problematic smartphone use according to a score of 32 or above out of 60 on the Smartphone Addiction Scale-Short Version (SAS-SV; Kwon et al., 2013; Lopez-Fernandez, 2017) and reside in the

province of Quebec (Canada). Recruitment was carried out through social media (e.g., Facebook, Instagram, LinkedIn) and on-site (e.g., university campuses, cafés, pharmacies) advertisements. Potential participants filled a pre-participation form to obtain their sociodemographic profile (e.g., sex, gender, age, ethnic/cultural identity, sexual orientation, education, region etc.), and their score on the SAS-SV. Eligible participants were contacted by email to schedule an online semi-structured interview with a member of the research team. An interview guide with open-ended questions was developed to gain a better understanding of participants' habits regarding their smartphone use (e.g., main applications used, notification management, social media use and reasons, daily smartphone screen time, etc.) and to explore the social norms surrounding smartphone among this population (e.g., perceptions of acceptability surrounding smartphone use in specific contexts). The present study focuses on participants' discourse regarding screen time and the participant's understanding of problematic smartphone use. The interview recordings were transcribed verbatim, and a thematic analysis will be carried out using NVivo (Software Version Release 15.3.2), following the method outlined by Braun and Clarke (2006) to draw a portrait of participants' experiences and opinions.

Results

The sample included 33 participants (11 men, 20 women, and two transgender individuals; mean age = 22 years) from ten administrative regions in the province of Quebec, Canada. The mean score on the SAS-SV among the sample was 38.6 (SD = 6.2). Participants indicated using their smartphones mainly for communication with friends, family and colleagues, and entertainment purposes such as social media, watching videos or listening to music. The main social media applications used by participants included Messenger (Meta), Snapchat, and WhatsApp, for communication; Instagram, TikTok, and YouTube for entertainment; and Facebook for event/marketplace browsing and school/work/leisure groups. Secondary uses included time management, GPS navigation, taking notes and playing games. Regarding notification management, practices generally varied, however notifications were rarely enabled for all applications. Some participants disabled notifications from applications that sent too many alerts and kept only those related to communication, such as messages and calls. We asked participants how many hours per day they thought it was 'acceptable' to spend on one's smartphone. While around 58% of participants indicated spending less than three hours on their smartphone was acceptable, only 18% reported that their daily average screen time was indeed under three hours (mean = 4.75 h), as indicated by their smartphone settings. Therefore, there seems to be a gap between what participants believe to be the norm (injunctive norm) and their real smartphone usage (descriptive norm; Cialdini et al., 1991). Participants also elaborated on certain nuances regarding the measurement of screen time, stating that not all screen time should be regarded as inherently negative, such as when applications run

in the background while the user is doing another activity (for example, listening to music while studying), or when digital platforms are used for learning, social connection, and other purposes judged as beneficial for participants. Furthermore, participants described other elements that for them, were indicative of a problematic smartphone use, aside from an elevated daily screen time. For example, some participants mentioned automatic checking behaviors (i.e., even when there are no new notifications), and neglecting other needs like sleeping or eating, were indicative of problematic smartphone use. These results constitute preliminary findings of the interview data, as analysis is currently underway. Final results of this study, including a thematic tree and main codes, will be presented at the International Conference on Social Media and Society 2026.

Future Work

The results of this study will allow a better understanding of the social norms surrounding smartphone use from the perspective of young adults. Results drawn from the first phase of the study will serve to inform intervention approaches and educational campaigns from an angle that considers the perspectives and needs of young adults. This will further be achieved through phase two of this research project, which involves the participation of young adults aged 18 to 25 years in a co-design workshop. The aim of this participative phase is to use research data from phase one of the study to propose, in conjunction with young adults, educational messages and interventions to reduce and prevent the harms associated with problematic smartphone use.

References

- Alotaibi, M. S., Fox, M., Coman, R., Ratan, Z. A., & Hosseinzadeh, H. (2022). Perspectives and Experiences of Smartphone Overuse among University Students in Umm Al-Qura University (UQU), Saudi Arabia: A Qualitative Analysis. *International Journal of Environmental Research and Public Health*, 19(7), 4397. <https://doi.org/10.3390/ijerph19074397>
- Asselin, G., Bilodeau, H., & Khalid, A. (2024). Digital well-being: The relationship between technology use, mental health and interpersonal relationships. *Statistics Canada*. <https://www150.statcan.gc.ca/n1/pub/22-20-0001/222000012024001-eng.htm>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Candussi, C. J., Kabir, R., & Sivasubramanian, M. (2023). Problematic smartphone usage, prevalence and patterns among university students: A systematic review.

- Journal of Affective Disorders Reports, 14, 100643.
<https://doi.org/10.1016/j.jadr.2023.100643>
- Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior. *Advances in Experimental Social Psychology*, 24, 1015–1026.
[https://doi.org/10.1016/S0065-2601\(08\)60330-5](https://doi.org/10.1016/S0065-2601(08)60330-5)
- Ciranka, S., & Van Den Bos, W. (2019). Social Influence in Adolescent Decision-Making: A Formal Framework. *Frontiers in Psychology*, 10, 1915.
<https://doi.org/10.3389/fpsyg.2019.01915>
- Giansanti, D. (2025). Smartphone Addiction in Youth: A Narrative Review of Systematic Evidence and Emerging Strategies. *Psychiatry International*, 6(4), 118.
<https://doi.org/10.3390/psychiatryint6040118>
- González-Ponce, B. M., Rojas-Tejada, A. J., Carmona-Márquez, J., Lozano-Rojas, Ó. M., Díaz-Batanero, C., & Fernández-Calderón, F. (2022). Harm Reduction Strategies among University Students Who Use Alcohol and Cannabis, and Related Psychological Variables: A Systematic Review. *Journal of Psychoactive Drugs*, 54(5), 403–418. <https://doi.org/10.1080/02791072.2021.2023240>
- Hong, F.-Y., Lin, C.-C., Lin, T.-J., & Huang, D.-H. (2021). The relationship among the social norms of college students, and their interpersonal relationships, smartphone use, and smartphone addiction. *Behaviour & Information Technology*, 40(4), 415–426. <https://doi.org/10.1080/0144929X.2019.1699959>
- Kwon, M., Kim, D.-J., Cho, H., & Yang, S. (2013). The Smartphone Addiction Scale: Development and Validation of a Short Version for Adolescents. *PLoS ONE*, 8(12), e83558. <https://doi.org/10.1371/journal.pone.0083558>
- Lopez-Fernandez, O. (2017). Short version of the Smartphone Addiction Scale adapted to Spanish and French: Towards a cross-cultural research in problematic mobile phone use. *Addictive Behaviors*, 64, 275–280.
<https://doi.org/10.1016/j.addbeh.2015.11.013>
- Mostyn Sullivan, B., George, A. M., & Rickwood, D. (2025). An exploration of motives that influence problematic smartphone use in young adults. *Behaviour & Information Technology*, 44(9), 1886–1900.
<https://doi.org/10.1080/0144929X.2024.2380102>
- Poushter, J., Gubbala, S., & Austin, S. (2024). 8 charts on technology use around the world. Pew Research Center. <https://www.pewresearch.org/short-reads/2024/02/05/8-charts-on-technology-use-around-the-world/>

Yang, Z., Asbury, K., & Griffiths, M. D. (2021). "A Cancer in the Minds of Youth?" A Qualitative Study of Problematic Smartphone Use among Undergraduate Students. *International Journal of Mental Health and Addiction*, 19(4), 934–946.
<https://doi.org/10.1007/s11469-019-00204-z>

12.11 A Scoping Review of Empirical Research on Social Exclusion and Responses to Misinformation

Ayana Cameron, Vanessa Volpe, MarQia Allen, Abbey Collins, Melissa Holland, Ore Badejoh, Julia Ross (North Carolina State University)

Introduction

Online misinformation is a pervasive digital phenomenon defined as unintentionally false or misleading information shared over the internet without malicious intent. It is distinct from disinformation, which involves intentional deception, and malinformation, which utilizes genuine information to cause harm (Aïmeur et al., 2023; Ecker et al., 2022; Karlova and Fisher, 2013). In the digital era, misinformation is uniquely virulent, traveling approximately six times faster than true information (Vosoughi et al., 2018). This was prominently observed during the COVID-19 "infodemic," which caused widespread confusion and eroded trust in public authorities (Aïmeur et al., 2023; Asai, 2023; Cinelli et al., 2021; Gupta et al., 2022).

Current psychological research has predominantly focused on individual-level characteristics to explain misinformation susceptibility, such as political ideology, conspiracist ideation, and reflexive open-mindedness (Bessi et al., 2015; Cinelli et al., 2021; Radzikowski et al., 2016; Schmidt et al., 2018; Sicilia et al., 2017; Wang et al., 2019). While valuable, these frameworks often overlook the fact that online users exist within complex social networks where interactions significantly shape behavior. There is a critical need to explore the socio-psychological context of misinformation engagement, specifically how social dynamics influence the way individuals evaluate and react to false content.

This review specifically examines social exclusion, a potent experience of being kept apart from others either physically (e.g., isolation) or emotionally (Riva et al., 2017, 2025), which threatens core psychological needs of belonging, self-esteem, control, and meaningful existence (Riva et al., 2025; Williams, 2007; Williams and Nida, 2022). Exclusion manifests in two primary forms: ostracism (e.g., being ignored or a lack of social attention) and social rejection (e.g., receiving negative social attention). Currently, the examination of social exclusion as it relates to online misinformation remains understudied, despite the inherently sociopsychological nature of social media and the significant role it plays in the diffusion of misinformation.

Objectives/Research Questions

The primary objective of this scoping review is to address the under examination of the socio-psychological influences that contribute to our evaluation and engagement with

online misinformation. Specifically, we evaluate the extent to which existing research explores the relationship between social exclusion and misinformation in digital social spaces. By synthesizing empirical research published between 2019 and 2024, our review identifies how various forms of social exclusion—ranging from anticipatory fears (e.g., FOMO) to structural isolation—have been examined as predictors of cognitive and behavioral responses to online misinformation. Ultimately, this work establishes a baseline to inform future research and interventions that account for the social motivations underlying the spread of misinformation. While previous psychological research has identified individual-level cognitive and motivational predictors of misinformation susceptibility, there remains a significant need to explore how the inherently social nature of online environments also shapes user behavior.

To address this, our review is structured around four primary research questions designed to report the extent, range, and nature of available evidence on this emerging topic:

1. How is online misinformation defined and operationalized in studies that specifically examine its intersection with social exclusion?
2. What types of online misinformation outcomes—spanning both cognitive evaluations (e.g., belief and credibility) and behavioral intentions (e.g., sharing and blocking)—have been utilized, and how are these outcomes measured?
3. What specific forms of social exclusion—ranging from anticipatory fears (e.g., Fear of missing out) to induced ostracism or systemic isolation (e.g., pandemic-related)—have been investigated as correlates of misinformation responses?
4. What are the key associations reported between social exclusion and online misinformation engagement, and what information gaps must be addressed to inform future research and methodology?

Methods

This scoping review utilized a systematic literature search protocol across seven major databases: PsycINFO, PsycARTICLES, Web of Science, JSTOR, Psychology Database, Communication Source, and PubMed. To capture the intersection of social exclusion and misinformation online, we generated specific search strings using eight terms for social exclusion (e.g., ostracism, isolation, and rejection), seven for misinformation (e.g., fallacy, and false news), and six for the online environment (e.g., online and social media).

The study followed six a priori inclusion criteria: articles had to be written in English, empirical, peer-reviewed, explicitly focused on misinformation outcomes, involve an online context, and investigate social exclusion as a predictor. Following the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) 2020 statement

flow protocol, the initial search yielded 562 sources (Page et al., 2021a). After removing duplicates and conducting a multi-stage screening of titles, abstracts, and full texts by independent reviewers, six peer-reviewed studies published between 2019 and 2024 were selected for final analysis. Data extraction was performed using a standardized codebook, with each article double-coded to ensure reliability regarding definitions, stimuli, measurements, and outcomes.

Results

The review identified considerable variability in how online misinformation and social exclusion are operationalized. Only one study provided an explicit definition of misinformation, though most characterized it as unintentionally false or misleading content (Pareek and Goncalves, 2024; Moravec et al., 2019; Wicks et al., 2023, Wasike, 2023). Participant responses to online misinformation were clustered into two domains: cognitive evaluations (e.g., belief and credibility) and behavioral intentions (e.g., sharing, blocking and deleting). Notably, sharing was the most common form of engagement investigated. Additionally, social exclusion was conceptualized as a multi-dimensional construct, ranging from anticipatory fears (e.g., FOMO or Fear of Isolation) to experimentally induced ostracism (e.g., receiving no "likes" in a mock social media environment) and systemic isolation (e.g., pandemic-related lockdowns). Finally, the studies in this review found several key associations between social exclusion and online misinformation. Two studies found that chronic isolation and the desire for social approval online were stronger drivers of misinformation belief and sharing than one-time experiences of being excluded or included (de Sousa, 2022; Wicks, 2023). On the other hand, two studies found that the anticipation of exclusion (e.g., fear of missing out and fear of isolation) motivated users to block misinformation sources to protect their social reputation (Ahmed et al., 2024; Wasike et al., 2023). Additionally, participants in one study were more likely to accept corrections about false information online when the correction came from someone in their own social group — corrections from outsiders often backfire and reinforce the false belief (Pareek and Goncalves, 2024). Similarly, the final study found that people processed false information more quickly (and less critically) when it aligned with their political identity, but resisted it more strongly (and took longer to think about it) when it challenged their political identity and beliefs. (Morevac, 2019).

Future Work

This scoping review identifies several ways to inform future research in this area. First, we advocate for a standardized definition of misinformation—distinct from disinformation and malinformation—to allow for better cross-study comparisons. There is also a clear need for a shared framework for validating misinformation stimuli. Additionally, we suggest moving away from viewing social exclusion as a static individual predictor and instead treating it as a dynamic process unfolding over time

and social contexts. Finally, this work underscores the importance of developing practical interventions that move beyond individual cognitive "corrections" (e.g., flagging content) to address the group dynamics and social motivations—such as the need for belonging and validation—that fundamentally drive the spread of misinformation.

The generalizability of the current literature on this topic is limited by four primary factors that should also be addressed in future work. First, the social media platform specificity used across studies restricts current insights to Facebook and X mock-ups, leaving it unclear if these processes apply to other platforms like TikTok, YouTube, or Reddit. Examining a variety of platforms in future studies may reveal unique socio-psychological dynamics due to known generational differences in platform usage (Gottfried & Park, 2025), as well as differences in graphical user interfaces (GUI) and levels of perceived anonymity (Mondal et al., 2020; Xu and Zhang, 2025). Additionally, while sample characteristics in the studies reviewed were globally diverse, not all studies measured the frequency of social media use or described the level of social media use in their sample. Third, stimuli heterogeneity without a standardized validation framework makes it difficult to determine if results stem from socio-psychological differences or disparate experimental environments. Taken together, this scoping review serves as a foundation for future research on the social processes (e.g., social exclusion and social identity) that play a role in how we evaluate and engage with online misinformation and highlights several important methodological considerations.

References

- Ahmed, S., Bee, A. W. T., Masood, M., & Han Wei, T. (2024). You have been blocked: Exploring the psychological, personality, and cognitive traits of blocking misinformation sources on social media. *Telematics and Informatics*, 89, 102123. <https://doi.org/10.1016/j.tele.2024.102123>
- Aïmeur, E., Amri, S., & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: A review. *Social Network Analysis and Mining*, 13(1), 30. <https://doi.org/10.1007/s13278-023-01028-5>
- Asai, T. (2023). Pandemic and infodemic: The role of academic journals and preprints. *Journal of Anesthesia*, 37(2), 173–176. <https://doi.org/10.1007/s00540-022-03149-1>
- Bessi, A., Zollo, F., Vicario, M. D., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015). Trend of Narratives in the Age of Misinformation. *PLOS ONE*, 10(8), e0134641. <https://doi.org/10.1371/journal.pone.0134641>
- Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National*

- Academy of Sciences, 118(9), e2023301118.
<https://doi.org/10.1073/pnas.2023301118>
- de Sousa, A., Schneider, G., de Carvalho, H., de Oliveira, L., Lima, S., de Sousa, A., de Araújo, T., Camargo, E., Oriá, M., Ramos, C., de Oliveira, R., Almeida, C., Ferreira, A., Teixeira, J., Lua, I., Souza, F., de Araújo, T., Fronteira, I., & Mendes, I. (2022). COVID-19 Misinformation in Portuguese-Speaking Countries: Agreement with Content and Associated Factors. *SUSTAINABILITY*, 14(1). (WOS:000751312700001). <https://doi.org/10.3390/su14010235>
- Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1(1), 13–29. <https://doi.org/10.1038/s44159-021-00006-y>
- Gottfried, J., & Park, E. (2025, November 20). Americans' Social Media Use 2025. Pew Research Center.
<https://www.pewresearch.org/internet/2025/11/20/americans-social-media-use-2025/>
- Gupta, A., Li, H., Farnoush, A., & Jiang, W. (2022). Understanding patterns of COVID infodemic: A systematic and pragmatic approach to curb fake news. *Journal of Business Research*, 140, 670–683.
<https://doi.org/10.1016/j.jbusres.2021.11.032>
- Karlova, N., & Fisher, K. (2013, March 15). A social diffusion model of misinformation and disinformation for understanding human information behaviour [Text]. Professor T.D. Wilson. <https://informationr.net/ir/18-1/paper573.html?fbclid>
- Marinucci, M., & Riva, P. (2025). Homelessness in Virtual Reality: Experiencing Social Exclusion Improves Attitudes and Behaviors Toward Unhoused People | *Cyberpsychology, Behavior, and Social Networking*. <https://www.liebertpub-com.prox.lib.ncsu.edu/doi/full/10.1089/cyber.2024.0027>
- Mondal, M., View Profile, Correa, D., View Profile, Benevenuto, F., & View Profile. (2020). Anonymity Effects. In *Proceedings of the 31st ACM Conference on Hypertext and Social Media* (pp. 69–74).
<https://doi.org/10.1145/3372923.3404792>
- Moravec, P. L., Minas, R. K., & Dennis, A. R. (2019). Fake News on Social Media: People Believe What They Want to Believe When it Makes No Sense At All. *MIS Quarterly*, 43(4), 1343–A13.
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson,

- E., McDonald, S., ... McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372, n160. <https://doi.org/10.1136/bmj.n160>
- Pareek, S., & Goncalves, J. (2024). Peer-supplied credibility labels as an online misinformation intervention. *INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES*, 188. (WOS:001217849200001). <https://doi.org/10.1016/j.ijhcs.2024.103276>
- Radzikowski, J., Stefanidis, A., Jacobsen, K. H., Croitoru, A., Crooks, A., & Delamater, P. L. (2016). The Measles Vaccination Narrative in Twitter: A Quantitative Analysis. *JMIR Public Health and Surveillance*, 2(1), e5059. <https://doi.org/10.2196/publichealth.5059>
- Riva, P., Marinucci, M., Telari, A., & Pancani, L. (2025). Updating the temporal need-threat model of ostracism: Challenges and future directions. *The Journal of Social Psychology*, 0(0), 1–13. <https://doi.org/10.1080/00224545.2025.2572648>
- Riva, P., Montali, L., Wirth, J. H., Curioni, S., & Williams, K. D. (2017). Chronic social exclusion and evidence for the resignation stage: An empirical investigation. *Journal of Social and Personal Relationships*, 34(4), 541–564. <https://doi.org/10.1177/0265407516644348>
- Schmidt, A. L., Zollo, F., Scala, A., Betsch, C., & Quattrocioni, W. (2018). Polarization of the vaccination debate on Facebook. *Vaccine*, 36(25), 3606–3612. <https://doi.org/10.1016/j.vaccine.2018.05.040>
- Sicilia, R., Giudice, S. L., Pei, Y., Pechenizkiy, M., & Soda, P. (2017). Health-related rumour detection on Twitter. 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 1599–1606. <https://doi.org/10.1109/BIBM.2017.8217899>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Wang, Y., McKee, M., Torbica, A., & Stuckler, D. (2019). Systematic Literature Review on the Spread of Health-related Misinformation on Social Media. *Social Science & Medicine*, 240, 112552. <https://doi.org/10.1016/j.socscimed.2019.112552>
- Wasike, B. (2023). You've been fact-checked! Examining the effectiveness of social media fact-checking against the spread of misinformation. *Telematics and Informatics Reports*, 11, 100090. <https://doi.org/10.1016/j.teler.2023.100090>
- Wicks, S. G., Hales, A. H., & Hennes, E. P. (2023). Does disseminating (mis)information restore social connection during a global pandemic? *Social and Personality Psychology Compass*, 17(10),

- Williams, K. D. (2007). Ostracism. *Annual Review of Psychology*, 58(Volume 58, 2007), 425–452. <https://doi.org/10.1146/annurev.psych.58.110405.085641>
- Williams, K. D., & Nida, S. A. (2022). Ostracism and social exclusion: Implications for separation, social isolation, and loss. *Current Opinion in Psychology*, 47, 101353. <https://doi.org/10.1016/j.copsyc.2022.101353>
- Xu, H., & Zhang, P. (2025). The Influence of Anonymity and Social Ties on Personal Experience Sharing: A Comprehensive Mixed-Methods Study. *Proceedings of the ACM on Human-Computer Interaction*, 9(1), 1–22. <https://doi.org/10.1145/3701210>

12.12 Is Outsourced Moderation Like Call Centre Work? An Examination of Commercial Content Moderation Ecosystem in India

Tania Chatterjee (The University of Queensland - Indian Institute of Technology Delhi)

Agam Gupta (Indian Institute of Technology, Delhi)

Introduction

Early journalistic investigations into content moderation, the practice of gatekeeping online content, highlighted the industry's dependence on outsourcing to low-wage labour markets (Chen, 2012, 2014; Stone, 2010). Characterised by 24/7 business cycles (Roberts, 2019) and tightly controlled management regimes where moderators' work is monitored to the second (Ahmad & Krzywdzinski, 2022; Newton, 2019), such arrangements have been described as operating within an existing call centre model (Newton, 2019; Roberts, 2019). This reliance offers platforms strategic advantages, including low-cost labour, round-the-clock coverage, and multilingual capability (Ahmad & Krzywdzinski, 2022; Newton, 2019; Roberts, 2019), enabling moderation to function efficiently at scale through infrastructures and workforces already trained to process high volumes of data and embody diverse user cultures (Roberts, 2016, 2019). However, Newton (2019), complicates this equivalence, arguing that moderators, though managed like call centre agents, shoulder far greater governance and psychological burdens, raising doubts about treating trust and safety as routine service transactions. This critique motivates our central question: to what extent is commercial content moderation comparable to call centre work?

Objectives/Research Questions

In this study we ask, to what extent is commercial content moderation comparable to call centre work?

Methods

With an emphasis on the Indian setting, we employed a qualitative methodology to investigate the parallels and discrepancies between call centre jobs and commercial content moderation. With ethical approval from the authors' institutional review board, we gathered data from moderators located in Delhi, Hyderabad, and Mumbai. The sole inclusion criteria for this study, given the opacity of the moderation industry, was prior expertise evaluating content for Indian social media platforms. Given the broad inclusion criteria, we were able to include moderators who joined the sector between 2017 and 2023, representing both local and worldwide platforms.

We used purposive snowball sampling, which is an effective technique for reaching sensitive or difficult-to-reach people (Biernacki & Waldorf, 1981). Eleven of the seventy-eight commercial content moderators who were first approached through a professional networking site consented to take part. Peer referrals led to the recruitment of five more moderators, for a total of sixteen participants (five females and eleven males). At one point in their lives, all of them worked for outsourcing companies. When no new information surfaced from the interviews, data saturation served as a guide for the ultimate sample size (Guest et al., 2006).

We conducted in-depth semi-structured interviews with participants from June 2023 to January 2024. The interviews were conducted virtual, over video conferencing software. The Delhi-based moderators were given the option of in-person interviews, but they chose to do virtual meetings instead, probably because of the anonymity and accessibility that such a format offers, especially when talking about delicate subjects (Braun & Clarke, 2013). Three of the sixteen interviews were recorded using handwritten notes, while thirteen were recorded with consent. Handwritten notes were turned into transcripts for coding, and verbatim transcriptions of recorded interviews were made in Bengali, Hindi, and English. The screensharing interviews were place over several sessions and lasted between 60 and 90 minutes. The outcome was a 96,076-word dataset that was 184 pages long. Internal rulebooks from three companies (two local and one multinational) and training films from two companies, all of which were supplied by moderators anonymously, were added to this dataset.

We analysed the data, following Braun and Clarke's (2006) theoretical thematic analysis, with Mirchandani's (2004) conception of the three practices that support global call centre labour as the guiding framework. Latent (conceptual) and semantic (data-driven) codes were merged in the first coding process, which was manually handled in NVivo. These codes served as the basis for our analytical comparison of call centre work and commercial content moderation. We arranged it into four candidate themes and 10 sub-themes, examined for consistency, and then further refined into four main themes and seven sub-themes.

Results

We identify four practices underpinning outsourced moderation: scripting, monitoring, synchronicity, and masking. First, scripting aims to script consistent decisions across the moderation supply chain through rulebooks and training. Yet it cannot eliminate discretion, as the complexity and variability of online content require contextual interpretation. As a result, scripting simultaneously constrains and enables consistency: it structures decisions through standardised procedures but also institutionalises contextual judgment.

Second, monitoring, enacted through performance metrics and technological oversight, becomes a central organising practice. Moderators are evaluated primarily on productivity (volume reviewed) and accuracy (alignment with platform guidelines), while the moderation tool structures their workflow, pacing, and breaks. Although these mechanisms are designed to ensure compliance with client Service Level Agreements (SLAs), they generate efficiency-driven workarounds that undermine the very consistency scripting aims to ensure.

Third, synchronicity is implemented through round-the-clock shifts and tight control over shrinkage to maintain continuous moderation is posed as necessary to ensure real-time moderation. However, synchronicity of human labour to real-time gatekeeping is only selectively relevant for cases such as livestreams or ex-ante content review (before publication), yet outsourcing firms must continuously sync the availability of human labour with the constant influx of content to meet client-defined SLAs.

Fourth, masking, used to obscure both platform policies and human involvement, is enforced through stringent non-disclosure obligations. Masking serves multiple functions for platforms: it prevents system gaming, shields platform values and moral logics, and maintains the illusion of seamless, automated governance while placing the labour itself out of public view.

Overall, in outlining the four practices that structure outsourced moderation, we argue that although moderation operates within call centre infrastructures, its organising practices diverge significantly in both form and function. In call centres, scripting minimises discretion; in moderation, it requires it, as workers must interpret complex and evolving content beyond what rulebooks can anticipate. Monitoring shifts from reinforcing scripts to becoming a central and autonomous layer of oversight, intensified by outsourcing pressures that incentivise exceeding client targets and adopting efficiency-driven shortcuts. Synchronicity, indispensable for real-time service in call centres, is only selectively necessary in moderation, yet firms maintain round-the-clock shifts to fulfil contractual obligations rather than technical demands. Masking extends beyond concealing location to obscuring the labour and decision-making processes behind platform governance.

Future Work

The differences lead us to contend that forcing the distinct operational logics of moderation into a call centre framework or repurposing call-centre-like mechanisms to outsourced moderation carries significant implications for both the gatekeeping of online spaces and the well-being of moderators. In terms of gatekeeping, while it is agreed that the work itself imbibes discretionary decision-making, in the current the guardrails for consistent decision making are broken due to uneven training and

stringent monitoring. As a result, moderation produces what can best be described as a “semi-consistent” process: decisions are structured by detailed scripts but simultaneously undermined by rushed training, efficiency pressures, and reliance on workarounds. Thus, any attempt to make moderation effective requires two areas of improvement. First, training must be treated as a priority rather than being sidelined by the pressure to bring moderators quickly to the production floor. Second, efficiency targets need to be recalibrated, since their current form undermines training and consistent decision-making by privileging speed over quality.

The layers of contracting and masking of labour and policies/processes, block any critical assessment of the current system and seeking accountability for how labour that unjustifiably burdened with cleaning the extremities of the internet. While large outsourcing firms often maintain “well-being departments,” yet these provisions are largely superficial. As ‘Moderator B’ observed, such initiatives did little to address deeper psychological strain: “Some companies have it as a formality. They will make you play games; how will it help you? The purpose of the wellness department needs to be improved. Most moderators don’t even know that they are suffering. As I said, it is a formality as they have to report it to the client, so they give these wellness breaks that allow moderators to go out for a smoke break or a short walk- how does it substantially improve anything?” Such accounts highlight the gap between the institutional performance of care and the lived reality of moderators who face sustained emotional stress. Thus, transparency and accountability around what happens behind the screens are essential, including clarity on investments in human resources, working conditions, and fair pay is need of the hour.

References

- Ahmad, S., & Krzywdzinski, M. (2022). Moderating in Obscurity: How Indian Content Moderators Work in Global Content Moderation Value Chains. In M. Graham & F. Ferrari (Eds.), *Digital Work in the Planetary Market* (pp. 77–96). The MIT Press. <https://doi.org/10.7551/mitpress/13835.003.0008>
- Chen, A. (2012, February 16). Inside Facebook’s Outsourced Anti-Porn and Gore Brigade, Where “Camel Toes” are More Offensive Than “Crushed Heads.” Gawker. <https://www.gawkerarchives.com/5885714/inside-facebooks-outsourced-anti-porn-and-gore-brigade-where-camel-toes-are-more-offensive-than-crushed-heads>
- Chen, A. (2014, October 23). The Laborers Who Keep Dick Pics and Beheadings Out of Your Facebook Feed. Wired. <https://www.wired.com/2014/10/content-moderation/>
- Klonick, K. (2018). The New Governors: The People, Rules, and Processes Governing Online Speech. *Harvard Law Review*, 131(6), 1598–1670.

- Mirchandani, K. (2004). Practices of global capital: Gaps, cracks and ironies in transnational call centres in India. *Global Networks*, 4(4), 355–373.
<https://doi.org/10.1111/j.1471-0374.2004.00098.x>
- Newton, C. (2019, February 25). The Trauma Floor: The secret lives of Facebook moderators in America. *The Verge*.
<https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona>
- Roberts, S. T. (2016). Commercial content moderation: Digital laborers' dirty work. In S. U. Noble & B. Tyne (Eds.), *The Intersectional Internet: Race, Sex, Class and Culture Online*. Peter Lang.
- Roberts, S. T. (2019). *Behind the Screen Content Moderation in the shadows of social media*. Yale University Press.
<https://yalebooks.yale.edu/book/9780300261479/behind-the-screen/>
- Stone, B. (2010, July 19). Policing the Web's Lurid Precincts. *The New York Times*.
<https://www.nytimes.com/2010/07/19/technology/19screen.html>

12.13 Mapping Canadian Immigration Discourse on YouTube: A Dashboard-Driven Analysis Using LLM-Assisted Multi-Label Sentiment Coding and Geo-Mention Extraction

Augustine Farinola (University of Alberta)

Introduction

Public debate about immigration increasingly unfolds on platformed spaces where news clips, political commentary, and influencer content attract large-scale participation through comment threads. YouTube is especially consequential because it combines algorithmic discovery, audiovisual framing, and persistent comment archives, yielding dense traces of public reaction to immigration events, policies, and identity narratives. Yet immigration discourse is difficult to study at scale because it is multilingual and highly contextual and mixes policy critique with lived-experience testimony and moralized conflict. Most of the comments users make on Youtube are shaped by platform dynamics (visibility, moderation, and engagement incentives), and researchers face additional barriers in transforming raw platform traces into analytically usable datasets while maintaining ethical care, methodological transparency, and reproducibility. To resolve some of these constraints, this project presents a research workflow and accompanying read-only dashboard for exploring YouTube discourse related to Canadian immigration across channels, videos, comments, and replies. The workflow combines (a) targeted keyword seeding with (b) channel expansion and incremental refresh to build a relational database, and then layers (c) LLM-assisted multi-label sentiment coding using a 25-category taxonomy with rationale strings, plus (d) optional geographic place-name extraction (country/province/city mentions). The dashboard [<https://www.youtubeanalytics.ca>] is designed to support exploratory analysis, methodological disclosure (including known limitations), and downstream research use through filtering, comparison views, and exportable tables.

Objectives/Research Questions

This submission has two goals:

To demonstrate a scalable and transparent workflow for building a research-ready dataset from YouTube content related to Canadian immigration; and

To report initial descriptive findings about how sentiment categories and geographic references vary across time and across content sources.

We focus on these three research questions:

What sentiment categories dominate YouTube comments about Canadian immigration, and how are categories distributed across polarity groups (negative/positive/ambiguous) within the collected corpus?

How do sentiment profiles differ across channels and videos?

Which countries and Canadian provinces/territories are most frequently referenced in titles/comments/replies, and how do place-mention patterns co-occur with particular sentiment categories over time?

Our findings, which are based on a purposive sample collected using seeding and channel expansion, should be interpreted as patterns observed within our corpus, not as estimates of the broader "Canadian public opinion." This contribution is therefore positioned as an empirically grounded poster, complemented by a transparent dashboard that serves as a methodological artifact to facilitate future validation and theory-driven research.

Methods

Data collection and structure:

We built a relational database from YouTube Data API v3 outputs, storing channels, videos, comments, and replies, linked by stable identifiers (video→channel; comment→video; reply→comment).

Collection proceeds in stages:

a targeted keyword seed (e.g., "Canadian immigration," "Canada visa," "deportation," "Punjabis/Indians in Canada") identifies candidate videos;

channel expansion retrieves additional videos from discovered channels to capture channel-level discourse; comments and replies are collected for each video, with pagination handled to preserve full threads;

incremental refresh uses per-video checkpoints to fetch new comments since the last run, supporting longitudinal updates without full rebuilds.

LLM-assisted multi-label sentiment coding:

Comment text is categorized using an LLM as an "assisted coder" against a 25-category taxonomy designed for immigration-related discourse. Coding is multi-label: a single comment may receive multiple categories. Each assignment includes an LLM-generated rationale string. Polarity (negative/positive/ambiguous) is treated as a property of categories rather than a separate per-comment label. The dashboard exposes category definitions (codebook), polarity distributions, and example rationales to support interpretability. We explicitly treat these labels as exploratory unless validated with human coding.

Geo-mention extraction:

We optionally extract place mentions from titles/comments/replies using named-entity recognition (spaCy where available) plus normalization to country and Canadian province/territory names/codes; results are aggregated by time buckets (day/week/month) and can be filtered by source type.

Analysis and presentation:

The Django-based dashboard is read-only (unmanaged models; database owned by the ETL) and supports filtering, sorting, and CSV exports for channels/videos/comments and derived views (timeline, compare, sentiment dashboards, and geo-mentions). Reported findings are descriptive aggregations from precomputed database views, focusing on patterns across time and across content sources.

Results

As of early March 2026 (access date recorded in the dashboard), the corpus contains on the order of hundreds of channels, hundreds of seed-and-expanded videos, and several hundred thousand top-level comments, with a subset of comments carrying LLM sentiment labels (multi-label assignments). Preliminary descriptive summaries show three consistent patterns.

First, the labeled subset exhibits a strongly negative-skewed polarity distribution, with negative categories substantially outweighing positive and ambiguous categories. This pattern is not interpreted as “overall public sentiment,” but as a feature of the collected discourse environment and sampling approach, which includes content likely to attract contention (policy announcements, news debates, and identity-laden narratives). The analysis reveals that negative discourse is not monolithic: it spans policy critique (institutional performance, fairness, bureaucratic delay), material anxieties (housing/jobs), and overt prejudice/xenophobia, as well as first-person accounts of discrimination and alienation. Positive categories appear but are comparatively concentrated in gratitude/hope/belonging frames and empathy/solidarity responses.

Second, comparisons across channels and videos show meaningful heterogeneity. High-engagement videos and certain channel types show sharper polarization and higher concentrations of conflict-laden categories, while other sources exhibit more mixed or ambivalent patterns. The dashboard’s side-by-side comparison views suggest that even within a single channel, sentiment profiles vary substantially across videos, implying that topical framing and event context shape comment dynamics beyond stable “audience ideology.”

Third, geo-mention aggregates indicate a strong spatial imagination in the discourse: commenters frequently reference Canada’s internal geographies (provinces/major cities) alongside origin and destination countries, often in ways that index migration

pathways, labor markets, and identity categories. However, geo-extraction introduces noise (false positives from entity recognition), so geographic counts are treated as indicative patterns rather than precise measurements. Together, these preliminary results motivate targeted validation and theory-driven analyses of how platformed publics narrate immigration through emotion, policy evaluation, and place-based reasoning

Future Work

Validation and reliability

The highest priority is systematic validation of LLM category assignments. We will draw a stratified sample of comments across channels/time.

report agreement metrics and error modes (confusions among related categories; sensitivity to sarcasm; handling of multilingual or coded speech).

examine whether model rationales help identify uncertainty and improve human interpretability.

Ethics and data minimization

Since the public comments were not originally intended for research, we will implement rigorous ethical safeguards throughout the analysis and dissemination process. This includes taking measures to avoid doxxing risks, minimizing the use of personal identifiers in shared outputs, paraphrasing any quoted comments in publications, and focusing on aggregate reporting. Should we share any data, our preference will be to release only comment and video IDs alongside derived features, wherever possible and in compliance with platform terms and community norms.

Method extensions

We will populate the richer per-comment label schema (e.g., stance, identity framing, toxicity, migration-stage cues) to move beyond sentiment into discursive structure, and we will evaluate how these dimensions interact with engagement and moderation dynamics. We also plan to improve sampling transparency by documenting seed selection more formally, tracking inclusion/exclusion criteria, and quantifying how channel expansion changes topical coverage.

Substantive research

With validated labels, we will develop theory-driven analyses connecting platformed affect to policy moments (time-series around salient events), channel ecosystems (news vs. creator spaces), and narratives of belonging/exclusion. The dashboard will remain a public-facing exploratory interface, while research claims will be grounded in validated subsets and clearly scoped inference.

References

- Akoka, J., Comyn-Wattiau, I., Lamassé, S. and du Mouza, C., 2019. Conceptual modeling of prosopographical databases integrating quality dimensions. *Journal of Data Mining and Digital Humanities*, (2019), pp.1–20.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*, 610–623. <https://doi.org/10.1145/3442188.3445922>
- boyd, d., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Bradley, J., 2014. Silk purses and sow's ears: Can structured data deal with historical sources? *International Journal of Humanities and Arts Computing*, 8(1), pp.13–27. <https://doi.org/10.3366/ijhac.2014.0117>
- Dönecke, A., 2020. From serial sources to modeled data: Changing perspectives on eighteenth-century court records from French Pondicherry. In: S. Schwandt, ed., *Digital Methods in the Humanities: Challenges, Ideas, Perspectives*. Göttingen: Perspectivia, pp.217–238.
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
- Higgins, S., Goddard, L. and Khair, S., 2024. Research data management in the humanities: Challenges and opportunities in the Canadian context. *Digital Studies/Le champ numérique*, 14(1), pp.1–22.
- Nyhan, J., Rockwell, G., Sinclair, S. and Ortolja-Baird, E., eds., 2023. *Making in the Digital Humanities: Textual Scholarship and Digital Humanities*. London: UCL Press.
- Pasin, M. and Bradley, J., 2014. Factoid-based prosopography and computer ontologies: Towards an integrated approach. *Digital Scholarship in the Humanities*, 30(1), pp.86–97.
- Quamen, H. and Bath, J., 2017. Digital humanities databases. In: C. Crompton, R. Lane and R. Siemens, eds., *Doing Digital Humanities: Practice, Training and Research*. London: Routledge, pp.145–162.
- Quamen, H., 2013. The limits of modelling: Data culture and the humanities. *Scholarly and Research Communication*, 3(4), pp.1–12.
- Roberts, S. T. (2019). *Behind the screen: Content moderation in the shadows of social media*. Yale University Press.

- Thaller, M., 2021. Can historical information be represented outside of a graph/hypergraph/network? In: GraphenTechnologien Hypotheses, pp.1–15.
- Townsend, L., & Wallace, C. (2016). Social media research: A guide to ethics. University of Aberdeen.
- McDonough, K., 2018. Relational databases in the humanities. Digital Scholarship in the Humanities, 33(3), pp.603–615.

12.14 Under the Influence: Beauty and Fashion Professionals' Understandings of Social Media Influencers in Promoting Diversity and Inclusion

Jordan Foster (MacEwan University) David Pettinicchio (University of Toronto)

Introduction

Social media influencers have grown tremendously in their reach and cultural authority (Leaver et al., 2020; Christin & Lu, 2024; Poell et al., 2022). Among them, fashion and beauty influencers stand out for their striking physical appearance and sartorial savvy (author, 2022) commanding significant attention not just online, but in the adjacent and notoriously exclusive fashion and beauty industries (Pham, 2015). The rise of influencers across digital platforms and mainstream campaigns and product advertisements is emblematic of a broader shift in creative labor across cultural industries (Abidin, 2018; Cunningham & Craig 2019; Duffy 2013; Duffy & Kang 2019), with platforms like Instagram, YouTube, and TikTok offering aspiring talent a “frictionless” point of entry (Parker et al., 2016; Cunningham & Craig, 2021). Among them, influencers of color, trans-influencers, and influencers with disabilities have been celebrated as indicators of growing diversity and inclusion in digital spaces including in the field of fashion and in the fashion-adjacent economy of beauty (authors, 2023; Trevisan & Farinosi, 2024). The presence of influencers with disabilities is especially noteworthy because disability is among the most unrepresented status characteristics in appearance-based industries—industries that have often reduced people with disabilities to “objects of stigma” (Heiss 2011; Garland-Thomson 2002). Existing scholarship on the topic of social media influencers and their ascent online has shed important light on the opportunities and challenges that marginalized influencers face, and on the distribution of influencers’ visibility more broadly (Hearn & Banet-Weiser, 2020; Pham, 2015; Poell et al., 2022). Crucially, these scholars show how cultural industries continue to differentially allocate resources to social media influencers, amplifying the visibility of some while marginalizing others including and especially influencers who embody visible markers of difference (Christin & Lu, 2023; Pham, 2015; Poell et al., 2022; author, 2022). Still, we know little about how industry figures make sense of or understand social media influencers, or the ways in which these digital mavericks shape field-level perceptions of diversity and inclusion. Our research builds on past work from critical media studies scholars and cultural sociologists, to show how industry figures perceive social media influencers located in fashion and beauty—the disruption digital platforms have brought—and how these perceptions shape professionals’ understandings of diversity and inclusion within the industries they work.

Objectives/Research Questions

Drawing on Ahmed's (2006, 2007) concept of diversity work and on scholarship surrounding platform labour and inequality, we address the role that established professionals working within the fashion and beauty industries play in shaping opportunities for inclusion. Diversity work represents a form of labour taken up to produce organizational or institutional change and remedy extant patterns of inequality (Ahmed, 2006, 2007). Problematically, this work is often "performed" rather than enacted, protecting industry figures and organizations from critique while eclipsing ongoing issues related to diversity and inclusion. Our approach, builds on Ahmed's conceptual framework, pairing diversity work alongside critical insights from among industry professionals in fashion and beauty. Using original survey data drawn from 244 industry professionals including talent agents, marketers, editors, stylists and photographers, we ask and answer three questions. First, how do established industry professionals make sense of diversity and inclusion in the field of fashion and in the fashion adjacent economy of beauty? Second, how are these sense making strategies shaped by the rise of social media influencers? And third, what can industry professionals' evaluations and perceptions of diversity and inclusion tell us about the future of appearance-based industries. In addressing these questions, we provide an empirically grounded explanation of the ways in which social media influencers propel industry change, the dynamic relationship that characterized change across cultural industries, and the limits of industry efforts toward inclusion and diversity broadly speaking.

Methods

To address our research questions, we employed an original survey which included numerous related modules around the state of the fashion and beauty industry as well as diversity and inclusion within traditional industry practices. The survey also contained specific questions about industry outlooks and professionals' views on emergent industry change brought on by social media and the work of influencers. Respondents were also asked to consider a series of text and image-based vignettes. Vignettes embedded in surveys enhance data collection by providing realistic scenarios that contextualize questions, reducing ambiguity in responses, and elicit nuanced attitudes and beliefs. For example, we asked respondents to examine photos of different fashion models and make decisions about how they would feature those models in a fashion magazine, as well as how others in the industry might posture those models. Both textual and image-based vignettes engage respondents in hypothetical reasoning, encouraging critical thinking not only about their values and decision-making but about others' as well. Our quota-based survey was administered in English and Spanish across the United States in February 2022. Quota sampling enables researchers to control the makeup of their survey sample by establishing targets for

demographic groups or characteristics – an especially important method when the group one wishes to analyze is specific and hard to reach through traditional sampling methods, as is the case with ours. We administered the survey through Qualtrics, an internet-based survey company that utilizes paid online research panels of respondents. This targeted sample includes professionals located in the fashion and beauty industry. These professionals are involved in the creation and assessment of appearance and include advertisers, brand managers, public relations professionals, designers, editors, writers, photographers, fashion buyers, marketing specialists, models, model agents/scouts, and stylists. To improve generalizability, we utilized detailed occupational codes from the Current Population Survey (CPS) to distribute our sample of 244 industry insiders across these roles. CPS is a monthly survey of about 60,000 households conducted by the U.S. Bureau of Labor Statistics (BLS). We also sought to distribute respondents across gender, disability status, and race to reflect as best possible the population of industry workers.

Given our sample size, and because all variables are categorical, we present our findings through a series of bivariate tables showing connections between key variables.

Results

Our analyses address possible ways in which social media influencers shape matters related to diversity and inclusion in appearance-based industries. We found that professionals overwhelmingly believe that fashion and beauty are diverse and becoming even more so. Their beliefs in this respect function as a form of “diversity work” through which professionals assert commitments toward the importance of inclusion all the while eclipsing important and ongoing issues related to representation in the fashion and beauty industries. Among the industry professionals we surveyed, those who follow social media influencers are more likely to believe the role of fashion is to reflect a realistic and diverse set of aesthetics and appearances. Read in one way, these attitudes among industry professionals align with popular narratives surrounding the democratic and inclusive potentials of social media platforms (Cunningham & Craig, 2019; Poell et al., 2022), that tend to commodify difference while obscuring systematic patterns in the maintenance and reproduction of inequality online. Read in another, industry attitudes may signal that social media influencers and the various forms of diversity they embody, are actively shifting industry perceptions in new, more inclusive directions.

That said, our analyses also reveal a set of complex and contradictory perceptions among industry professions working in fashion and beauty. In spite of broad consensus surrounding their industry’s movement toward diversity and inclusion, industry professionals are not uniformly positive about the representation of difference nor are they ignorant of ongoing developments in the fashion and beauty space. Professionals

acknowledge that specific markers of bodily difference remain underrepresented. They identify, for example, extremely narrow opportunities for models with disabilities both on the runway and across editorial covers. When asked whether a fashion editor might opt in favor of a model with disabilities, the professionals we surveyed were highly doubtful. Their assertions align closely with Ahmad's critical analyses of "diversity work," as professionals make rhetorical and discursive moves to place blame among their industry counterparts (see also Godart and Mears 2009). These same professionals were somewhat more positive in their appraisal of existing opportunities for racial minorities, suggesting that strides toward diversity and inclusion in the fashion and beauty industries may be uneven across status characteristics and markers of difference.

Future Work

Our research builds on existing work from communications and media studies scholars as well as cultural sociologists, who highlight the social and economic influence of platform environments and the opportunities and challenges these environments pose to marginalized creators (Hearn and Banet-Weiser, 2020; Pham, 2015; Poell et al., 2022). Specifically, we shed light on the role that industry personnel play in enabling and constraining opportunities for diversity and inclusion, tracing how taken-for-granted norms and perceptions limit the production and promotion of more varied beauty imagery and representations of bodily difference. Future work should examine how industry change and diversity efforts among multinational brands are understood across different social and cultural contexts. Second, although quota-based sampling is especially useful in targeting a specialized population and, where we ensured that our sample reflected the distribution of characteristics in the population, paid respondents do not always fully reflect the population of industry insiders. To our knowledge, there is no existing probability-based sample of this group. Furthermore, existing large-scale surveys that ask about occupation do not also ask about industry outlooks or about diversity and inclusion.

Borrowing from existing ethnographic approaches to the study of the appearance industries (Mears, 2011), future work might also consider ethnographic approaches to the study of professionals' decision making. This work should account for changes in the organization and marketing of fashion and beauty products in-step with the rise of social media influencers and their wide platform reach. Ethnographic work is well suited to map industry decision making and professionals' sense-making strategies, illuminating fine-grained details related to the processes and real-time negotiations that shape the production and promotion of diversity. Combining observations from among industry professionals and pairing these alongside influencers' own accounts would be especially beneficial, allowing researchers to comment on the reciprocal

dynamics that shape the negotiation and production of platformed content in the digital era.

References

- Abidin, C. (2018). #familygoals: family influencers, calibrated amateurism, and justifying young digital labor. *Social Media + Society* 3(2), 1–15.
- Ahmed, S. (2006). Doing diversity work in higher education in Australia. *Educational Philosophy and Theory*, 38(6), 745-768.
- Ahmed, S. (2007). The language of diversity. *Ethnic and racial studies*, 30(2), 235-256.
- Christin, A., & Lu, Y. (2023). The influencer pay gap: Platform labor meets racial capitalism. *New Media & Society*, 26(12), 7212-7235.
- Cunningham, S., & Craig, D. (2019). *Social media entertainment: The new intersection of Hollywood and Silicon Valley*. New York University Press.
- Duffy, B. E. (2013). *Remake, remodel: Women's magazines in the digital age*. University of Illinois Press.
- Duffy, A., & Kang, H. (2020). “Follow me, I’m famous: travel bloggers’ self-mediated performances of everyday exoticism.” *Media, Culture & Society* 42(2), 172-190.
- Godart F., & Mears, A. (2009). How do cultural producers make creative decisions? Lessons from the catwalk. *Social Forces* 88(2), 671–692.
- Hearn, A., & Banet-Weiser, S. (2020). The beguiling: glamour in/as platformed cultural production. *Social Media + Society* 6(1), 1–11.
- Mears, A. (2011). *Pricing Beauty: The Making of a Fashion Model*. Berkeley, CA: University of California Press.
- Parker, G., Van Alstyne, M., & Choudary, S.P. (2017). Platform ecosystems: how developers invert the firm. *MIS Quarterly*, 41(1), 255–266
- Pham, M. (2015). *Asians wear clothes on the internet: Race, gender, and the work of personal style blogging*. Duke University Press.
- Poell, T., Neiborg, D., & Duffy, B. (2022). *Platforms and Cultural Production*. Cambridge: Polity Press.

12.15 'Preparatory Social Capital' and the Gendered Platform Economy: Chinese Young Mothers Navigating WeChat-Based Multi-Level Marketing

Jingjing Fu (Queen Mary University of London)

Introduction

The flexible working patterns associated with multi-level marketing (MLM) models have long attracted female participants, closely tied to women's socio-cultural and economic positions within society and the household (Blakemore, 2017; Charity, 2012; Herbig & Yelkurm, 1997). The figure of the 'mompreneur' has become a prominent MLM recruitment narrative, constructing an idealised image of women who fulfil caregiving roles and achieve financial independence (Pearce, 2024). However, this discourse reinforces traditional gender expectations under the guise of postfeminist empowerment, while obscuring structural inequalities.

In contemporary China, the state promotes gender equality as part of its socialist legacy while simultaneously restricting independent feminist activism and reinforcing traditional gender roles. Feminist ideas are frequently depoliticised and reframed through individualised narratives of self-improvement and personal success, particularly within digital media and platform economies (Zhang & Tang, 2026; Li, 2024; L. Zhang, 2018; Chen & Lu, 2025), transforming structural gender issues into lifestyle guidance that foregrounds personal responsibility over collective struggle.

Existing studies examining gender, politics, and economy in digital China have predominantly drawn on public figures whose social status is already secured, such as Professor Shen Yifei (Zhang & Tang, 2026) or Olympic athlete Eileen Gu (Li, 2024). By contrast, this study focuses on young mothers involved in WeChat-based MLM activities, an area historically stigmatised in China and further complicated by the gig economy. This positioning means that gendered expression becomes simultaneously more invisible and more intense, as participants' social status itself requires justification. The study, therefore, examines how neoliberal postfeminist values are negotiated not from established authority, but from the precarious margins of the platform economy.

This paper examines how Chinese young mothers navigate these gendered structures through their social practices on WeChat, in the context of multi-level marketing facilitated by the Chinese coupon app Mi Yuan. Mi Yuan relies on users' social networks to acquire and maintain consumer networks through 'social fission', positioning young mothers as both targets and agents of platform commerce.

The paper firstly contextualises how conditions of contemporary China position young mothers in the (digital) labour market. It then examines how they segment social activities across WeChat's features, including silent chat groups for commercial links, curated feed posts without comments, and intensive self-expression in private settings. In particular, it draws on recurring themes around money/economic value from participants' narratives to examine how financial success suggests a broader reduction of social value to economic value, driven not only by platform infrastructure but by cultural and gendered structures.

Objectives/Research Questions

How do Chinese young mothers' gendered social practices on WeChat-based multi-level marketing platforms produce and negotiate the economic and affective value of social relations within the platform economy? In particular, the study examines how the segmentation of social activities across WeChat's features, ranging from silent commercial chat groups to intensive private self-expression, reveals the ways in which social relations are simultaneously claimed as platform 'private traffic' and experienced as personal friendship, producing what this study terms 'preparatory social capital'.

Methods

To address the research question, the study draws on 26 semi-structured interviews with young mothers who participate in at least two of the observed segmented social activities on WeChat, and digital ethnographic approaches such as the walkthrough method (Light et al., 2018) and the media go-along technique (Jørgensen, 2016) conducted in 2019 and 2020. Through discourse analysis of the participants' testimonies, alongside the examination of the digital infrastructure of this MLM model enabled by apps including Mi Yuan and WeChat, this study examines how participants perceive and attach meaning to the activities they conduct, and how their everyday lives are structured and framed by gender norms, cultural traditions, and the trajectory of state economic development. Simultaneously, the study attends to individual-level agency: how young women manoeuvre around and bypass established rules and norms, choosing behaviours that best serve their personal circumstances.

The study aims to contribute to existing debates on governance and power relations in platformisation (e.g., de Kloet et al., 2019; Nieborg, 2020; Nieborg & Poell, 2018; Poell et al., 2019; van Dijck, 2020; Van Dijck et al., 2021) by exploring how, from users' perspectives, the values associated with digital platform activities are reduced to economic values — a reduction fuelled not only by the digital infrastructure of platforms, but also by the cultural structures embedded within Chinese society. In doing so, the study seeks to further engage with discussions around exploitation in the digital economy and to offer the concept of preparatory social capital as a framework

for understanding how social relations are continuously created, maintained, and held in reserve within the gendered platform economy.

Results

First, the study situates participants within a historical and developmental context. The most active users of Mi Yuan, young mothers born after or raised during the 1980s, grew up during rapid economic growth following China's market reforms (Link & Liu, 1998). The one-child policy enabled only daughters to access significant educational opportunities (Chang, 2020; Ding, 2023; Lee, 2012). However, a relatively gender-relaxed upbringing has not insulated these women from Confucian cultural norms or structural labour market disadvantages. Policies protecting women's employment rights have paradoxically increased hiring costs, leading to discrimination at critical career stages. The lag in social welfare and childcare services has further intensified the burden on young mothers. (Dasgupta et al., 2015; Du & Dong, 2010).

Second, the findings reveal that participants' online social relations are framed through two competing terms: 'private traffic' (economic value), and 'friend' (affective value). Both embody contradictory meanings: the tension between private and commercial, and the contest between individual and platform over who controls the economic value of social networks. The perceived empowerment from Mi Yuan is linked to flexibility, financial sustainability, and the enhancement of family and female status. However, this autonomy frequently contradicts what the platform demands, exposing tensions of networked individualism within platform governance (Rainie & Wellman, 2012; Walkerdine & Ringrose, 2006). Nevertheless, I argue that what is telling in this case is that it extends the discussion around networked governance by questioning the nature of networks. By situating the money talk among my participants, I found that it is not only about economics; it is rather the language through which social legitimacy, gendered autonomy and self-worth are negotiated in a context where other avenues for status are foreclosed.

Such an argument is further sustained by the concept of 'preparatory social capital' that I use to capture how social relations function as a networked resource continuously created, perpetuated, reassembled, and restored by both individuals and organisations, held in reserve awaiting moments of redemption. Unlike conventional understandings of social capital (e.g., Bourdieu, Putnam and Lin), this concept foregrounds temporality and speculation: capital fuelled by hope and passion, structured by socio-cultural norms and state economic development, and mediated through gender. Crucially, the platform economy provides 'exits' for young mothers encountering gender inequality, yet these exits simultaneously draw participants deeper into layered power structures, complicating rather than resolving their subjectivity and subject position. This dynamic cannot be adequately captured by

conventional frameworks of atomisation, neoliberal individualism, or exploitation-empowerment binaries.

Future Work

First, longitudinal research tracking participants over time would enable a deeper examination of how 'preparatory social capital' is actually redeemed and what happens to social relations when their anticipated value does not materialise. Such research could reveal the temporal rhythms of hope, exhaustion, and reinvestment that sustain participation in the platform economy, and whether the 'exits' identified in this study ultimately lead to new forms of entrapment.

Second, the concept of 'preparatory social capital' has the potential to invite theoretical development through posthuman feminism. This study has shown that social capital is not simply possessed by individuals but is continuously shaped by platform architectures, algorithmic logics, and data infrastructures. A posthuman feminist framework would push this further by treating these nonhuman elements not as background conditions but as active agents co-producing gendered subjectivities and value. This would reposition preparatory social capital as emerging from entanglements between human affect, platform code, and socio-cultural norms, offering a more relational ontology that moves beyond the individualist assumptions embedded in conventional social capital theory.

References

- Blakemore, E. (2017). Multi-Level Marketing's Feminine Mystique: A Reading List. Longreads. <https://longreads.com/2017/10/16/multi-level-marketings-feminine-mystique-a-reading-list/>
- Bourdieu, P. (1986). The Forms of Capital. In R. J.G. (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241–258). Greenwood Press.
- Chang, S. (2020). Chinese Women, Marriage and Gender: Exploring the idea of women and marriage over time in the context of China [University of Windsor]. https://scholar.uwindsor.ca/major-papers/109/?utm_source=scholar.uwindsor.ca%2Fmajor-papers%2F109&utm_medium=PDF&utm_campaign=PDFCoverPages
- Charity. (2012, July 10). Statistically, the Majority of Direct Sellers are Women. MLM, Network Marketing, Direct Selling News, Videos, Articles, Legal Updates, and More. <http://mlmlegal.com/MLMBlog/statistically-the-majority-of-direct-sellers-are-women/>
- Chen, T., & Lu, Y. (2025). Discursive state feminism in the Xi era navigating gender discourse, nationalism, and global ambitions. *Journal of Gender Studies*, 0(0), 1–25. <https://doi.org/10.1080/09589236.2025.2580394>

- Dasgupta, S., Matsumoto, M., & Xia, C. (2015). Women in the labour market in China (ILO Asia-Pacific Working Paper Series, p. 34). ILO Regional Office for Asia and the Pacific.
- de Kloet, J., Poell, T., Guohua, Z., & Yiu Fai, C. (2019). The platformization of Chinese Society: Infrastructure, governance, and practice. *Chinese Journal of Communication*, 12(3), Article 3.
<https://doi.org/10.1080/17544750.2019.1644008>
- Ding, Y. (2023). *The World of Work for Women in China*. Henley Business School.
<https://www.henley.ac.uk/news/2023/the-world-of-work-for-women-in-china>
- Du, F., & Dong, X. (2010). Women's Labor Force Participation and Childcare Choices in Urban China During the Economic Transition. *Departmental Working Papers*, Departmental Working Papers, Article 2010–04.
<https://ideas.repec.org//p/win/winwop/2010-04.html>
- Herbig, P., & Yelkurm, R. (1997). A Review of the Multilevel Marketing Phenomenon. *Journal of Marketing Channels*, 6(1), Article 1.
https://doi.org/10.1300/J049v06n01_02
- Jørgensen, K. M. (2016). The media go-along: Researching mobilities with media at hand. *MedieKultur: Journal of Media and Communication Research*, 32(60), Article 60. <https://doi.org/10.7146/mediekultur.v32i60.22429>
- Lee, M.-H. (2012). The One-Child Policy and Gender Equality in Education in China: Evidence from Household Data. *Journal of Family and Economic Issues*, 33(1), 41–52. <https://doi.org/10.1007/s10834-011-9277-9>
- Li, Q. (2024). “Eileen Gu fetish” as a feminist phenomenon: The intertwining of feminist, neoliberal, and nationalist discourses on Chinese social media. *Feminist Media Studies*, 24(4), 869–882. <https://doi.org/10.1080/14680777.2023.2229049>
- Light, B., Burgess, J., & Duguay, S. (2018). The walkthrough method: An approach to the study of apps. *New Media & Society*, 20(3), Article 3.
<https://doi.org/10.1177/1461444816675438>
- Lin, N. (2001). *Social Capital: A Theory of Social Structure and Action*. Cambridge University Press. <https://www.cambridge.org/core/books/social-capital/E1C3BB67419F498E5E41DC44FA16D5C0>
- Link, P., & Liu, B. (1998, October 8). A Great Leap Backward? *The New York Review of Books*, 45(15). <https://www.nybooks.com/articles/1998/10/08/a-great-leap-backward/>
- Nieborg, D. B. (2020). Apps of Empire: Global Capitalism and the App Economy. *Games and Culture*, 155541202093782. <https://doi.org/10.1177/1555412020937826>

- Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society*, 20(11), Article 11. <https://doi.org/10.1177/1461444818769694>
- Pearce, J. (2024). Achieving motherhood: A qualitative analysis of MLM success stories online. *Feminist Media Studies*, 24(1), 1–17. <https://doi.org/10.1080/14680777.2023.2172056>
- Poell, T., Nieborg, D., & van Dijck, J. (2019, November 29). Platformisation [Info:eu-repo/semantics/article]. Alexander von Humboldt Institute for Internet and Society gGmbH. <https://doi.org/10.14763/2019.4.1425>
- Putnam, R. D. (2000). *Bowling Alone*. Simon & Schuster.
- Rainie, L., & Wellman, B. (2012). *Networked: The New Social Operating System*. The MIT Press. JSTOR. <http://www.jstor.org.ezproxy.sussex.ac.uk/stable/j.ctt5vjq62>
- van Dijck, J. (2020). Governing digital societies: Private platforms, public values. *Computer Law & Security Review*, 36, 105377. <https://doi.org/10.1016/j.clsr.2019.105377>
- Van Dijck, J., de Winkel, T., & Schäfer, M. T. (2021). Deplatformization and the governance of the platform ecosystem. *New Media & Society*, 146144482110456. <https://doi.org/10.1177/14614448211045662>
- Walkerdine, V., & Ringrose, J. (2006). Femininities: Reclassifying Upward Mobility and the Neo-Liberal Subject. In C. Skelton, B. Francis, & L. Smulyan, *The SAGE Handbook of Gender and Education* (pp. 31–46). SAGE Publications Ltd. <https://doi.org/10.4135/9781848607996.n4>
- Zhang, L. (2018). The contradictions of “women’s work” in digital capitalism: A “non-Western”/Chinese perspective. *Feminist Media Studies*, 18(1), Article 1. <https://doi.org/10.1080/14680777.2018.1407121>
- Zhang, Y., & Tang, L. (2026). Postfeminism with Chinese characteristics: The rise of the postfeminist intellectual celebrity in digital China. *Feminist Media Studies*, 0(0), 1–18. <https://doi.org/10.1080/14680777.2026.2622442>

12.16 Populist Posts vs. Community Notes : Strategies for Countering Election Misinformation

Hiroyuki Fujishiro (Department of Media and Communication Studies /Hosei University)Mone Saito (Uniquevision)

Introduction

The spread of misinformation during elections has become a critical challenge for democratic societies. As a countermeasure, the "Community Notes" feature on X has attracted significant attention, and similar systems are currently being tested on platforms such as YouTube and Instagram. Community Notes are designed to reflect diverse perspectives on misleading content through contributor input. A consensus-based algorithm and participant ratings determine whether a note is published. Research indicates that Community Notes effectively suppress the diffusion of misleading posts, and early publication is particularly crucial for maximizing this deterrent effect (Slaughter et al., 2025). However, a major challenge remains: the low publication rate of election-related notes. According to the Center for Countering Digital Hate (CCDH), in their study of the 2024 U.S. presidential election, approximately 74% of "accurate" Community Notes—those aligned with independent fact-checks and reputable sources—were not displayed to users. The CCDH further highlights that when contributors submit notes but the system does not publish them, misleading content about candidates continues to circulate unchecked.

In Japan, the dissemination of misinformation during elections is also a pressing concern. In the 2025 House of Councillors election, Sanseito—a party campaigning on a "Japan First" platform—expanded its parliamentary presence. International media outlets, including CNN (2025), classified Sanseito as a right-wing populist party, noting its similarities to European parties such as Alternative for Germany (AfD) and "Reform UK."

The rise of far-right populism has been identified as a key factor contributing to the spread of misinformation on social media (Törnberg & Chueri, 2025). Observers frequently noted that Sanseito disseminated misinformation via social media. According to the Mainichi Shimbun, Sanseito was the most frequent subject of fact-checking related to false or misleading information among all political parties during the 2025 election.

Notably, the publication rate of Community Notes in Japan initially remained at an extremely low level. In the 2024 House of Representatives election, the rate was a mere 6.1%, and in the Hyogo gubernatorial election held the same year, the final publication rate was 0% (although 3% were temporarily displayed). However, this figure rose to

13.6% during the 2025 House of Councillors election. Among all political parties, Sanseito was the primary subject of these notes. In particular, notes regarding the party leader, Sohei Kamiya, accounted for approximately 30% of all published notes during this period (Fujishiro, 2025).

Objectives/Research Questions

The objective of this study is to identify methods for increasing the publication rate of “Community Notes,” with a particular focus on addressing the challenge of low publication rates for election-related content.

To achieve this goal, the research focuses on Sohei Kamiya, the leader of Sanseito, whose posts received an exceptionally high rate of published notes compared to the general trend. By analyzing the interactions between the disseminator and the responders, this study addresses the following research questions:

RQ1: What specific types of posts did Kamiya disseminate during the election period?

RQ2: What strategies and content did “Community Notes” contributors employ in response to those posts to achieve a high publication rate?

By analyzing the strategies of both the disseminators and the contributors, this study aims to identify the factors that increase the likelihood of Community Notes successfully countering populist misinformation. Ultimately, it seeks to establish practical methods for enhancing the effectiveness of Community Notes in mitigating the influence of such claims

Methods

Data Collection Period and Political Context

The 2025 Japanese House of Councillors election was officially announced on July 3 and held on July 20. The election produced a significant shift in the political landscape: while the ruling Liberal Democratic Party lost seats, Sanseito expanded its parliamentary presence from one to fifteen seats. This study focuses on this critical period to analyze the interaction between populist communication and consensus-based contextualization through Community Notes.

Hybrid Data Collection

A hybrid approach was employed to ensure both data comprehensiveness and accuracy. First, all data regarding Community Notes were retrieved programmatically from X’s public dataset. This allowed for the precise collection of 824 notes related to the 2025 House of Councillors election. Second, posts from Sohei Kamiya’s X account were manually collected to ensure accuracy and completeness. For these 141 posts, text content, post types, and engagement metrics were recorded to ensure data precision.

Analysis of Community Notes

From the dataset, notes specifically targeting Kamiya's posts were isolated for detailed verification. The study examined the rating status, information sources, and textual content of these notes. Among Kamiya's 141 posts, 65 received Community Notes, with a total of 84 notes identified and analyzed. This high density of notes relative to the number of posts allows for a detailed examination of the strategies employed by contributors in a politically challenging environment.

Results*Analysis of Kamiya's X Posts and Dissemination Strategy*

Data analysis was conducted on all 141 posts made by Sohei Kamiya during the election period. Among these, the 65 posts that received "Community Notes" were selected for detailed analysis. The dataset revealed a strategic reliance on the "Quote" function, with 42 quotes (64.6% of the 65 posts) and 13 original posts.

Kamiya's strategy centered on two main tactics: positioning himself as a "victim of misinformation" and engaging in "response battles" with high-profile accounts to leverage their reach.

The "Victim" Tactic

The highest engagement (13,436,191 impressions) occurred when Kamiya quoted an anonymous account with only 200 followers, claiming, "This is a hoax... spread by those trying to disparage Sanseito."

Leveraging Influencers

A post criticizing political influencer Toru Hashimoto reached 9.5 million impressions. By initiating confrontations with prominent figures, Kamiya successfully amplified his visibility through their audiences.

Community Notes Publication Rates and Status

Of 84 submitted Community Notes, 56% reached the public: 33 marked "Helpful" (Published) and 14 "Showing for some" (Temporarily Published). Remaining notes were either "Needs more ratings" (33) or "Not Helpful" (4). This 56% publication rate significantly exceeds the 6.1% recorded during the 2024 House of Representatives election, suggesting contributors effectively maximized the tool's impact against the party leader's claims.

Correlation Between Information Sources and Publication

The effectiveness of notes varied according to the type of source cited:

News media: 27 notes (9 published, 6 temporary)

Politician/Party archives: 20 notes (14 published, 1 temporary)

Public Institutions: 15 notes (8 published, 1 temporary)

Internet-only media: 16 notes (1 published, 5 unpublished)

Notes citing primary archives (party/politician websites) or official public institutions achieved the highest publication rates. In contrast, notes relying solely on internet-only media often failed to reach consensus. Notably, formal fact-checking organizations were cited in only one note, which remained under evaluation. Notes that stayed "under evaluation" despite citing news media often contained subjective phrasing, such as "repeating false claims," which likely hindered the consensus-based publication algorithm.

Impressions and the Presence of Notes

The average impression count for posts with Community Notes (3.37 million) was more than double that of posts without notes (1.52 million). Although this appears to contradict prior research suggesting that notes suppress diffusion (Slaughter et al., 2025), it likely reflects that notes are more frequently attached to highly viral content. Because post-publication impression growth was not tracked, this observation remains preliminary.

Future Work

This study demonstrates that, while the right-wing populist leader maximized reach through strategic use of "victimhood" and "response battles," "Community Notes" contributors successfully countered these efforts by prioritizing high-quality evidence from news media and political archives. The 56% publication rate—substantially higher than in previous elections—indicates that decentralized fact-checking can serve as a potent deterrent against populist misinformation when contributors focus on verifiable primary sources. These findings provide a clear roadmap for how users can strategically navigate algorithmic requirements to provide corrective context in highly polarized political environments.

Although news media served as the primary source used for these notes, its application requires careful consideration. Journalism is often dismissed by populists as "spin" (Yi & Gaskin, 2025); however, it remains a vital resource for contributors. The study found that notes citing news media often remained "under evaluation" when contributors employed aggressive language to label a post definitively as a "hoax." Because X's evaluation criteria prioritize neutral, unbiased language and high-quality sources, the success of the analyzed notes suggests that contributors effectively aligned their phrasing with these algorithmic preferences to achieve consensus.

Despite these insights, the study is limited by its focus on a single party and its leader. To generalize the findings, future research should expand the scope to include other

political parties and international electoral contexts. As populist communication tactics continue to evolve, further investigation is necessary to determine whether these contributor strategies remain a sustainable defense for democratic discourse on social media platforms.

References

- Fujishiro, H. (2025). The effectiveness of Community Notes in addressing misinformation and disinformation II: A case study of the 2025 House of Councillors election. *Information Processing Society of Japan Electronic Intellectual Property* 110, 110(5), 1–7.
- Slaughter, I., Peytavin, A., Ugander, J., & Saveski, M. (2025). Community notes reduce engagement with and diffusion of false information online. *Proceedings of the National Academy of Sciences of the United States of America*, 122(38), e2503413122.
- Törnberg, P., & Chueri, J. (2025). When do parties lie? Misinformation and radical-right populism across 26 countries. *The International Journal of Press/Politics*. 19401612241311886.
- Yi, Z., & Gaskin, T. (2025). Populist narrative power in a globalized infosphere: a cross-language analysis. *Humanities and Social Sciences Communications*, 12(1), 1–13.

12.16 How TikTok influencers frame sustainability across cultures

Lina Gomez-Vasquez, Sebastian Madero (The University of Tampa)

Introduction

The rapid growth of social media platforms has created significant opportunities to raise awareness and encourage communication and engagement around sustainability (Kannengießer et al., 2021). Sustainability communication refers to the practice of making contributions and debates for a better ecological, economic, and social life (Ziemann, 2011). Communication of sustainability issues can use a deliberative approach (open and transparent discourse), an instrumental approach (one-way communication to achieve a specific goal), or a deliberative approach (call-to-action) (Weder et al., 2021). These approaches offer creative and committed solutions to construct meaning and motivate behavior (Golob et al., 2023). Sustainability communication encompasses topics such as climate change, mindful consumption, biodiversity/conservation, waste reduction, the circular economy, slow fashion, food systems, and policy change, among others.

Within this evolving landscape, TikTok is notable: its visibility, editability, relatability, and interactivity translate complex issues into accessible, support-building, engaging communities (Hautea et al., 2021). Prior work shows social media may encourage the adoption of sustainable behaviors (Tao et al., 2024). These behaviors may include lifestyle changes, such as adjustments in habits and consumption, participation in sustainability conversations (Tao et al., 2024), and involvement in public actions or sustainability advocacy (Confetto et al., 2023). Moreover, peer influence on social media is a deciding factor that can affect an individual's decision about sustainable consumption (Salciuviene et al., 2022), as influencers, by frequently sharing content, address the lack of awareness about adopting sustainable consumption. Literature shows that influencers are perceived as approachable, relatable actors whose recommendations resonate more effectively with audiences than traditional marketing techniques, thereby enhancing trust (Kim et al., 2021). Influencers are often treated as “ordinary” experts by their audiences, who accept their content as trustworthy and friendly; this perceived impartiality helps influencers, for instance, to promote zero-waste practices and other eco-friendly behaviors through concrete, everyday examples (e.g., reducing single-use plastics, using reusable bags) (Buczyńska-Pizoń, 2015). Social media influencers form emotional bonds with followers and help influence purchase decisions authentically (Sharma, 2023), recommending products as friendly advice (Kaabachi et al., 2021). As parasocial opinion leaders, influencers can simplify complex environmental issues, provide orientation, raise awareness, and mobilize

behavior, especially among younger or less-aware audiences (Huber et al., 2022; Stehr et al., 2015). Guided by this background, we examine how TikTok influencers frame sustainability across cultural contexts.

Objectives/Research Questions

This exploratory study investigates how sustainability-focused TikTok influencers in the United States, Germany, and Colombia employ framing strategies to shape audience perceptions. Although corporations can frame their messages in various ways, limited research has explored multiple framing conditions simultaneously, with most studies concentrating on specific corporate social responsibility dimensions, such as environmental efforts (Kwon et al., 2024). Therefore, this study seeks to contribute to the existing literature by identifying how TikTok influencers frame sustainability communication across multiple framing conditions to foster audience awareness and engagement. The project's theoretical framework is grounded in framing theory, which enables a deeper examination of how specific information presentations can influence audience attitudes and behaviors toward sustainability. Framing means choosing which aspects of a message (e.g., promoting a particular problem definition or specific subjects) to highlight on a website or social media post, and making them more salient than other aspects of the message, thereby shaping the audience's understanding of the topic (Entman, 1993). How a topic is presented in a message can shape audiences' attitudes toward it (Aarøe, 2011). Therefore, framing has emerged as a key strategy in marketing, particularly for corporate social responsibility advertising, where companies employ different frames to emphasize their socially responsible and sustainable initiatives (Chung & Lee, 2019; Kim et al., 2012). The literature also shows that most empirical work on sustainability influencers often remains descriptive or qualitatively oriented, underscoring the need for comparative and platform-specific analyses, such as examining how different framing strategies vary across countries and cultural contexts (Confetto et al., 2023; Huber et al., 2022). Furthermore, research indicates that mixing frame types on content may strengthen persuasive effects on attitudes and behavior (Davis, 1995). Kwon et al. (2024) observed gaps in the literature: few studies test multiple framing conditions simultaneously, most work centers on a single domain (typically environmental issues), and there is a lack of studies evaluating corporate social responsibility and sustainability framing on social media platforms. Therefore, to contribute to addressing some of these gaps in the literature, this study answers the following research questions: RQ1: How do sustainability influencers in different countries frame sustainability messages on TikTok? RQ2: How do sustainability messages vary by country in terms of topical emphasis, particularly environmental versus other sustainability issues? In this study, framing is operationalized using a set of deductively and inductively developed strategies designed to capture key dimensions of sustainability communication, as outlined in the method section.

Methods

We conducted a cross-cultural content analysis of 75 TikTok videos by sustainability influencers in Colombia, Germany, and the United States (25 videos per each). Countries were selected to represent sustainability communication across culturally distinct environments. Although their approaches differ, Germany's coordinated sustainability policies, the United States' fragmented and often polarized sustainability communication discourse, and Colombia's development-oriented sustainability agenda collectively offer meaningful variation in sustainability narratives on TikTok for this study. We employed a purposive sampling strategy to identify sustainability-focused TikTok influencers with a minimum of 20K followers. This threshold ensured that selected creators had established engaged communities. Five influencers from each country were selected based on article recommendations of top sustainable influencers in each country. This sample allows for analysis of influencers who vary in niche emphasis (environmental, social, economic), communication style, and strategic approach. We randomly selected five videos from each influencer. To avoid recency bias, videos were randomly selected rather than limited to the most recent posts. Pre-testing involved 10 videos to test the codebook's variables and categories, and both authors participated in refining the codebook. Second, we conducted intercoder reliability using Cohen's Kappa. Results showed high scores ranging from 0.84 to 1, with 40% of the sample (30 videos) randomly coded. Guided by framing theory and sustainability communication literature, videos were analyzed combining deductive constructs (Davis, 1995; Golob et al., 2022; Iyengar, 1991; Overton, 2018; Wang & Anderson, 2008) with inductive refinement from the sample. We used several variables to measure sustainability framing communication deductively, such as a) Advocacy (diagnostic framing and categories such as practical lifestyle, cultural/identity, economic/consumer choice, and similar) vs Non-advocacy (prognostic framing, policy/governance issues); b) Communications strategy (deliberative vs others such as constitutive, instrumental, engagement, etc.); c) Responsibility attribution (individual vs organizational); d) Behavior appeals (doing more or taking less frames, coded as present/absent), and e) Other framing types (gain/loss). At the same time, inductive categories, particularly those capturing platform-specific practices were recurrent in the data and analyzed such as a) De-influencing (e.g., discouraging behaviors, products, critiquing trends coded as yes/no), b) Topic depth (surface vs substantive), c) Tone (positive/playful vs other/negative), and d) Niche (environmental vs other-economical, social, or holistic). These variables/categories were selected because they capture key dimensions of sustainability messaging that are both relevant across cultural contexts and meaningful for comparing influencer communication.

Results

To answer RQ1, we conducted descriptive analysis to discover cross-national variation framing. Advocacy framing was dominant across the sample (84%), with identical distributions across the three countries. Positive/Engaging tone appeared in 55% of posts overall, but the distribution varied sharply: Colombia and the United States each had roughly two-thirds of posts coded as Positive/Engaging (64% and 68%, respectively), while Germany had only 32%, suggesting a more informational communication style. Depth of content also differed meaningfully. 45% of posts were substantive, but Germany had the highest proportion (56%), followed by the United States (40%) and Colombia (36%). De-influencing content appeared in 36% of posts, with Colombia showing the lowest prevalence (36%), the United States at 36%, and Germany at 48% the highest. Although not statistically significant in inferential testing, these descriptive differences suggest emerging cross-national patterns in de-influencing messaging. Responsibility attribution showed notable variation: 44% of posts overall emphasized individual responsibility, but Colombia displayed the strongest emphasis (56%), followed by the United States (44%) and Germany (32%).

Niche also varied: environmental-focused posts (52%) were predominant, but Germany had the lowest proportion of environmental content (40%), compared to 56% in the United States and 60% in Colombia. Other framing strategies, such as behavioral appeals (48% overall) and deliberative communication strategy (41%), were more evenly distributed across countries. Variables such as gain/loss and episodic/thematic framing were extremely rare, meaning influencers across all three countries rarely used them in their framing. To provide more in-depth analysis to RQ1 and to answer RQ2, we performed bivariate analyses that confirmed several statistically significant cross-national differences, using Pearson chi-square when assumptions were met and Monte-Carlo chi-square (10,000 replicates) when small cells were present. Depth differed strongly by country (Pearson $\chi^2 = 13.00$, $p = 0.0015$, $V=0.42$), driven by Germany's higher proportion of substantive posts. Tone also varied significantly (Monte-Carlo $p = 0.002$, $V \approx 0.41$), reflecting Germany's lower use of positive tone. Responsibility attribution showed significant variation (Monte-Carlo $p = 0.003$, $V \approx 0.39$), with Colombia more likely to emphasize individual responsibility, suggesting meaningful cultural differences in how influencers assign agency for sustainability action. The topical niche also differed significantly (Monte-Carlo $p = 0.0013$, $V \approx 0.42$), with Germany posting fewer environmental-focused videos, suggesting that German influencers may prioritize broader or more diverse sustainability issues. De-influencing showed a marginal pattern (Pearson $\chi^2 = 5.56$, $p=0.062$), suggesting emerging but not statistically confirmed cross national differences.

Future Work

The findings of this exploratory study provide initial answers about how influencers frame sustainability messages on TikTok and how they vary by country and

sustainability domain. Results show that while advocacy framing is consistently dominant across all three countries, other framing dimensions, such as tone, depth, responsibility attribution, and niche, vary meaningfully. The significant differences in tone, depth, and responsibility attribution suggest that affective and cognitive dimensions of sustainability messaging may be particularly sensitive to national context even within a global platform. Findings also indicate cross-national variation in topical emphasis (particularly environmental vs. other issues), but insufficient representation of economic and social sustainability content to support deeper comparative analysis. Future research should build on these insights and address several gaps identified in the literature. First, scholars have noted that sustainability communication is inherently complex and multidimensional (Ziemann, 2011; Weder et al. 2021). Expanding the dataset to include a larger, more diverse sample of influencers, additional countries, and longitudinal coverage would allow researchers to test a broader range of framing strategies and to examine how these strategies evolve over time. Such expansion would also enable more robust statistical modeling and cross-cultural comparisons aligned with calls for more comparative, platform-specific analyses (Confetto et al., 2023; Huber et al., 2022). Second, scholars have argued that framing effects depend not only on message content but also on how frames are combined or layered (Davis, 1995; Entman, 1993). Future research could incorporate mixed methods designs to test the causal effects of tone, depth, and responsibility attribution cues and related frames on audience attitudes. Several limitations on this study should be noted that are avenues for future work. The sample size limits statistical power and restricts the ability to analyze rare categories. Sparse cells required category collapsing and the use of Monte Carlo or Fisher's exact test, thereby reducing granularity. By integrating framing theory with empirical content analysis, the results of this exploratory study underscore the complexity of sustainability communication as a global social process (Ziemann, 2011) and highlight the role of influencers as relatable, persuasive actors capable of shaping public understanding of sustainability issues (Kim et al., 2012; Huber et al., 2022). Although exploratory, the study contributes to filling gaps identified in prior research regarding the examination of multiple framing conditions and the need to move beyond single-issue analysis (Kwown et al., 2024). As social media continues to evolve as a key platform for sustainability discourse, this work provides a foundation on the diverse framing strategies to communicate about sustainability across cultural environments.

References

- Aarøe, L. (2011). Investigating Frame Strength: The Case of Episodic and Thematic Frames. *Political Communication*, 28(2), 207–226.
<https://doi.org/10.1080/10584609.2011.568041>

- Buczyńska-Pizoń, N. (2021). The Promotion of the Zero-Waste Concept by Influencers in Social Media. *Journal of Public Governance*, 52(2), 63-74.
<https://doi.org/10.15678/ZP.2020.52.2.06>
- Chung, S., & Lee, S.Y. (2019). Visual CSR Messages and the Effects of Emotional Valence and Arousal on Perceived CSR Motives, Attitude, and Behavioral Intentions. *Communication Research*, 46(7), 926-947.
<https://doi.org/10.1177/0093650216689161>
- Confetto M.G., Covucci C., Addeo F., & Normando M. (2023). Sustainability advocacy antecedents: how social media content influences sustainable behaviours among Generation Z. *Journal of Consumer Marketing*, 40 (6), 758–774,
<https://doi.org/10.1108/JCM-11-2021-5038>
- Davis, J.J. (1995). The effects of message framing on response to environmental communications. *Journalism & Mass Communication Quarterly*, 72(2), 285-299.
<https://doi.org/10.1177/107769909507200203>
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Golob, U., Podnar, K., & Zabkar, V. (2022). Sustainability communication. *International Journal of Advertising*, 42(1), 42–51.
<https://doi.org/10.1080/02650487.2022.2144035>
- Hautea, S., Parks, P., Takahashi, B., & Zeng, J. (2021). Showing They Care (Or Don't): Affective Publics and Ambivalent Climate Activism on TikTok. *Social Media + Society*, 7(2). <https://doi.org/10.1177/20563051211012344>
- Huber, B., Lepenies, R., Quesada Baena, L., & Allgaier, J. (2022). Beyond Individualized Responsibility Attributions? How Eco Influencers Communicate Sustainability on TikTok. *Environmental Communication*, 16(6), 713–722.
<https://doi.org/10.1080/17524032.2022.2131868>
- Iyengar, S. (1991). *Is anyone responsible?* University of Chicago Press.
- Kaabachi, S., Mohamed, K. K., Monyédo, R. K., & Ahmed, A. C. 2021. Celebrity Endorsement Vs Influencer Endorsement for Financial Brands: What Does Gen-Z Think? In *Management & Sciences Sociales*, 31, 55-81. <https://hal.science/hal-03767446v1>
- Kannengießer, S. (2021). Media reception, media effects and media practices in sustainability communication: State of research and research gaps. In F. Weder (Ed.), *The sustainability communication reader* (pp. 323–338). Springer Fachmedien

- Kim, E.E.K., Kang, J., & Mattila, A.S. (2012). The impact of prevention versus promotion hope on CSR activities. *International Journal of Hospitality Management*, 31(1), 43-51. <https://doi.org/10.1016/j.ijhm.2011.05.004>
- Kim, M., Doori, S., & Ahnlee, J. 2021. Consumer Response Toward Native Advertising on Social Media: The Roles of Source Type and Content Type. *Internet Research*, 31(5), 1656–1676. doi:10.1108/INTR-08-2019-0328
- Kwon, K., Lee, J., & Wang, C. (2024). Message Framing in CSR Advertising on Social Media during the COVID-19 Pandemic. *Journal of Sustainable Marketing*, 5(1), 50-65. <https://doi.org/10.51300/JSM-2024-119>
- Overton, H.K. (2018). Examining the impact of message frames on information seeking and processing: A new integrated theoretical model. *Journal of Communication Management*, 22(3), 327-345. <https://doi.org/10.1108/JCOM-10-2017-0114>
- Salciuvienė, L., Banytė, J., Vilkas, M., Dovalienė, A., & Gravelines, Z. (2022). Moral identity and engagement in sustainable consumption. *Journal of Consumer Marketing*, 39(5), 445-459. <https://doi.org/10.1108/JCM-03-2021-4506>
- Sharma, D. (2023). How not who: Message strategies adopted by celebrities v/s social media influencers. *Journal of Marketing Communications*, 31(1), 99–123.
- Stehr, P., Rössler, P., Leissner, L., & Schönhardt, F. (2015). Parasocial opinion leadership media personalities' influence within parasocial relations: theoretical conceptualization and preliminary results. *International Journal of Communication*, 982. <https://ijoc.org/index.php/ijoc/article/view/2717/1350>
- Tao, W., Tian, S., Sunny Tsai, W-H, & Seelig, M. (2024). The Power of Emotional Appeal in Motivating Behaviors to Mitigate Climate Change among Generation Z. *Journal of Nonprofit & Public Sector Marketing*, 36(1), 37-64, DOI: 10.1080/10495142.2022.2133058
- Wang, A., & Anderson, R.B. (2008). Corporate Social Responsibility Priming and Valence of CSR Framing on CSR Judgments. *Public Relations Journal*, 2(1), 1-19.
- Weder, F., M. Karmasin, L. Krainer, and D. Voci. 2021. Sustainability Communication as Critical Perspective in Media and Communication Studies—an Introduction. In *The sustainability communication reader*, eds. F. Weder, M. Karmasin, L. Krainer, and D. Voci, pp. 1–12. Wiesbaden: Springer VS.
- Ziemann, A. 2011. Communication theory and sustainability discourse. In *Sustainability communication*, eds. L. T. Godemann and G. Michelsen, pp. 89–96. Dordrecht: Springer

12.17 The medium's agenda or the audience's feedback? Tensions between editorial lines, metrics and comments in social media in digital newspapers in Chile

Francisca Greene (Universidad de los Andes), Cristian Muñoz (Universidad de la Serena)

Introduction

This study examines the tension between audience interests and editorial lines in the major national and regional digital media outlets in Chile. It analyses how editors and journalists incorporate metrics and user feedback from social media into content selection and prioritization processes. The sample included the five websites with the largest national reach according to the 2024 ComScore ranking (emol.com, biobiochile.cl, latercera.com, megamedia.cl, and chilevision.cl), along with digital media outlets from the country's five most populous cities without counting the capital (La Serena, Rancagua, Antofagasta, Valparaíso, and Temuco). The discussion raised is somewhat reminiscent of the two models of journalism identified by Schudson (2003): the fiduciary and the market models. In the former, journalists prioritize journalistic and editorial criteria to decide which news stories are important to the public. In contrast, the market model aims to respond to audience preferences. As will be seen, this tension, identified more than two decades ago, remains fully relevant in the discourse of editors of Chilean digital media outlets.

Previous studies have examined how editors of nationally distributed websites make decisions influenced by audience feedback (Greene et al., 2022 and González-Trujillo et al., 2022). Gronemeyer's research (2013) indicates that, in this context, the editor becomes more of a broker than a traditional gatekeeper, with click ratings ultimately influencing news coverage and publication decisions. This generally translates into "(...) greater engagement with audiences in the news selection process (...)" (Lee et al. 2014, p. 505). The study we propose is novel because it extends from the perceptions of editors of the main national digital newspapers to the editors of the most-read digital newspapers in the five most populous cities in the country, according to the latest census of 2024. In this sense, the study allows us to understand whether the relationship between audience feedback and editorial decisions is similar in national newspapers and in local media.

Objectives/Research Questions

In this regard, three questions guide this research: (1) To what extent does audience feedback influence editorial decisions in the most widely read digital newspapers in Chile? (2) Do the main Chilean digital newspapers tend to follow a fiduciary or a market

model? (3) Do national digital media outlets exhibit similar behaviour in incorporating audience feedback into editorial decisions compared to local digital media outlets?

Methods

The proposed research is a qualitative study examining the responses of 11 directors/editors of Chilean digital news outlets. Directors and editors were chosen because they are the ones who decide what and how to publish. In the case of digital media, they are particularly relevant because they make other decisions, such as the duration of a particular news story on the front page (see Greene et al., 2022). The editor's role in Chilean media is usually associated with selecting and prioritizing the most relevant news for citizens' decision-making (Salinas Muñoz & Stange Marcus, 2015; Puente et al., 2014; Greene et al., 2022; Greene-González & Lecaros-Menéndez, 2015). The editors or directors interviewed had similar roles: news selection, directing the editorial meeting, and making the final decision on what and how to publish. In one case (see Table 1), the journalist who made the final decisions about what and how to publish did not hold the position of director or editor, but rather an equivalent role.

Through in-depth, semi-structured interviews, we recorded the experiences of 11 directors/editors or journalists with similar roles in Chilean digital news outlets, regarding their perception of the influence of metrics on editorial decisions. According to Fuchs (2010), this was the number of interviews after which the authors considered that no new topics emerged, and data saturation had been reached. The interviews for this study were conducted by the three authors between March and June 2025, via Zoom or in person. While in-person interviews can facilitate a stronger connection between the researcher and the participant, we found no impediments to maintaining the key principles of indepth semi-structured interviews (Morris, 2015). These principles include flexible and fluid interaction between the interviewer and the interviewee, space to obtain further clarity and details from the interviewee regarding the question asked, and a balance between a structured interview guide and room for digression. Through semi-structured questions, a thematic discussion was achieved regarding the incorporation of audience feedback into editorial decisions.

Each interview lasted approximately 40 min. They were conducted in Spanish, transcribed, and translated into English. Thematic analysis was used to analyze the interviews (Braun & Clarke, 2006). This analysis was influenced by the research questions and the literature review, but it also included an inductive element that allowed the themes to evolve during the process. We analyzed each interview, developing general themes through a coding process (Coffey & Atkinson, 1996). All authors participated in the data analysis process. The qualitative data collected through the interviews were analyzed using a three-step thematic analysis performed

by three independent coders. The first step involved the emergent and inductive coding of four transcripts by the three coders to identify the main thematic areas.

Results

The findings presented and the tensions mentioned confirm what previous research has indicated regarding the value of audience feedback metrics and comments, but also the difficulty of balancing these resources with journalistic intuition and the normative value of journalism and quality information (Zamith, 2018). Thus, many editors see as a risk the fact that news is no longer just 'what journalists do', but also what the public wants it to be (Nguyen, 2013).

Likewise, the imposition of market rules on journalism is recognized as a reality that forces many media outlets, especially smaller ones, to abandon or change editorial lines to perceive greater economic benefits associated with higher readership and clicks (Vu, 2013). In this sense, it confirms what recent research has already indicated: journalists are more likely to incorporate audience feedback when media norms and/or policies guide them in this commercial dimension (Christin, 2020; Petre, 2021; Walters, 2022).

The tension highlighted in the proposed study thus allows for a renewed discussion and debate surrounding two opposing viewpoints. One viewpoint is to the fact that a greater audience presence in the newsroom (through feedback and readership metrics) enhances pluralism and the inclusion of diverse voices, while the other expresses concern about the decline in content quality and a sort of dictatorship of metrics.

It is confirmed that, while the incorporation of audience metrics can contribute to the economic sustainability of media outlets by attracting advertising investment, this dynamic significantly undermines normative journalistic practice, sometimes eroding the gatekeeper role, given the subordination of editorial decisions to digital traffic indicators. It also affects the watchdog role, compromising the selection and prioritization of content, and weakens journalism's oversight function, especially in relation to political and economic powers.

The results of this study reveal that the structural dependence on metrics not only reconfigures the news agenda but also erodes fundamental ethical principles, including accuracy, editorial independence, and commitment to the public interest. Professional practice is progressively transformed into one primarily driven by commercial criteria over normative journalistic values. This transformation is particularly acute in regional or local media outlets with limited resources, where the need to generate advertising revenue through increased traffic and engagement imposes a "dictatorship of metrics." The difference in audience participation in editorial decisions is striking: while national media associated with large brands tend to ignore

audience feedback, for journalists and editors of small local media, audience feedback is crucial in generating new content.

At the same time, the emergence of new professional roles dedicated to interpreting and acting upon analytical data may signal that the quantified view of the audience.

Future Work

By focusing solely on the perspective of some editors, this study leaves out other dimensions such as, for example, an ethnographic study of the newsroom with its routines, focus groups with audiences, content analysis of publications and written editorial policies, which would clearly enrich the study and are the subject of subsequent research. Moreover, the problem outlined here can be examined through additional theoretical approaches that further underscore the importance of editors' and journalists' discourse and its connection to journalistic practice. In this regard, theoretical work on metajournalistic discourse (Carlson, 2016), as well as its relationship to the problem of journalistic authority (Carlson, 2017) and the struggles over journalistic ways of knowing (Carlson, 2025), offers promising avenues for further exploration of this topic.

References

- Anderson, C. W. (2011). Between creative and quantified audiences: Web metrics and changing patterns of newswork in local US newsrooms. *Journalism*, 12(5), 550–566. [CrossRef]
- Anderson, C. W., Bell, E., & Shirky, C. (2015). Post-industrial journalism: Adapting to the present. *Geopolitics, History & International Relations*, 7(2), 32.
- Bastos, M. T. (2014). Shares, pins, and tweets. *Journalism Studies*, 16(3), 305–325. [CrossRef]
- Belair-Gagnon, V., Zamith, R., & Holton, A. E. (2020). Role orientations and audience metrics in newsrooms: An examination of journalistic perceptions and their drivers. *Digital Journalism*, 8(3), 347–366. [CrossRef]
- Bennett, W. L. (2003). The internet and global activism. In *Contesting media power: Alternative media in a networked world* (pp. 17–37). Rowman & Littlefield.
- Boczkowski, P. J. (2010). *News at work: Imitation in an age of information abundance*. University of Chicago Press.
- Boczkowski, P. J., Mitchelstein, E., & Walter, M. (2011). Convergence across divergence: Understanding the gap in the online news choices of journalists and consumers in western Europe and Latin America. *Communication Research*, 38(3), 376–396. [CrossRef]

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. [CrossRef]
- Bruns, A. (2011a). Gatekeeping, gatewatching, real-time feedback: New challenges for journalism. *Brazilian Journalism Research*, 7(2), 117–136. [CrossRef]
- Bruns, A. (2011b). News produsage in a pro-am mediasphere: Why citizen journalism matters. In *News online: Transformations and continuities* (pp. 132–147). Palgrave Macmillan.
- Carlson, M. (2016). Metajournalistic discourse and the meanings of journalism: Definitional control, boundary work, and legitimation. *Communication Theory*, 26(4), 349–368. [CrossRef]
- Carlson, M. (2017). *Journalistic authority: Legitimizing news in the digital era..* Columbia University Press.
- Carlson, M. (2025). Epistemic contests in journalism: Examining struggles over journalistic ways of knowing. *Digital Journalism*, 13(3), 362–377. [CrossRef]
- Chakraborty, A., Paranjape, B., Kakarla, S., & Ganguly, N. (2016). Stop clickbait: Detecting and preventing clickbaits in online news media. In *2016 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)* (pp. 9–16). IEEE.[CrossRef]
- Christin, A. (2020). *Metrics at work: Journalism and the contested meaning of algorithms.* Princeton University Press.
- Coffey, A., & Atkinson, P. (1996). *Making sense of qualitative data: Complementary research strategies.* Sage Publications, Inc.
- Deuze, M. (2008). Understanding journalism as newswork: How it changes, and how it remains the same. *Westminster Papers In Communication and Culture*, 5(2), 4–24. [CrossRef]
- Domingo, D., Quandt, T., Heinonen, A., Paulussen, S., Singer, J. B., & Vujnovic, M. (2008). Participatory journalism practices in the media and beyond. *Journalism Practice*, 2(3), 326–342. [CrossRef]
- Elsheikh, D., Jackson, D., & Jebril, N. (2024). The power of numbers: Four ways metrics are transforming the news. *Digital Journalism*, 12(8), 1165–1183. [CrossRef]
- Ferrer-Conill, R., & Tandoc, E. C. (2018). The audience-oriented editor. *Digital Journalism*, 6(4), 436–453. [CrossRef]
- Ferrucci, P. (2020). It is in the numbers: How market orientation impacts journalists' use of news metrics. *Journalism*, 21(2), 244–261.[CrossRef]

- Ferrucci, P., & Painter, C. E. (2016). Market matters. *Critical Studies in Television the International Journal of Television Studies*, 11(1), 41–58. [CrossRef]
- Focás, B. M. (2023). Audience editors: Between metrics and journalistic routines. *Universitas XXI*, 39, 161–179. [CrossRef]
- Fuchs, C. (2010). Alternative media as critical media. *European Journal of Social Theory*, 13(2), 173–192. [CrossRef]
- Gans, H. J. (1979). Deciding what's news: Story suitability. *Society*, 16(3), 65–77. [CrossRef]
- Gans, H. J. (2004). *Democracy and the news*. Oxford University Press.
- Gimmler, A. (2001). Deliberative democracy, the public sphere and the internet. *Philosophy & Social Criticism*, 27(4), 21–39. [CrossRef]
- González-Trujillo, R., Olate-Hidalgo, C., & Grassau, D. (2023). Impacto del entorno digital en los medios tradicionales chilenos: Percepciones y actitudes predominantes de sus protagonistas. *Palabra Clave*, 25(4), 1–35. [CrossRef]
- Goode, L. (2009). Social news, citizen journalism and democracy. *New Media & Society*, 11(8), 1287–1305. [CrossRef]
- Greene, M. F., Lecaros, M. J., & Diez, M. F. C. (2022). Medios de prensa digitales en Chile: Influencia del editor y de la audiencia en la reunión de pauta. *Cuadernos Info*, 51, 93–113. [CrossRef]
- Greene-González, M. F., & Lecaros-Menéndez, M. J. (2015). El trabajo del editor y la pauta en medios online chilenos. *Palabra Clave*, 18(2). [CrossRef]
- Gronemeyer, M. E. (2013). Digitalization and its effect on journalism products and practices in Chile. *Palabra Clave*, 16(1), 101–128. [CrossRef]
- Harrington, S. (2008). Popular news in the 21st century time for a new critical approach? *Journalism*, 9(3), 266–284. [CrossRef]
- Hirst, M. (2011). *News 2.0. can journalism survive the internet?* Allen & Unwin.
- Jofré Larenas, C. O. (2015). *Journalism and news cultures: Journalistic practices and online media in the Chilean newsroom* [Doctoral dissertation, Monash University].
- Lee, A. M., Lewis, S. C., & Powers, M. (2014). Audience clicks and news placement: A study of timelagged influence in online journalism. *Communication Research*, 41(4), 505–530. [CrossRef]

- Lee, E., & Tandoc, E. C. (2017). When news meets the audience: How audience feedback online affects news production and consumption. *Human Communication Research*, 43(4), 436–449. [CrossRef]
- Lewis, S. C., Holton, A. E., & Coddington, M. (2014). Reciprocal journalism. *Journalism Practice*, 8(2), 229–241. [CrossRef]
- Loosen, W., & Schmidt, J. (2012). (Re-) discovering the audience. *Information Communication & Society*, 15(6), 867–887. [CrossRef]
- MacGregor, P. (2007). Tracking the online audience: Metric data start a subtle revolution. *Journalism Studies*, 8(2), 280–298. [CrossRef]
- McChesney, R. D. (2004). *The problem of the media: US communication politics in the twenty-first century*. NYU Press.
- McManus, J. H. (1994). *Market driven journalism: Let the citizen beware?* Sage.
- Miller, S. (2019). Citizen journalism. In *Oxford research encyclopedia of communication*. Oxford University Press. [CrossRef]
- Morris, J. W. (2015). Curation by code: Infomediaries and the data mining of taste. *European Journal of Cultural Studies*, 18(4–5), 446–463. [CrossRef]
- Neheli, N. B. (2018). News by numbers. *Digital Journalism*, 6(8), 1041–1051. [CrossRef]
- Nguyen, A. (2013). Online news audiences: The challenges of web metrics. In K. Fowler-Watt, & S. Allan (Eds.), *Journalism: New challenges* (pp. 146–161). Bournemouth University.
- Örnebring, H., & Jönsson, A. M. (2004). Tabloid journalism and the public sphere: A historical perspective on tabloid journalism. *Journalism Studies*, 5(3), 283–295. [CrossRef]
- Petre, C. (2021). *All the news that's fit to click: How metrics are transforming the work of journalists*. Princeton University Press.
- Preston, P. (2009). Making the news. In *Journalism and news cultures in Europe*. Routledge.
- Puente, S., Edwards, C., & Delpiano, M. O. (2014). Modelamiento de los aspectos intervinientes en el proceso de pauta periodística. *Palabra Clave*, 17(1), 188–210. [CrossRef]
- Rebillard, F., & Touboul, A. (2010). Promises unfulfilled? 'Journalism 2.0', user participation and editorial policy on newspaper websites. *Media Culture & Society*, 32(2), 323–334. [CrossRef]

- Salinas Muñoz, C., & Stange Marcus, H. (2015). Burocratización de las rutinas profesionales de los periodistas en Chile (1975–2005). *Cuadernos.info*, (37), 121–135.
- Schlesinger, P. (1978). Putting «reality» together: BBC news. Available online: https://archive.org/details/puttingrealityto0000schl_f9q9 (accessed on 14 December 2025).
- Schudson, M. (2003). Click here for democracy: A history and critique of an information-based model of citizenship. In *The MIT press eBooks* (pp. 49–60). MIT Press. [CrossRef]
- Shoemaker, P. J., & Vos, T. (2009). Gatekeeping theory. Routledge. [CrossRef]
- Singer, J. B. (2005). The political j-blogger. *Journalism*, 6(2), 173–198. [CrossRef]
- Singer, J. B. (2013). User-generated visibility: Secondary gatekeeping in a shared media space. *New Media & Society*, 16(1), 55–73. [CrossRef]
- Steensen, S. (2009). What’s stopping them? *Journalism Studies*, 10(6), 821–836. [CrossRef]
- Tandoc, E. C. (2014). Journalism is twerking? How web analytics is changing the process of gatekeeping. *New Media & Society*, 16(4), 559–575. [CrossRef]
- Tandoc, E. C., & Ferrucci, P. R. (2017). Giving in or giving up: What makes journalists use audience feedback in their news work? *Computers in Human Behavior*, 68, 149–156. [CrossRef]
- Tandoc, E. C., Hellmueller, L., & Vos, T. P. (2012). Mind the gap: Between journalistic role conception and role enactment. *Journalism Practice*, 7(5), 539–554. [CrossRef]
- Tandoc, E. C., & Thomas, R. J. (2015). The ethics of web analytics. *Digital Journalism*, 3(2), 243–258. [CrossRef]
- Turner, G. (1999). Tabloidization, journalism and the possibility of critique. *International Journal of Cultural Studies*, 2(1), 59–76. [CrossRef]
- Usher, N. (2012). Going web-first at the Christian science monitor: A three-part study of change. *International Journal of Communication*, 6, 20
- Usher, N. (2018). Breaking news production processes in US metropolitan newspapers: Immediacy and journalistic authority. *Journalism*, 19(1), 21–36. [CrossRef]
- Vobić, I. (2015). From one-man band to integrated newsroom. *Journalism Studies*, 16(2), 175–190. [CrossRef]

- Vu, H. T. (2014). The online audience as gatekeeper: The influence of reader metrics on news editorial selection. *Journalism*, 15(8), 1094–1110. [CrossRef]
- Walters, P. T. (2023). Pushing for social change: How collaborations are recalibrating the journalistic mission. *Journalism Practice*, 19(2), 321–337. [CrossRef]
- Welbers, K., Van Atteveldt, W., Kleinnijenhuis, J., Ruigrok, N., & Schaper, J. (2016). News selection criteria in the digital age: Professional norms versus online audience metrics. *Journalism*, 17(8), 1037–1053. [CrossRef]

12.18 Short Videos under Conditions of Epistemic Uncertainty: How Young People Interpret Social Challenges

Elisa Haralampides (University of the Bundeswehr Munich)

Introduction

Short-videos are among the most popular social media formats used by young people. Short-video formats such as TikTok, Instagram Reels, and YouTube Shorts have become an important part of everyday meaning-making. Short videos do not serve entertainment purposes alone, nor are they merely part of a purely escapist form of media use. Rather, they increasingly engage with complex social and political issues. Short videos therefore play a significant role in shaping patterns of interpretation and meaning.

In contrast to other media formats, short videos are characterised in particular by their algorithmic curation. The selection of content that users are exposed to is highly individualised: no two users see exactly the same videos in the same order. As a result, knowledge about society is not only mediated through media content, but is also structured by algorithmic selection logics. Agency thus does not lie exclusively with either users or platforms, but emerges within a complex interaction between algorithms and users. Consequently, constructions of social reality may become further differentiated and fragmented.

This qualitative research contribution examines how young people perceive society, which images of society they develop, and to what extent content encountered in short videos influences their interpretative patterns regarding social challenges. A particular focus lies on how users negotiate the credibility of short videos and which heuristics they develop and apply in this process.

Objectives/Research Questions

The aim of this contribution is to explore how young people integrate short-videos into their everyday engagement with social challenges and what role this format plays in processes of meaning-making. Particular attention is paid to how interpretations are formed under conditions of epistemic uncertainty, how credibility is negotiated, and how short-video platforms are understood as social and communicative spaces.

Against this background, the paper addresses the following research questions:

RQ1: How do short videos influence young people and their reception of social challenges?

RQ2: How do young people negotiate the credibility of short-video content under conditions of epistemic uncertainty?

RQ3: How do young people interpret short-video platforms as spaces of information, discourse, and experience?

Methods

The study employs qualitative methods of data collection and analysis. The methodological framework is grounded in Grounded Theory, which, through its interpretative and constructivist research paradigm, provides the basis for how the data are approached. Data were collected using a semi-structured interview guide (Reiter & Witzel, 2022). In addition to audio recordings, video recordings were also collected in order to capture not only the perspectives of the participants but also their short-video feeds themselves. This resulted in a rich data corpus comprising 30 interviews, with approximately 32 hours of audio material and five hours of video material.

To analyse this heterogeneous dataset, the Grounded Theory approach according to Strauss and Corbin is applied (Corbin & Strauss, 2015). The audio and video material is transcribed, integrated, and analysed using coding procedures. In order to adequately address the multimodal nature of the data, the analysis combines classical Grounded Theory, originally developed primarily for textual data, with video analysis methods from Visual Grounded Theory (Konecki, 2019) and interpretative video analysis (Knoblauch et al., 2014).

Results

The social challenges most frequently mentioned by the participants include insecurity in public spaces, migration, processes of social polarisation, the risk of poverty, and war (listed in alphabetical order and not as a ranking of perceived urgency).

With regard to research question one, preliminary findings indicate that short videos are not the only actors shaping patterns of interpretation. Rather, participants interpret social phenomena within a complex interplay of various sources of information and experience. Everyday observations, shared experiences from family and friends (Katz & Lazarsfeld, 1964), news media, and political sources such as election programmes also play an important role in how social issues are interpreted.

Within the short-video format, three analytically distinct dimensions of meaning-making can be identified: the level of the creator (who posted the video), the level of the video content itself (which topics and narratives are addressed), and the level of the comment section (how other users respond to the video).

Regarding research question two, the findings show that the reception of short videos activates a complex interpretative process that does not simply involve the uncritical incorporation of new knowledge into existing stocks of knowledge (Berger & Luckmann,

1991). Instead, it often generates epistemic uncertainty among participants. To cope with this uncertainty, young people employ various strategies, with plausibilisation strategies being used far more frequently than validation strategies.

Along the different dimensions of meaning-making, users attempt to reduce epistemic uncertainty by plausibilising the short video with reference to different points of orientation. They assess the creator (e.g. familiarity, reach, perceived authenticity), the message conveyed in the video (its compatibility with their stock knowledge and its subjective imaginability), and the comment section as a form of social follow-up communication (to what extent comments by other users help to contextualise the content).

In addition, participants report further strategies for dealing with epistemic uncertainty, including online research, asking reference persons, critical reflection, and reliance on what they describe as a “gut feeling”, understood as common sense knowledge (Berger & Luckmann, 1991).

With regard to research question three, preliminary results indicate that content overload, which cannot be processed in its entirety, as well as platform logics perceived as addictive, are viewed critically. Many participants also describe a perceived decline in their own attention span as problematic. Moreover, short-video platforms, particularly Instagram and TikTok, are perceived as arenas of highly polarised social debates.

Future Work

In the next phase of the project, the analysis of the data material will be continued. So far, 20 out of 30 interviews have been analysed in line with the logic of theoretical sampling. A third and final round of data collection is planned to be completed by July 2026 in order to achieve theoretical saturation.

Methodologically, the analysis will build on the completed stages of open and axial coding and will be further developed through selective coding. The aim is to identify core categories and systematically elaborate their relationships in order to develop a theory grounded in the data that allows for a robust response to the research questions. Subsequently, the developed Grounded Theory will be situated in relation to existing media-theoretical and sociology-of-knowledge approaches. This step will examine the theoretical compatibility of the findings and clarify the specific contribution the study makes to existing research.

Furthermore, the results suggest that participants’ attribution of agency to algorithmic systems could be further explored through perspectives from Actor-Network Theory (Latour, 2007), particularly with regard to the interplay of human and non-human actors in the constitution of collective meaning-making processes.

Finally, the reconstructed strategies of credibility assessment and plausibilisation point towards potential connections with debates on media literacy. In this sense, the findings may be relevant beyond the immediate analytical context, contributing to broader discussions on reflexive media use and dealing with uncertainty in digital information environments.

References

- Berger, P. L., & Luckmann, T. (1991). *The social construction of reality: A treatise in the sociology of knowledge* (Repr. in Penguin Books). Penguin Books.
- Corbin, J. M., & Strauss, A. L. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (Fourth edition). SAGE.
- Katz, E., & Lazarsfeld, P. (1964). *Personal influence: The part played by people in the flow of mass communications* (1st Free Press paperback ed). The Free Press of Glencoe ; Collier-Macmillan.
- Knoblauch, H., Tuma, R., & Schnettler, B. (2014). Video Analysis and Videography. In U. Flick (Hrsg.), *The SAGE Handbook of Qualitative Data Analysis*. SAGE Publications, Inc. <https://doi.org/10.4135/9781446282243>
- Konecki, K. T. (2019). Visual Images and Grounded Theory Methodology. In A. Bryant & K. Charmaz (Hrsg.), *The SAGE Handbook of Current Developments in Grounded Theory*. SAGE reference.
- Latour, B. (2007). *Reassembling the social: An introduction to Actor-Network-Theory* (1. publ. in pbk). Oxford Univ. Press.
- Reiter, H., & Witzel, A. (2022). *Das problemzentrierte Interview – eine praxisorientierte Einführung* (1. Auflage). Juventa Verlag.

12.19 Origins of City Reputation: Measuring Bias in Online Reviews Using Agentic AI and NLP

Rafiazka Hilman, Júlia Koltai (ELTE Centre for Social Sciences)

Introduction

Online reviews and other user-generated content have been shown to exert a powerful influence on consumer behavior and on the collective perception of places. Numerous studies document that consumers rely heavily on peer reviews when making purchasing and travel decisions. For example, Lackermair (2013) finds that restaurant reviews are a primary information source for diners while Bae and Lee (2011) report that consumers treat online community reviews as highly credible when researching products. In the tourism context, online travel reviews (OTRs) are frequently used in option comparison and destination choice and become an important indicator of a destination's post-visit image (Marine-Roig and Huertas, 2020; Guo et. al., 2021; Zelenka et. al., 2021). Therefore, by aggregating thousands of individual opinions, online reviews deliver strong signals that can sway purchasing and travel intentions on a large scale (Chen et. al., 2022).

Beyond individual perception, aggregated user reviews contribute to the broader reputational landscape of cities and places. Tourism researchers use OTRs as a form of electronic word-of-mouth to study destination image and place branding. In this context, tourists' written reviews have been used to infer the 'received image' of a city based on its perceived quality, appeal, and safety after visiting. More broadly, a city's visibility and popularity on social media and review platforms correlates with its reputation (Guo et. al., 2021). This suggests that the volume and tone of online content represented by reviews, ratings, and comments can shape outsiders' perceptions of a city's overall quality and attractiveness. Consequently, user reviews act as aggregated signals of perceived urban characteristics that feed into the collective reputation of places.

These reputational dynamics are grounded in the perceived trustworthiness of reviews, which, however, may contain implicit bias or uneven credibility. Existing computational approaches for the analysis of such content rely predominantly on corpus-based natural language processing (NLP) techniques, including sentiment analysis and embedding-based bias measures. Recent advances in agentic AI capable of autonomous evaluation, explanation, and deliberation suggest new possibilities for capturing bias and trust in ways that more closely resemble human interpretation. However, it remains unclear how reputation is understood as an emergent, place-based construct in agentic assessments compared to established corpus-based

approaches. This study addresses this gap by systematically comparing agentic and NLP-based measurements of bias and trust in online reviews and examine their implications for city-level reputation formation.

Objectives/Research Questions

This study aims to evaluate whether agentic AI systems provide distinct and complementary measurements of bias and trust in online reviews relative to traditional NLP-based approaches, and how these measurements relate to the construction of city reputation. Specifically, we ask:

1. To what extent do agentic AI assessments of bias and trust align with corpus-based NLP measurements?
2. How does the two shape our understanding of city reputation and urban context as an aggregate outcome of trusted (or distrusted) online reviews?

Methods

We propose a methodological framework that operationalises agentic AI (Bandi et. al., 2025) as a structured evaluative instrument for online review analysis. Using the Yelp Open Dataset covering Las Vegas and Toronto, we construct a corpus of user-generated reviews linked to business identifiers and city-level metadata, enabling reputation to be analyzed as a place-based aggregate outcome derived from review trustworthiness. These two cities contain the largest volumes of businesses, reviews, and user-business interactions in the dataset, while exhibiting distinct urban characteristics. Las Vegas is widely known as a tourism-driven city shaped by the entertainment and hospitality sectors, whereas Toronto exhibits more routine and locally grounded consumption patterns associated with its large residential metropolitan population. This data coverage allows us to assess reputational dynamics across contrasting urban systems while maintaining platform consistency.

Each review is evaluated along two conceptual dimensions namely bias and trust using a multi-agent AI design. Autonomous agents are tested independently and we follow fixed evaluation protocols to ensure reliability and comparability. Specifically, we introduce two primary agents: i) a Bias Agent, tasked with identifying evaluative skew, stereotyping, or unjustified generalization; and ii) a Trust Agent, tasked with assessing perceived credibility, internal consistency, and evident-based grounding. Agents receive only the review text and are explicitly prevented from accessing metadata such as ratings or business popularity. This role-conditioned prompting procedure is the basis to train the agents in the shared base model of GPT-5.4 and Sonnet 4.6. Model temperature is fixed at a low value to minimize stochastic variability, and outputs are constrained to a structured schema to facilitate downstream analysis.

Evaluation protocols are implemented via standardized prompts presented in Table 1. The Bias Agent is instructed to: i) read the review; ii) identify biased framing or language; and iii) assign a bias score on a continuous scale from 0 (no bias) to 1 (high bias). The Trust Agent follows an analogous protocol, producing a trust score on the same scale along with a short explanation. Operationalise reputation at the city level c , we define a bias-adjusted trust aggregation. This formulation reflects the premise that reputational signals are driven by trustworthiness but attenuated by detected bias.

In parallel, we compute corpus-based baselines, including sentiment polarity (Alvarez-Morales et. al., 2024), embedding based bias indicators (Petreski and Hashim, 2022), and text-derived trust proxies (Huang and Liang, 2021). We assess agreement and systematic divergence between agentic and statistical measurements and analyze how these differences are present when aggregated into city-level reputation indicators.

Results

Preliminary results show alignment between agentic and corpus-based measurements, alongside systematic divergences in context-sensitive cases such as mixed-valence narratives and experiential descriptions tied to urban settings. Agentic AI tends to identify implicit bias and credibility cues that are weakly captured by sentiment or embedding-based metrics. When aggregated at the city level, these differences lead to distinct reputational profiles, suggesting that measurement choice directly affects how city reputation is inferred from review data.

Future Work

Future work will extend this framework to examine cross-platform such as Google Reviews, as well as and cross-city generalisability, for instance comparing Pittsburgh (a post-industrial city) and Berkeley (a college town). We also plan to explore how agentic explanations can support oversight of reputation systems. From a governance perspective, this research opens pathways for designing platform and urban analytics tools that combine scalable statistical monitoring with interpretable, context-aware AI assessments, supporting more accountable and socially grounded representations of city reputation.

References

- Alvarez-Morales, L., Solera, O., \& Travieso-González, C. M. (2024). Sentiment analysis methods, applications, and challenges: A systematic literature review. *Journal of King Saud University — Computer and Information Sciences*, 36(4), Article 102048.
- Bandi, A., Kongari, B., Naguru, R., Pasnoor, S., \& Vilipala, S. V. (2025). The rise of agentic AI: A review of definitions, frameworks, architectures, applications, evaluation metrics, and challenges. *Future Internet*, 17(9), Article 404.

- Guo, X., Pesonen, J., \& Komppula, R. (2021). Comparing online travel review platforms as destination image information agents. *Information Technology \& Tourism*, 23(2), 159–187.
- Huang, G., \& Liang, H. (2021). Uncovering the effects of textual features on trustworthiness of online consumer reviews: A computational-experimental approach. *Journal of Business Research*, 126, 1–11.
- Marine-Roig, E., \& Huertas, A. (2020). How safety affects destination image projected through online travel reviews. *Journal of Destination Marketing \& Management*, 18, 100469.
- Petreski, D., \& Hashim, I. C. (2022). Word embeddings are biased: But whose bias are they reflecting? *AI \& Society*.
- Yelp Open Dataset. (2025, January 15). Yelp for Business.
<https://business.yelp.com/data/resources/open-dataset/>
- Zelenka, J., Azubuike, T., \& Pásková, M. (2021). Trust model for online reviews of tourism services and evaluation of destinations. *Administrative Sciences*, 11(2), 34.

12.20 Constructing Masculinity in Cyberspace: Male Sex Worker Usernames as Performative Identities

Kristopher Jackson, Martha Whitfield (University of Massachusetts Dartmouth)

Introduction

The digital sex work economy has transformed the ways in which sex workers engage with clients and construct professional identities. Online advertisements serve not only as platforms for marketing and client acquisition but also as sites for the deliberate curation of identity. The rise of dedicated advertising platforms and mobile applications has both expanded access to clients and lowered barriers to entry into the profession, emphasizing the need to examine how male sex workers construct and perform their identities in digital contexts. In this environment, digital identity functions as a highly curated performance, offering a symbolic and purposeful, yet profoundly condensed representation of a worker's persona, services, and sexual availability.

While existing research has explored aspects of digital sex work, including pricing strategies, service negotiation, and platform use, little attention has been paid to the username, one of the most immediate and publicly visible markers of online identity. Usernames act as identifiers and function as the first point of contact between sex workers and potential clients. They are imbued with symbolic meaning, linguistic creativity, and cultural references that can communicate a range of professional, sexual, and emotional traits. By analyzing usernames, we can gain insight into how male sex workers who have sex with men (MSMSW) construct public personas, navigate client expectations, and negotiate the performance of masculinity and sexual availability in the digital marketplace. This study addresses this gap in the existing literature through a qualitative examination of 2,554 usernames, highlighting the symbolic and strategic elements of digital identity construction among MSMSW.

Objectives/Research Questions

The primary objective of this qualitative study was to conduct a thematic analysis of usernames adopted by MSMSW advertising services on a major internet advertising platform.

Methods

We compiled a dataset of 2,554 usernames from male sex workers who sell services to men on one of the largest internet platforms for commercial sex work in California. Data were collected at biweekly intervals over a four-month period (July–October 2024) using the ParseHub (Parsehub.com) web-scraping application. The dataset was organized

and cleaned to remove duplicates, inactive accounts, and incomplete profiles, resulting in a representative sample of publicly available usernames.

Usernames were selected as the unit of analysis because they represent the most condensed and immediate form of self-presentation in the digital sex work context. Each username reflects conscious choices regarding language, symbolism, and performative identity, offering insight into how MSMSW convey professional, sexual, and emotional aspects of themselves through a single digital marker.

Analysis and theme development were guided by Reflexive Thematic Analysis (RTA) as outlined by Braun and Clarke (2023), allowing for iterative exploration of patterns and meanings within the data. Our analytical approach was further informed by R.W. Connell's Theory of Hegemonic Masculinity (Connell, 2020; Connell & Messerschmidt, 2005) which provided a lens for understanding how participants construct personas that may reinforce, negotiate, or subvert dominant masculine norms. Through this framework, we considered how usernames encode cultural ideals of strength, sexual dominance, nurturance, and marginality, and how these ideals intersect with the occupational and emotional labor associated with commercial sex work.

To ensure methodological trustworthiness, we engaged reflexively with the data, held regular coding discussions, and iteratively refined our themes. This approach allowed us to remain closely aligned with participants' digital self-presentations while producing nuanced interpretive insights. Findings are reported in accordance with the Consolidated Criteria for Reporting Qualitative Research (COREQ) guidelines.

Results

Our analysis of 2,554 usernames revealed three core themes that illustrate the performative and strategic nature of online identity construction among MSMSW:

1. **Familiarity vs. Fantasy:** Usernames reflected a strategic oscillation between presenting an accessible, "boy next door" or familiar persona (e.g., using common names, professional titles, or demographic markers) and constructing an explicit, sometimes hyper-sexualized fantasy figure (e.g., referencing mythological figures, dominant roles, or extreme physical characteristics). This negotiation highlights the worker's effort to simultaneously cultivate trust and market an alluring, commodified persona.
2. **Dominance vs. Nurturance:** This theme captured the deliberate choice between usernames signaling authority, control, and aggressive masculinity versus those emphasizing care, intimacy, or nurturing traits. Findings highlight the wide spectrum of emotional and sexual labor that MSMSW advertise, often adapting to niche client demands that either reinforce or challenge rigid masculine roles.

3. Mainstream vs. Marginal: Usernames reflected a tension between aligning with conventionally "mainstream" or heteronormative masculinity (e.g., "straightness" references to sports/finance) and overtly referencing marginalized, subcultural, or fetishized sexual identities that exist outside typical societal norms. This theme underscores the delicate balance sex workers maintain between leveraging societal acceptance for broader appeal and catering to specific, niche clientele. Interpreted collectively through the lens of hegemonic masculinity, these themes demonstrate that MSMSW strategically construct personas that may superficially reinforce dominant masculine ideals, yet simultaneously subvert them to accommodate diverse sexual roles and emotional labor. These findings highlight the fluidity and complexity of identity in this digital marketplace.

Future Work

Findings from this study offer direct, actionable insights for public health initiatives and lay the groundwork for further research. For example, the themes identified could inform targeted digital outreach and health communication strategies. Messaging around PrEP uptake could be tailored to align with specific username themes (e.g., addressing the "Dominance" group with messages centered on control and self-protection, and the "Nurturance" group with messages emphasizing partner protection and care). Public health communication could leverage the language of the "Familiarity" persona to build trust and rapport, recognizing that simple, relatable language may be more effective than generic, clinical terminology. Conversely, understanding the "Marginal" personas represented in this dataset may help public health practitioners identify and reach sub-populations that are less visible, face greater stigma, and/or face greater barriers to care.

Future research might examine client-side perceptions of MSMSW usernames and digital personas. A follow-up study could analyze profile traffic patterns and client engagement in response to different usernames, as well as profiles with and without photos. Additionally, analyzing client reviews could provide insight into how usernames and self-presentation strategies influence client expectations, satisfaction, and interactions. Further work could also explore potential associations between online personas (usernames) and MSMSW engagement with health resources, risk behaviors, and broader occupational strategies.

References

- Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE.
- Connell, R. W. (2020). *Masculinities*. Routledge.

Connell, R. W., & Messerschmidt, J. W. (2005). Hegemonic masculinity: Rethinking the concept. *Gender & Society*, 19(6), 829–859.
<https://doi.org/10.1177/089124320527863>

12.21 I demand you cancel your cancellation! When politics meets entertainment: The effects of canceling a late night show host

Olivia Johnson (University of Houston)

Introduction

Political instability and unrest increasingly shape mediated public discourse, allowing politics to permeate almost every aspect of life including entertainment. In this context, media platforms become tools through which audiences communicate approval or dissent (Johnson et al., 2019) shaping viewing behavior, collective action, and online discourse. Such practices—often labeled as “cancellation”—are amplified through social media and frequently directed at high-profile and influential figures.

Recent events surrounding the killing of conservative political figure Charlie Kirk in September 2025 illustrate such phenomenon. On September 10, 2025, Kirk, the founder of TurningPoint USA, a non-profit that advocates for conservative politics, was killed during an outdoor student event at Utah Valley University. Many condemned the act of violence, but were critical of Kirk’s polarizing views and how conservatives were politicizing his death. Increased discourse and the silencing of media morphed this into a political consumerism movement. Political consumerism has traditionally been understood as either individualized lifestyle choices or collective actions such as boycotts and buycotts, yet existing research underestimates how deeply these practices have become embedded in everyday cultural consumption (Stolle & Micheletti, 2011). One commentary came from Jimmy Kimmel, a late-night talk show host, who discussed President Trump’s reaction to the killing of Charlie Kirk in his opening monologue on September 15, 2025. Subsequently, ABC announced the cancellation of the Jimmy Kimmel Live! Show indefinitely a couple of days later after pressure from the US Federal Communications Commission chair. After receiving backlash on social media using hashtags such as #cancelABC, #cancelDisney, ABC announced the reinstallation of Jimmy Kimmel Live! a week after their decision to cancel the show.

Some cases of cancel culture represent a digitally networked mode of political consumerism in which individual acts of withdrawing support (e.g., refusing to watch, subscribe, or endorse) are publicly articulated, morally framed, and amplified through social media. Entertainment figures and media brands increasingly function as political symbols, transforming acts of support or cancellation into routine expressions of political identity. Cancel culture reflects a hybrid form of political consumerism in which lifestyle politics and contentious connective action converge through networked

media platforms (Gotlieb & Cheema, 2017). In this way, cancel culture transforms individualized consumer choices into visible collective action without requiring formal organization. By analyzing both the top-down cancellation (i.e, pressured by FCC, initiated by ABC) and bottom-up consumer responses (i.e, consumer-initiated cancellation against broadcasters), this research advances understanding of how political instability reshapes the relationship between politics, media, and consumption.

Objectives/Research Questions

Efforts to remove shows, channels, or public figures from public visibility, as illustrated by the Jimmy Kimmel incident, are commonly framed as instances of cancel culture, understood as “the withdrawal of any kind of support (viewership, social media follows, purchases of products endorsed by the person, etc.) for those who have said or done something problematic, generally from a social justice perspective” (Ng 2020, p. 623). Some scholars consider the act of cancellation a form of partisanship-driven boycotting (i.e., political consumerism) that allows marginalized voices to be heard but also risks silencing opposing opinions. Others view consumers' demands to cancel an influential target as collective activism through which consumers collectively validate their feelings and experience (Traversa et al., 2023). The cancellation is not always initiated bottom-up (i.e., by concerning citizens) but can be top-down (i.e., state-initiated) (Li & Ng, 2024), in which the dynamic between the authorities, media, social media platforms, and the consumers plays an important role in the development and outcomes of the cancellation.

Cancel culture is a global occurrence shaping cultural consumption, accountability, and practices of restitution, largely driven by the connective power of digital platforms (Stec et al., 2023; Symons, 2022, Lo et al., 2025). These platforms facilitate the rapid diffusion of discourse, allowing fringe ideological content to migrate into mainstream discourse and turning social media into a contested space for competing values and identities (Van Dijck et al., 2018). The objective of this study is to explore public sentiment and discourse through the lens of cancel culture literature, responding to a politically charged entertainment event. More specifically, this study will examine the top-down cancellation of the Jimmy Kimmel Show in September 2025 and the bottom-up reactions to this incident (e.g., #cancelhulu, #canceldisney). Due to unrest political consumerism becomes even more important when there are so many instances particularly as they start to play out in the digitally networked environments.

Furthermore, this study seeks to understand not only consumer reactions but also media response prior to, during, and post the cancellation of Jimmy Kimmel's Late Night Show. We aim to identify the leaders, actors (hub and brokers), of the top-down cancellation and the bottom-up cancellation processes. The following research questions are proposed:

RQ1: How do top-down and bottom-up cancellation dynamics manifest in digital discourse surrounding the Jimmy Kimmel incident, and what roles do media institutions and consumers play in shaping cancellation outcomes?

RQ2: How does public sentiment evolve across pre-, during-, and post-cancellation phases of the Jimmy Kimmel incident, and to what extent do consumer-driven boycott movements (e.g., #cancelhulu, #canceldisney) reflect political consumerism versus collective activism?

Methods

To answer the research question, this study adopts a mixed-methods design integrating media content analysis, social network analysis, topic modeling, and sentiment analysis to examine public discourse surrounding the cancellation of late-night host Jimmy Kimmel's television show. Social network analysis (SNA) examines interactions and relationships within a social network (Borgatti et al., 2013), while content analysis, sentiment analysis, and topic modeling will be processed using natural language processing methods. We used the Access World News database to collect news media articles written about the event, from major U.S. media outlets, including CNN, NPR, and Fox News, which are the 2025 top US news outlets based on digital reputation and web metrics rankings (Scimago.com), regardless of political leanings and published between September 10–28, 2025. User generated content from social media users was collected from TikTok and Threads within the same time frame for messages with #jimmykimmel or #kimmel. Data was collected via two third-party AI web scrapers.

These two platforms allow the review of social media posts made on two distinctly different platforms. TikTok is designed for predominantly short-form videos, while Threads is designed for text-based message, similar to X. First, we will examine what users are saying about the event – determining what types of messages are part of this conversation. Using NodeXL, an open-source network visualization tool (Smith et al., 2010), we identified network characteristics for the social media dataset. Social network analysis was then used to identify clusters, (including actor hubs and brokers, frequently replied or reposted messages and their associated actors, in addition to the interactions between actors. Next, we will focus on the content and sentiment of these messages through a political consumerism lens, including the role of accompanying hashtags. Finally, we will analyze responses from major news media outlets and compare the topics surrounding the discourse to that of users online on social media.

Results

As a pilot data collection, we conducted a topic modeling of news articles collected from the New York Times using keyword “Jimmy Kimmel” during the predetermined timeframe. Of the 31 articles, the primary topic is centered on the freedom of speech and ABC's decision to cancel and to reinstall Jimmy Kimmel Live! Some other topics

emerged from the analysis show media's attention on President Trump's role in the top-down cancellation; public figures' criticism on the cancellation decision as well as on Kimmel; and the return of Kimmel's show. Interestingly, when comparing the topics that emerged from the articles published before Kimmel's return (i.e., 09/10/2025 to 09/22/2025, n=14) and those after his return (i.e., 09/23/2025-09/28/2025, n=17), it can be observed that topics shifted from focusing more on Charlie Kirk to Trump and that the topic around ABC's role in this incident and freedom of speech grew more prominent. The preliminary results, although only from one news outlet, showed the distinct dynamics in the top-down and the bottom-up cancellations, and the influences of key authorities and media in this process.

Future Work

The project which sits at the intersection of political consumerism, cancel culture, and digital activism, can lead to several future studies. First, the differences between top-down cancellation and bottom-up reactions need to be further explored in additional contexts. Particularly within digitally mediated networks, actors aim to voice their opinions in real time as consumers are increasingly using digital platforms to express their dissatisfaction. The Jimmy Kimmel incident is a great example of how the lines are blurred, denoting no clear line between social and political issues, which can turn a simple topic (or opinion) into a firestorm. Much of the existing literature on cancel culture focuses on political speech and media debates surrounding free expression.

However, relatively little research has examined cancel culture within the context of consumer boycotts. Additionally, this study seeks to provide a more nuanced understanding of cancel culture, as the term is often used interchangeably with related concepts such as call-out culture, digital shaming, and deplatforming.

For marketers, results from this study might be interesting for social listening. The conversations of discourse allow for a more realistic picture of who their customers are. That is, being able to identify performative actions versus authentic ones. Further, the findings from this study begin to fill the gap on cancel culture and how consumers can incite or prevent negative fallout. For example, one group of consumers can initiate this connective action, while another group is attempting to counter it; showing digital connections primarily made of similar actors with different objectives. Additionally, future work may also translate these insights into educational resources or industry-facing frameworks that help media professionals, brands, and civic organizations navigate public backlash, discourse polarization, and consumer activism in an increasingly politicized media environment.

References

Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2013). *Analyzing social networks*. London, England: SAGE

- Gotlieb, M. R., & Cheema, S. E. (2017). From consumer to producer: Motivations, internet use, and political consumerism. *Information, Communication & Society*, 20(4), 570–586. <https://doi.org/10.1080/1369118X.2016.1202301>
- Johnson, O., Hall-Phillips, A., Chung, T.-L. (Doreen), & Cho, H. (2019). Are you connected through consumption? The role of hashtags in political consumption. *Social Media + Society*, 5(4). <https://doi.org/10.1177/2056305119883427>
- Li, X., & Ng, E. (2025). “Block (封杀)!”: State-Netizen Constructions of Cancel Culture in China. *Television and New Media*, 26(1), 104–118. <https://doi.org/10.1177/15274764241277478;JOURNAL:JOURNAL:TVNA;PAGE:STRING:ARTICLE/CHAPTER>
- Lo, S. M., Araque, O., Sharma, R., & Stranisci, M. A. (2025, July). That is unacceptable: The moral foundations of canceling. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 6625-6639).
- Ng, E. (2020). No grand pronouncements here...: Reflections on cancel culture and digital media participation. *Television & new media*, 21(6), 621-627.
- Scimago Media Rankings. (n.d.). Media rankings: United States. Retrieved January 25, 2026, from <https://www.scimagomedia.com/rankings.php?country=United%20States>
- Smith, M., Ceni, A., Milic-Frayling, N., Shneiderman, B., Mendes Rodrigues, E., Leskovec, J., et al. (2010). NodeXL: A free and open network overview, discovery and exploration add-in for excel 2007/2010/2013/2016, from the social media research foundation. <https://www.smrfoundation.org>.
- Stec, P., & Jagielska-Burduk, A. (2023). Digital restitution of cultural goods: In search of a working model. *International Journal for the Semiotics of Law*, 36, 2207–2218.
- Stolle, D., & Micheletti, M. (2013). *Political consumerism: Global responsibility in action*. New York, NY: Cambridge University Press.
- Symons, A. (2022). *Women Comedians in the Digital Age: Media Work and Critical Reputations After Trump*. Routledge.
- Traversa, M., Tian, Y., & Wright, S. C. (2023). Cancel culture can be collectively validating for groups experiencing harm. *Frontiers in Psychology*, 14, 1181872. <https://doi.org/10.3389/FPSYG.2023.1181872/BIBTEX>
- van Dijck, J., Poell, T., & de Waal, M. (2018). *The Platform Society: Public Values in a Connective World*. Oxford University Press.

Zukin, C., Keeter, S., Andolina, M., Jenkins, K., & Carpini, M. X. D. (2006). A new engagement?: Political participation, civic life, and the changing American citizen. Oxford University Press.

12.22 Platform-Driven Right to Information in India: A Mixed-Methods Analysis of Politics and Participation

Shravani K, Neha Murthy, Maitreya Prafulla Chitale, Rajorshi Ray (International Institute of Information Technology Hyderabad)

Introduction

The Right to Information (RTI) Act, 2005, was established in India in response to citizens' growing demand for transparency. The Act empowered individuals to access governmental records and hold the public authorities accountable. However, over time, the implementation of this law has also produced troubling consequences. According to the data compiled by the Commonwealth Human Rights Initiative (CHRI), since 2006, there have been at least 109 murders, 187 physical assaults, 109 cases of harassment, and 7 suicides linked to RTI-related activities. Although The Whistleblowers Protection Act, 2014, was introduced to protect individuals who expose corruption, it has not been very effective in ensuring their safety.

One important reason for this risk lies in the structure of the RTI application process itself. Section 6(2) of the RTI Act states that applicants are not required to provide "other personal details except those that may be necessary for contacting him." In practice, however, even this limited disclosure exposes the requester's identity. This visibility can make the requester vulnerable to harassment, intimidation, or violence. Such social realities can be understood through the conceptual lens of The Political Information Order (Srinivasan, 2022), in which access to information is shaped by local power structures, class relations, and institutional authority. In this context, requesting official documents can challenge existing power hierarchies, placing the requester at risk of retaliation. As a result, a paradox has emerged: a law designed to empower citizens can also place them in significant danger.

To counter the vulnerability faced by individual RTI requesters, information collectives are being formed. These collectives, that operate across Indian states and act as intermediaries between citizens and the state by submitting RTI requests on behalf of the citizens. The collective becomes the RTI requester itself, thereby invisibilising the individual citizen. Citizens who rely on such intermediaries participate in what Benjamin (2000) describes as a Politics of Stealth. This concept refers to strategies through which marginalized or vulnerable actors pursue accountability while deliberately minimizing their visibility to state power.

These developments raise an important question about how citizens can safely exercise their right to information within a political environment, where acts of inquiry may invite retaliation. While the RTI Act formally guarantees access to state records,

the risks associated with filing requests have reshaped how citizens engage with this legal mechanism. Emerging digital intermediaries and information collectives, therefore, represent an important shift in the RTI landscape. By mediating the interaction between citizens and the state, such platforms may enable new forms of collective accountability while protecting individual requesters from direct exposure.

Objectives/Research Questions

To understand how this form of collective mediation operates in practice, this study examines a non-governmental organization referred to here as “Public-Forum.” Active in India for more than fifty years, the organization expanded its work in 2016 by developing a digital platform that we refer to as “RTI-raahi.” The platform allows citizens to submit RTI requests online without revealing their identities. NGO staff review and, if necessary, reframe these submissions offline before submitting them to government authorities. They cover the required fees themselves and send the requests via physical postal orders signed by an organizational representative. All requests and responses are archived in a public repository, creating a shared informational resource while avoiding the commercialization of user data.

This study examines the platform through the framework of Platform Cooperativism (Scholz, 2016). Platform cooperativism refers to digital platforms that are designed to serve collective social interests rather than profit-driven data extraction. By mediating the RTI process, RTI-raahi functions as an infrastructural shield that separates the identity of the citizen from the formal legal request.

Accordingly, this study asks the following research questions:

1. How does platform-enabled anonymity lead to a safer RTI landscape in India?
2. What role does platform cooperativism play in making the State accountable via RTI filings?

Methods

This study adopts a mixed-methods approach, combining situated ethnography with computational text analysis to examine patterns in India’s RTI ecosystem. The two methods were used iteratively: computational analysis of RTI records helped identify patterns and themes, which then informed interview questions and field observations. In turn, the ethnographic insights helped interpret the patterns observed in the digital data.

The ethnographic component included semi-structured interviews with five NGO personnel and 12 individual RTI activists; participant observation to understand both offline and online workflows within the NGO; and three focus group discussions to explore shared concerns and anxieties within the RTI community, including NGO employees, activists, and ordinary citizens.

To identify thematic patterns within approximately 13,000 RTI records, we applied unsupervised topic modeling techniques based on sentence embeddings. Requests were grouped using K-Means clustering (Sia et al., 2020) to identify semantically similar filings. To interpret the resulting clusters, we used a TF-IDF framework (Grootendorst, 2022) to extract the most informative keywords and phrases within each group. This process was further refined using a manually curated, domain-specific stopwords list derived through frequency analysis guided by Zipf's Law, which states that in natural language corpora a small number of words occur extremely frequently while the majority occur rarely. Using this distributional property, we identified highly frequent but semantically uninformative terms such as legal boilerplate and formal salutations and removed them to better surface the substantive intent of the requests.

To analyze the geographic focus of RTI requests, we developed a spatial pipeline that extracted locations associated with target government departments. Using Natural Language Processing (NLP), we applied Named Entity Recognition (NER) to identify Geopolitical Entities (GPE) within agency names and administrative jurisdictions in each filing. These locations were geocoded using OpenStreetMap and visualized through marker clusters and density heatmaps, allowing us to identify regional bureaucracies most frequently audited through RTI requests

Our data management followed a collaborative ethical protocol developed jointly between the researchers and the NGO to safeguard both participants and the organization. To prevent the exposure of vulnerable requesters, all sensitive identifiers, including individual names, contact details, and institutional affiliations, were anonymised before analysis. In addition, no raw personally identifiable information (PII) was stored on institutional servers. The data was processed through a secure local pipeline designed to maintain the anonymity of the original requesters at every stage of the research process.

Results

Our findings indicate that platforms such as RTI-raahi fundamentally reconfigure the materiality of the RTI process by bridging the digital-physical divide. The organization acts as an intermediary, keeping the citizen's identity anonymous. They do so by assuming the visibility and legal responsibility of the filing on behalf of the citizen. In this context, anonymity functions not merely as a privacy feature but enables citizens to exercise the politics of stealth, allowing them to safely engage in democratic participation.

Interviews with NGO personnel suggest that their collective structure and social capital position provide them a degree of protection from retaliation. However, we found that the risk is not entirely eliminated, with them receiving threatening phone calls in response to particular RTI filings, though these incidents did not escalate into in-person

harassment. These findings suggest that while the platform substantially mitigates the risk of direct bodily harm to individual citizens, the friction involved in challenging power structures does not disappear; rather, the risks associated are mediated through the organization.

Our computational analysis of the requests further illustrates how this cooperative structure shapes patterns of civic inquiry. Semantic clustering of the corpus shows that RTI requests concentrate around key domains of public governance. Many queries focus on welfare schemes and the distribution of social benefits, indicating that citizens use RTIs to monitor whether public programs reach intended beneficiaries. Other prominent themes include infrastructure and municipal services, public recruitment and employment, education governance, policing and legal processes, land administration, and public health systems. This analysis indicates that individual requests often reflect systemic concerns related to public interests.

Spatial analysis of the corpus demonstrates that RTI filings originate from multiple states and districts across India, even though Public-Forum operates from a single physical location and maintains no regional branches. This pattern shows how the digital interface of RTI-raahi enables geographically dispersed citizens to access the platform and participate in the RTI process.

Thus, we see that the design of RTI-raahi protects the anonymity of individual requesters and shapes the broader ecology of RTI use. The protective capacity of the platform is closely tied to its cooperative structure. Unlike models of platform capitalism that treat user data as a commodity, RTI-raahi operates on a non-monetized, philanthropic model oriented toward collective safety. By removing financial barriers to filing requests and prioritizing open access, the platform aligns with the principles of platform cooperativism. Such collectives, therefore illustrate how platform cooperativism can enable safer and more collectively oriented forms of democratic oversight within the RTI system.

Future Work

Future research will involve a more granular NLP-based analysis of the RTI requests to better understand the role of the human-in-the-loop in the reframing process carried out by NGO personnel. Examining how citizen-submitted queries are translated into formally structured RTI applications can help reveal how institutional mediation shapes the final form of civic inquiry.

We plan to analyze how request patterns evolve over time and across different user groups, we aim to better understand the changing relationship between digital citizenship and state engagement in India. A comparative analysis of cooperative and commercial RTI platforms operating in India would help clarify how different platform

governance models influence the nature of RTI filings, the issues raised by citizens, and the broader dynamics of public accountability.

Finally, future work will explore the policy implications of platform-enabled anonymity in RTI systems. In particular, it will examine how digital intermediaries and cooperative platforms might inform discussions on strengthening legal protections for RTI users and improving institutional safeguards for whistleblowers within the broader transparency framework. This includes examining how the forthcoming Digital Personal Data Protection (DPDP) Act, 2023, may affect the RTI landscape for both individual citizens and collectives, shaping the future of transparency and digital civic participation in India.

References

Government of India. (2005). The Right to Information Act, 2005 (No. 22 of 2005).

Commonwealth Human Rights Initiative. (n.d.). Attacks on RTI users.

<http://attacksonrtiusers.org/>

Srinivasan, J. (2022). The political lives of information: Information and the production of development in India. MIT Press.

Benjamin, S., & Bhuvaneswari, R. (2006). Urban futures of poor groups in Chennai and Bangalore. In *Local Governance in India* (pp. 221-267).

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794.

Nissenbaum, H. (2009). Privacy in context: Technology, policy, and the integrity of social life. Stanford University Press.

Nissenbaum, H. (1999). The meaning of anonymity in an information age. *The Information Society*, 15(2), 141–144.

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review*, 79(1), 119-157.

Scholz, T. (2016). Platform cooperativism: Challenging the corporate sharing economy. Rosa Luxemburg Stiftung.

Sia, S., Dalmia, A., & Mielke, S. J. (2020, November). Tired of topic models? Clusters of pretrained word embeddings make for fast and good topics too!. In *Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP)* (pp. 1728-1736).

Zhu, J., & Marjanovic, O. (2021). A different kind of sharing economy: A literature review of platform cooperatives. *Proceedings of the 54th Hawaii International Conference on System Sciences*.

12.23 Is “Chat” our friend? A study on conversational AI and parasocial relationships

Trevor Kartalija, Julia Belmiro (Drexel University)

Introduction

As of 2025, 62% of Americans indicated that they interact with AI at least a few times a week (Kennedy, et al., 2025), and according to OpenAI, 800 million people interact with ChatGPT every week (De Vynck & Merrill, 2025). People may interact with conversational AI platforms, for a variety of purposes—to gather news, get help with their homework, or to talk. In fact, the Washington Post claims that out of 47,000 conversations with ChatGPT, 13% of those were “musings and abstract discussion” with 10% being about emotional support, with users addressing the AI through nicknames (De Vynck & Merrill, 2025). Additionally, OpenAI estimates that one million of its weekly users show signs of emotional reliance (De Vynck & Merrill, 2025). Through users’ casual and emotional conversations with these chatbots, it is possible that they are forming strong one-sided bonds, or parasocial relationships (PSR), with these AI. The purpose of this study is to understand how individuals form PSRs with conversational AI such as ChatGPT, Claude, or any other conversational AI platform.

We intend to examine individuals’ self-reported interactions with AI using PSR Theory to examine what dimensions of PSR theory influence users’ relationships with AI. Particularly, we are interested in individuals’ attraction, perceived authenticity, perceived homophily, perceived trustworthiness, and emotional connection to/with the AI. These are all frequently investigated dimensions of PSRs in the context of various media personae, and our goal is to extend these to AI. We are also interested in users’ loneliness and introversion, and whether these factors influence PSR strength.

Current research has examined the relationship between technology attraction and PSI as well as PSR with AI influencers, AI service agents, and some other human-AI interactions. However, none have specifically investigated conversational AI. Similarly, few studies have looked at the dimensions we are proposing, such as attraction, authenticity, loneliness, introversion, or emotional connection in the context of PSRs and AI. This research also adds to existing literature on human-computer interaction which initially conceptualized emergent technologies as social actors, emphasizing that humans often respond to computers and digital media as if they were human social agents (Reeves & Nass, 1996). By building from this framework, the present study expands the social actor framework to conversational AI systems, positioning these chatbots as actors capable of sustaining interaction and even parasocial relationships.

We aim to expand the discussion of why individuals form emotional attachments and relational bonds with conversational AI systems, thereby enabling further examination of AI dependence, overreliance, and attachment.

Objectives/Research Questions

RQ1: How do individuals form PSRs with AI chatbots?

RQ2: What role does emotional perception play in the development of parasocial relationships?

RQ3: How does the level of trust affect the strength of parasocial relationships?

H1: Social attraction is positively related to PSR

H2: Task attraction is positively related to PSR

H3: Authenticity is positively related to PSR

H4: Homophily is positively related to PSR

H5: Loneliness mediates the relationship between PSR and social attraction (higher levels of loneliness will indicate higher social attraction, which will indicate higher PSR/ lower levels of loneliness indicate lower social attraction and lower PSR).

H6: Introversion is positively related to PSR

Methods

These research questions and hypotheses will be addressed by administering a survey to 400 participants. Using purposive sampling, the study will focus on selecting avid AI users, and participants will be recruited via Reddit, providing a large pool of potential subjects.

The survey will include established and reliable measures, with questions sourced from previous validated scales, such as Rubin et al.'s (1985) Parasocial Interaction Scale, McCroskey and McCain's (1974) measurement of interpersonal interaction to measure social attraction, adapted items from Zheng et al. (2020) to measure task attraction, and measures on perceived authenticity adapted from Moulard et al. (2015). Items from Zhang et al. (2021) and McCroskey et al. (2006) will be used to assess homophily. We will also use McCroskey and Teven (1999) scale on goodwill, adapted to assess competence and trustworthiness, as well as items from Dibble et al.'s (2012) Unidimensional Relationship Closeness Scale, the UCLA Loneliness scale, and Pittsburgh Introversion-Extroversion scale.

Results

Based on prior research and literature, this study anticipates that individuals will form parasocial relationships (PSRs) through their interactions with AI agents, such as ChatGPT, Claude, or any other conversational AI platform (Noor et al., 2022; Qi et al., 2025). Additionally, this study aims to examine the strength of PSRs and whether it is positively associated with social and task attraction (Zheng et al., 2020), perceived authenticity (Moulard et al., 2015), perceived homophily (McCroskey, 2006), and trustworthiness (McCroskey & Teven, 1999).

Our preliminary findings may anticipate that individuals who perceive their conversational AI partners as reliable, credible, or helpful may be more likely to report stronger emotional connections and more frequent engagement with the AI, consistent with prior parasocial relationship (PSR) literature on media figure and digital agents (Rubin et al., 1985; Reeves & Nass, 1996). We also anticipate that users who perceive conversational AI as emotionally responsive, supportive or empathetic will indicate stronger PSR, in accordance with previous literature that observed parasocial relationship formation and the fulfillment of emotional needs (Lotun et al., 2024).

Another variable of interest is loneliness, which we expect to mediate social attraction and parasocial relationship strength with conversational AI. We expect that individuals reporting higher levels of loneliness will experience greater social attraction toward their conversational AI, and consequently this increased social attraction will predict stronger parasocial relationships, that is, loneliness may drive people to seek social connection through conversational AI, which enhance the development of PSR. In contrast, individuals with lower levels of loneliness are expected to feel less socially engaged with AI chatbots, leading to weaker social attraction and lower PSR. In terms of our assessment on introversion, we expect it to be positively correlated with parasocial relationships. That is, introverted individuals may feel more prone to engage with AI chatbots and form parasocial relationships.

These anticipated findings extend the Parasocial Relationship Theory framework into conversational AI, indicating a novel approach, as most of PSR and AI studies remain unexplored as most current studies on PSR focuses on parasocial relationships with real influencer and Human Virtual Influencers (Stein et al., 2022), parasocial relationships and social media (Hoffner & Bond, 2022), and parasocial relationships and video game characters (Milman & Mills, 2023). The novelty of the study lies on the assumption that AI models can function as social actors capable of eliciting emotional attachment from its users, opening a discussion about relational bonds in the current era, as well as social support in an AI driven age.

Future Work

Based on our findings, future research should explore the depth and qualitative dimensions of parasocial relationships (PSRs) that users form with AI chatbots.

Although the current study assesses general emotional connection and social attraction toward conversational AI, these constructs do not fully capture the different relational experiences that may emerge from human-AI interactions. Therefore, future work could assess the different types of connections formed by these interactions, to examine whether users are engaging in deeply romantic connections or just developing friendships with AI chatbots.

In addition, future research could observe the functional role that these AI-mediated relationships play in users' lives, particularly in terms of social and emotional support, allowing one to assess whether AI chatbots can be a source of comfort, validation, or stress management. Similarly, it may be worth examining how individuals' relationships with conversational AI may be associated with particularly harmful effects, such as the recent phenomenon of "AI psychosis" (Hudon & Stip, 2025).

As human- AI interaction is a novel and evolving process, longitudinal studies could be useful to track how these relationships evolve over time, and whether AI usage increases or decreases throughout time and age differences. Such approaches could also clarify whether engagement with conversational AI becomes integrated into user's daily routine, or it is just a side interaction, and what these effects must show about social interaction.

References

- Dibble, J. L., Levine, T. R., & Park, H. S. (2012). The Unidimensional Relationship Closeness Scale (URCS): Reliability and validity evidence for a new measure of relationship closeness. *Psychological Assessment*, 24(3), 565–572.
<https://doi.org/10.1037/a0026265>
- De Vynck, G., & Merrill, J. B. (2025, November 12). We analyzed 47,000 ChatGPT conversations. Here's what people really use it for. *The Washington Post*.
<https://www.washingtonpost.com/technology/2025/11/12/how-people-use-chatgpt-data/>
- Hoffner, C. A., & Bond, B.J. (2022). Parasocial relationships, social media, & well-being. *Current opinion in psychology*, 45, 101306.
<https://doi.org/10.1016/j.copsyc.2022.101306>
- Hudon, A., & Stip, E. (2025). Delusional Experiences Emerging From AI Chatbot Interactions or "AI Psychosis." *JMIR Mental Health*, 12(1), e85799.
<https://doi.org/10.2196/85799>
- Kennedy, B., Yam, E., Kikuchi, E., Pula, I., & Fuentes, J. (2025, September 17). AI in Americans' lives: Awareness, experiences and attitudes. Pew Research Center.
<https://www.pewresearch.org/science/2025/09/17/ai-in-americans-lives-awareness-experiences-and-attitudes/>

- Lotun, S., Lamarche, V. M., Matran-Fernandez, A., & Sandstrom, G. M. (2024). People perceive parasocial relationships to be effective at fulfilling emotional needs. *Scientific reports*, 14(1), 8185. <https://doi.org/10.1038/s41598-024-58069-9>
- McCroskey, J. C., & McCain, T. A. (1974). The measurement of interpersonal attraction. *Speech Monographs*, 41(3), 261–266. <https://doi.org/10.1080/03637757409375845>
- McCroskey, L. L., McCroskey, J. C., & Richmond, V. P. (2006). Analysis and Improvement of the Measurement of Interpersonal Attraction and Homophily. *Communication Quarterly*, 54(1), 1–31. <https://doi.org/10.1080/01463370500270322>
- McCroskey, J. C., & Teven, J. J. (1999). Goodwill: A reexamination of the construct and its measurement. *Communication Monographs*, 66(1), 90–103. <https://doi.org/10.1080/03637759909376464>
- Milman, D., & Mills, D. (2023). The relationship between parasocial friendship quality with non-playable video game characters, gaming motivations, and obsessive vs harmonious passion. *Telematics and Informatics Reports*, 10, 100057. <https://doi.org/10.1016/j.teler.2023.100057>
- Moulard, J.G., Garrity, C.P. and Rice, D.H. (2015), What makes a human brand authentic? identifying the antecedents of celebrity authenticity. *Psychol. Mark.*, 32: 173-186. <https://doi.org/10.1002/mar.20771>
- Noor, N., Rao Hill, S., & Troshani, I. (2022). Artificial intelligence service agents: Role of parasocial relationship. *Journal of Computer Information Systems*, 62(5), 1009–1023. <https://doi.org/10.1080/08874417.2021.1962213>
- Qi, T., Liu, H., & Huang, Z. (2025). An assistant or A friend? The role of parasocial relationship of human-computer interaction. *Computers in Human Behavior*, 167. <https://doi.org/10.1016/j.chb.2025.108625>
- Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Center for the Study of Language and Information; Cambridge University Press.
- Rubin, A. M., Perse, E. M., & Powell, R. A. (1985). Loneliness, Parasocial Interaction, and Local Television News Viewing. *Human Communication Research*, 12(2), 155–180. <https://doi.org/10.1111/j.1468-2958.1985.tb00071.x>
- Stein, J.-P., Linda Breves, P., & Anders, N. (2022). Parasocial interactions with real and virtual influencers: The role of perceived similarity and human-likeness. *New*

Media & Society, 26(6), 3433-3453.

<https://doi.org/10.1177/14614448221102900> (Original work published 2024)

Zhang, K., Zhang, M., & Li, C. (2021). Effects of Celebrity Characteristics, Perceived Homophily, and Reverence on Consumer-Celebrity Para-Social Interaction and Brand Attitude. *Frontiers in Psychology*, 12.

<https://doi.org/10.3389/fpsyg.2021.711454>

Zheng, X., Men, J., Xiang, L., & Yang, F. (2020). Role of technology attraction and parasocial interaction in social shopping websites. *International Journal of Information Management*, 51, 102043.

<https://doi.org/10.1016/j.ijinfomgt.2019.102043>

12.24 Curating the Continent: Instagram and the Visual Promotion of African Tourism

Ammina Kothari (Simmons University) Sally Ann Cruikshank (Middle Tennessee State University)

Introduction

Social media is increasingly used by individuals, influencers, and tourism companies to inspire travel and drive engagement (Herzallah et al., 2022), with platforms like Instagram becoming conduits of social e-commerce (Yahia et al., 2018), where the goal is to entice followers to travel to promoted destinations. While the African continent contains a breathtaking tapestry of landscapes, diverse cultures, and rich agricultural traditions, traditional tourism narratives often portray it through a limited lens (Mutipforo et al., 2025). Stereotypical images of Africa prevalent in Western media, from the iconic acacia tree to persistent depictions of conflict and poverty, have been examined by scholars for decades (Garuba, 2002; McClintock, 1995) and have also been identified on social media. Hungerford et al. (2023) found that, despite being founded to challenge such stereotypes, the popular Instagram account @everydayafrica continued to perpetuate the “colonial gaze” (p. 1883). Smith (2018) similarly identified three motifs prevalent in African tourism posts: the tropical exotic, the promontory gaze, and fantasized assimilation. The reproduction of these motifs is “indicative of the unique embodiment afforded by Instagram, where tourists’ right to consume a destination and their privileged position within it may be more deeply inscribed” (Smith, 2018, p. 189).

Instagram therefore serves as a critical site for investigating how users leverage visual affordances, algorithmic functions, sociocultural contexts, and globalized commercial practices to represent African countries as tourist destinations. The platform hosts approximately 1.7 billion users worldwide, including more than 91 million across Africa (DataReportal, 2025). While Instagram has received criticism for manipulating algorithms in ways that encourage addictive use and amplify harmful content (Abrahms, 2022), it simultaneously facilitates what Arthur (2022) describes as “digital cultural bearing.” In this capacity, the platform enables the circulation of nuanced information about countries across the African continent, particularly among Black populations (Arthur, 2022).

Given these dynamics, analyses of tourism-focused Instagram accounts must also acknowledge the growing market of African social media influencers. This includes both established influencers and micro-influencers who collaborate with tourism companies and, in some cases, create niche digital spaces to showcase their countries

and cultures (Njathi & Lee, 2021). Their participation further complicates the visual representation of tourism-focused content on social media and raises important questions about how African destinations are visually framed and promoted within the platform's attention-driven economy. Accordingly, this study investigates how tourism-focused Instagram content constructs and circulates visual narratives about African travel destinations.

Objectives/Research Questions

This study aims to examine how tourism-focused Instagram accounts curate and circulate visual narratives about Africa. Despite growing scholarship on social media representations of the continent, relatively little research has systematically analyzed how tourism-focused accounts use platform features to promote travel to African destinations. To address this gap, the present study analyzes the use of African tourism-focused hashtags on Instagram and the visual content associated with them.

Given the persistence of stereotypical visual tropes about African countries on social media, as well as the emergence of generative AI and increasing pressure to capture potential clients' attention in a crowded social e-commerce landscape, this analysis seeks to understand how tourism-focused accounts leverage Instagram's marketing affordances to reframe the continent's tourism potential and promote more diverse, culturally rich, and environmentally conscious travel narratives.

RQ1: How are African tourism-focused hashtags used on Instagram to promote travel to destinations across the continent?

RQ2: What visual themes and narratives are most commonly represented in posts associated with African tourism-focused hashtags?

RQ3: To what extent do Instagram posts associated with African tourism reproduce or challenge stereotypical visual tropes historically associated with representations of Africa?

RQ4: How do tourism-focused Instagram accounts leverage the platform's visual and marketing affordances to frame African destinations as attractive travel experiences?

Methods

A mixed-method approach, previously employed in Instagram research to capture both quantitative trends and qualitative nuances (Reichert Smith & Sanderson, 2015), was adopted for this study. Researchers created a corpus using the first 50 posts returned by Instagram's search algorithm for each of the following hashtags: #africatourism, #africantourism, #africatravel, #travelinafrica, #travelingafrica, and #africantravel, resulting in a total sample of 300 posts.

The first 50 posts were selected because Instagram's algorithm prioritizes content based on recency, popularity, and engagement, making these posts the most visible to typical users and thus the most influential in shaping public perceptions of African tourism. Each post will be coded for indicators including account location, account type, source of visual content, and engagement metrics.

To ensure consistency and reliability, two researchers independently coded a subset of 30 posts, for inter-coder agreement for discrete data such as account type, follower count, account origin, type of content (video, stills, etc) with discrepancies discussed and resolved through consensus. Following the quantitative coding process, a textual analysis was conducted to examine the visual representation of African locations and individuals, as well as accompanying caption text and additional contextual elements such as secondary hashtags. Both researchers reviewed a subset of posts to establish criteria for textual analysis. This dual-method strategy enabled the identification of both measurable patterns and recurring thematic narratives within the dataset (Reichert Smith & Sanderson, 2015).

Results

Preliminary results show that when searching African tourism hashtags, the majority of results come from travel companies or influencers with substantial followers, generally more than 25,000. Very few personal accounts were returned in the hashtag search. Accounts often combine African focused hashtags with other tourism and travel concepts, including sustainable travel, hidden gem, and travel tips. One departure from previous media depictions of Africa is that accounts tended to focus on one country or place, including specific countries and locations, rather than just referring to general Africa.

Despite a wide variety of countries across the continent being represented in posts, the images and text used are fairly homogeneous. Videos and images tend to be a mix of big game, safaris, luxury resorts, and scenic vistas. The promotional messaging associated with these images usually follow a script of how Africa is a place to "slow down" and "appreciate nature." Smaller countries are regularly referred to as "hidden gems." For example, one travel influencer with 141,000 followers posted a reel showing various images of her enjoying amenities at a luxury resort interspersed with locals performing dancing shows with the caption "Africa has a way of slowing you down and waking you up, all at once... Africa is a feeling," along with the hashtag #thisisafrika.

Although the posts are largely positive depictions of Africa, rather than scenes of poverty or war, they do little to challenge traditional tourism stereotypes. When local people are included, they are most often seen in traditional clothing in rural spaces or doing something "exotic." One reel posted by a company called Rebrand Africa, which in its biography to be the "first platform dedicated to Africa's soft power," posted a reel

from Cameroon that featured a man swallowing fire. A safari company posted a 10-day itinerary for visiting Kenya, and all 10 days suggested something to do with nature and wildlife. The caption, however, touted, “Kenya is not what the world assumes... This country is beauty in contrast: endless savannah and modern skylines, ancient culture and rising innovation.”

In these posts, Africa’s image of a continent of poverty and conflict has been replaced by the opposite. Instead, Africa is depicted as a place to enjoy extreme luxury. The majority of travel companies and influencers featured high end resorts that have opulent rooms and fine dining. One reel highlighted the Chem Chem Lodge in Tanzania, which had been named one of the best resorts in the world by Condé Nast Travel. Rates for the Chem Chem Lodge begin at \$1420 a night per person, with a two night minimum. Another place often featured by travel companies and influences is Giraffe Manor in Kenya, where rooms can cost nearly \$5,000 a night. Finally, most of the popular accounts in the analyzed sample originated from non-African travel companies or influencers.

Future Work

The researchers are currently completing the analysis of the remaining data using established coding categories. Once the visual data analysis is finalized, we will conduct a thematic analysis of the associated promotional texts to examine how content creators leverage Instagram’s visual and marketing affordances to frame African destinations as attractive travel experiences. Our goal is to identify differences across Instagram account types, account origins, and the visual strategies used to promote tourism attractions in Africa.

Our preliminary analysis suggests that by promoting a narrow set of attractions such as wildlife safaris, luxury lodges, and expansive natural landscapes, tourism companies often present African destinations through a limited and homogenized lens. This emphasis risks erasing the cultural, social, and geographic diversity of African countries and can reinforce stereotypes that portray the continent primarily as a site of wilderness and exotic adventure rather than as a complex set of modern, culturally rich societies.

References

- Abrams, Z. (2021). How can we minimize Instagram’s harmful effects? *Monitor on Psychology*, 53(2). <https://www.apa.org/monitor/2022/03/feature-minimize-instagram-effects>
- Arthur, T. O. (2022). “We bring home the roots”: Black women travel influencers, digital culture bearing, and African internationalism in Instagram. *Social Media+ Society*, 8(2), 20563051221103843.

DataReportal. (2025, March 12). Instagram users, stats, data & trends for 2025.

<https://datareportal.com/essential-instagram-stats>

Garuba, H. (2002). Mapping the land/body/subject: Colonial and postcolonial geographies in African narrative. *Alternation*, 9(1), 87-116.

Herzallah, D., Munoz-Leiva, F., & Liebana-Cabanillas, F. (2022). Selling on Instagram: Factors that determine the adoption of Instagram commerce. *International Journal of Human-Computer Interaction*, 38(11), 1004-1022.

Hungerford, H., Subulwa, A.G., & Chakravarty, D. (2023). The Insta-Gaze: Investigating the endurance of stereotypes of Africa. *Social & Cultural Geography*, 24(10), 1883-1902.

McClintock, A. (1995). *Imperial Leather. Race, Gender, and Sexuality in the Colonial Contest*. Routledge.

Mutipforo, G., Muwani, T. S., Katsande, C., Ranganai, N., Denhere, P. T., & Mugijima, S. (2025). Data-Powered Social Media Marketing and Agritourism: A Catalyst for Rebranding African Tourism. In *Agritourism Marketing in Africa: Exploring Digital and Social Media Strategy* (pp. 287-318). Cham: Springer Nature Switzerland.

Njathi, A. W., & Lee, N. M. (2021). Social media influencers in Africa: an analysis of Instagram content and brand endorsements. In *Research perspectives on social media influencers and their followers* (pp. 91-110). Lexington Books, Bloomsbury Publishing Plc.

Reichart Smith, L., & Sanderson, J. (2015). I'm going to Instagram it! An analysis of athlete self-presentation on Instagram. *Journal of Broadcasting & Electronic Media*, 59(2), 342-358.

Smith, S. P. (2018). Instagram abroad: Performance, consumption and colonial narrative in tourism. *Postcolonial studies*, 21(2), 172-191.

Yahia, I. B., Al-Neama, N., & Kerbache, L. (2018). Investigating the drivers for social commerce in social media platforms: Importance of trust, social support and the platform perceived usage. *Journal of Retailing and Consumer Services*, 41, 11-19.

12.25 Redesigning Transparency Cues on Instagram for Agency

Manish Kumar, Ryan Moore (University of Texas Austin)

Introduction

Transparency cues on social media are interface signals that render algorithmic curation partially legible. Who is connected to whom, who did what, and why particular content is being surfaced (e.g., “Liked by...,” “Mutuals,” “People you may know”) (Marwick & boyd 2014; Oeldorf-Hirsch et al. 2025). Social media has made it easy to maintain weak ties – networks full of acquaintances, former classmates, and “friends-of-friends” (Harrison & Wellman 2012; boyd 2010). Platforms capitalize on these latent connections by deploying social transparency cues to activate interaction, and relational maintenance at scale (Dijck et al. 2018).

Social signals like “mutuals,” “liked by,” or “people you may know” can genuinely help people find communities of shared interest, reconnect with old friends, or feel socially anchored in large networks (Ellison et al., 2007). The problem is that these cues are increasingly optimized through an engagement and profitability first logic (Alaimo & Kallinikos 2019; National Academies of Sciences, Engineering, and Medicine, 2024), rather than designed through a humanistic lens that anticipates how visibility can produce interpersonal and embodied risk. When cues surface social context directly under content: making one’s engagement legible to others, they can enable unintentional identity disclosure (Marwick & boyd 2014; Vitak 2012). These risks are amplified for vulnerable groups (e.g., LGBTQ+ users) (Geeng & Hiniker 2021), since visible engagement can enable identity inferences about politics, sexuality, and affiliations (DellaPosta 2014).

Moreover, scholars have long pushed for greater explainability arguing that platforms should provide users with intelligible cues about how visibility is produced (Zhang & Chen 2020; Oeldorf-Hirsch et al. 2025; Haque et al. 2025). Rather than assuming transparency is normative or uniformly beneficial, we should use a contextual integrity perspective (Nissenbaum, 2011). Increased visibility does not automatically yield understanding or justice, and can redistribute burdens and risks onto those being made legible (Ananny & Crawford, 2018). Compounding this, platforms change these features so rapidly and opaquely, that the users have limited agency to fully understand, contest, or opt out of what new forms of social legibility entail (Fox & McEwan 2017).

Objectives/Research Questions

Prior work on likes focuses on self-esteem, social comparison, and status based on number of likes (Diefenbach & Anders 2022; Jaidka 2021; Noon 2024), but less on how visible likes shape perceptions of others and relational consequences. In this paper, I focus on Instagram because its follower network is overwhelmingly composed of known individuals (McClain, Rainie & Bell 2023). This means that transparency cues on Instagram are not directed into an abstract public but are posted into one's actual social network, where relational histories and interpersonal obligations are already in play. I specifically focus on these transparency cues:

1. Liked by ____
2. Commented by ____
3. Followed by ____
4. Liked by floating heads
5. Reshared by floating heads
6. 'Friends' tab on reels
7. Likes' List

Research questions:

1. How do users engage with these transparency cues in everyday life? How do these cues shape users' perceptions of each other and what are their behavioural and relational impacts?
2. Do these cues carry harsher implications for vulnerable users particularly LGBTQ+ users in non-majoritarian contexts where queer identities remain stigmatized and visibility can be consequential?
3. What strategies or design modifications could mitigate potential harms and enhance user agency over the privacy and relational risks produced by these cues?

Methods

Methodologically, I pursue this through a mixed approach combining a survey (n=300) and in-depth interviews (n=23). The survey was launched on the CloudResearch platform and it was aimed at adult Instagram users. The interview participants were recruited through the survey. The interview included questions about their daily engagement patterns, effect of cues on perception and behavioural adaptations. The interview also had a participatory design element aimed at generating alternative cue designs grounded in users' embodied constraints and privacy needs. Data were anonymized and analyzed using reflexive thematic analysis.

Results

I analyze participants' experiences through the lens of the social translucence framework (Erikson & Kellogg 2000). It has three components—visibility, awareness and accountability. Instagram's transparency cues do not show everything, but they

produce an abstracted visibility: partial signals that are still sufficient for users to notice patterns, infer interests, and draw conclusions about others. Most survey respondents (over 80%) reported being aware of these cues and noticing the kinds of content their friends engage with, suggesting that the interface routinely turns passive engagement into socially interpretable information.

At the level of awareness, participants described both connective and consequential interpretations. Some experienced positive effects, such as identity and affiliation: discovering shared tastes, recognizing “their kind of people,” and deepening weak ties through low-stakes cultural alignment. As one participant noted, “Seeing what my partner’s brother liked allowed me to tailor the memes I was sending him and now we send each other memes on Instagram.” Yet the same interpretability also produced judgment and discomfort. Participants reported reassessing others’ character, values, or wellbeing based on what they saw them liking. For example, worrying about a friend’s mental health after repeated engagement with nihilistic content, or changing how they viewed someone after seeing them interact with misogynistic posts.

Finally, accountability emerged once visible actions became subject to others’ interpretation. Around three-quarters of survey respondents described being conscious of leaving traces when liking, and many reported behavioral adaptations: saving posts instead of liking, sending content privately, liking less, or using alternate/dummy accounts. These accountability pressures were intensified for vulnerable users, particularly those who were not out about their sexual identities or who lived in geographies where queer identities are stigmatized. As one participant from south asia put it: “I am not out and to ensure my colleagues do not find out I avoid liking queer content.”

Taken together, these findings suggest that Instagram’s transparency cues do more than support social discovery. They also shape impression management, interpersonal judgment, and self-censorship, especially where networked visibility intersects with stigma and relational risk. The findings further point toward the value of exploring alternative cue designs that preserve connection while offering users greater control over how their engagement becomes legible to others.

Future Work

Building from the participatory design part of the interviews, I propose design modifications grounded in social-psychological concepts, including the privacy paradox (Barnes 2006) and the hyperpersonal model (Walther 1996).

1. Sensitive Content Classifier (Settings > Privacy): Users can tag topics, hashtags, or profiles as “Sensitive”. When they like content that matches these tags, the like remains in private history but is not surfaced publicly through transparency cues reducing involuntary disclosure while preserving personal curation.

2. Custom Interest Groups via long-press: On a long-press “Like”, users choose a circle (e.g., Friends, Book Club, Queer Community) to determine which audience can see that like. This makes social legibility reflect how users already mentally segment networks, aligning interface behavior with everyday impression management.

3. Temporary Visibility option: Users can make a like public for a limited window (e.g., 24 hours) after which it auto-archives into private history Or turn the cues off for 24 hours. This supports lightweight social signaling without forcing permanent exposure, and it reduces the long-tail risk of future reinterpretation or context collapse.

This work points toward redesigning transparency cues so that visibility becomes selective, situated, and user-controllable rather than a fixed default. Future research should evaluate these design directions through usability studies and related user-centered assessments. More broadly, this study offers an empirically grounded framework for rethinking transparency cues in ways that support agency and connection while mitigating privacy and relational harms. These findings also offer implications for social media platform design.

References

- Alaimo, C., & Kallinikos, J. (2019). Social media and the infrastructuring of sociality. In *Thinking infrastructures* (pp. 289-306). Emerald Publishing Limited.
- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *new media & society*, 20(3), 973-989.
- Barnes, S. B. (2006). A privacy paradox: Social networking in the United States. *First Monday*.
- boyd, d. (2010). Social network sites as networked publics: Affordances, dynamics, and implications. In *A networked self* (pp. 47-66). Routledge.
- DellaPosta, D., Shi, Y., & Macy, M. (2015). Why do liberals drink lattes?. *American Journal of Sociology*, 120(5), 1473-1511.
- Diefenbach, S., & Anders, L. (2022). The psychology of likes: Relevance of feedback on Instagram and relationship to self-esteem and social status. *Psychology of Popular Media*, 11(2), 196.
- Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook “friends:” Social capital and college students’ use of online social network sites. *Journal of computer-mediated communication*, 12(4), 1143-1168.
- Erickson, T., & Kellogg, W. A. (2000). Social translucence: an approach to designing systems that support social processes. *ACM transactions on computer-human interaction (TOCHI)*, 7(1), 59-83.

- Fox, J., & McEwan, B. (2017). Distinguishing technologies for social interaction: The perceived social affordances of communication channels scale. *Communication Monographs*, 84(3), 298-318.
- Geeng, C., & Hiniker, A. (2021). LGBTQ privacy concerns on social media. *arXiv preprint arXiv:2112.00107*.
- Haque, A. B., Islam, N., & Mikalef, P. (2025). To Explain or Not To Explain: An Empirical Investigation of AI-based Recommendations on Social Media Platforms: AKMB Haque et al. *Electronic Markets*, 35(1), 2
- Jaidka, K. (2021). The Well-being effects of Social Media differ across platforms and ethnicities: Evidence from the United States. Jaidka, K.(2022). Cross-platform- and subgroup-differences in the well-being effects of Twitter, Instagram, and Facebook in the United States. *Scientific reports*, 12(1), 1-11.
- Noon, E. J., Yang, C. C., Pesout, O., Stefanczyk, M. M., & Seiler, G. (2024). Insta-identity: The longitudinal reciprocal relationship between authentic and positive self-presentation on Instagram and the sexual identity development of sexual minority youth. *Computers in Human Behavior*, 158, 108278.
- Oeldorf-Hirsch, A., Romann, L. R., Witkowich, I., & Chen, J. (2025). Exploring social media users' experiences with algorithmic transparency cues. *new media & society*, 14614448251339493.
- Marwick, A. E., & boyd, d. (2014). Networked privacy: How teenagers negotiate context in social media. *New media & society*, 16(7), 1051-1067.
- McClain, C., Anderson, M., & Gelles-Watnick, R. (2024). How americans navigate politics on TikTok, X, Facebook and Instagram. *Pew Research Center*.
- National Academies of Sciences, Engineering, and Medicine. (2024). Social media and adolescent health (Chapter 2, "How social media work"). National Academies Press. <https://www.ncbi.nlm.nih.gov/books/NBK603437/>
- Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus*, 140(4), 32-48.
- Rainie, H., & Wellman, B. (2012). *Networked: The new social operating system* (Vol. 10). Cambridge, MA: MIT press.
- Van Dijck, J., Poell, T., & De Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford university press.
- Vitak, J. (2012). The impact of context collapse and privacy on social network site disclosures. *Journal of broadcasting & electronic media*, 56(4), 451-470.

- Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication research*, 23(1), 3-43.
- Zhang, Y., & Chen, X. (2020). Explainable recommendation: A survey and new perspectives. *Foundations and Trends® in Information Retrieval*, 14(1), 1-101.

12.26 Cultures of celebration around femicide perpetrators in Italy: a digital methods research

Valentina Lagati Sirino (Sapienza University of Rome)

Introduction

Recently, Italy has witnessed prominent cases of femicide, generating debate on social media. Online communities such as the manosphere and dark fandoms have increasingly celebrated perpetrators on TikTok and Instagram. This study examines such celebration, addressing femicide as a social fact (Durkheim, 1982) and situating its mediated representation within cultural platformization (Duffy et al., 2019). Existing studies have examined femicide perpetrators' portrayals in newspapers (Morrone & Capecchi, 2024; Dell'Anno, 2021; Saccà, 2021), overlooking social media communities' role. Newspapers emphasize perpetrators' relationships with victims, psychological states, jealousy, reputation, and raptus, portraying them as monstrous and framing femicide as a crime of passion, thereby reinforcing a psycho-individualistic rather than feminist perspective (Morrone & Capecchi, 2024).

Within the platformization of news, sensationalized portrayals reflect the attention economy; meanwhile, platformization facilitates practices through which users spread hate on social media. In femicide cases, comments often glorify perpetrators, reflecting misogynistic narratives promoted by dark fandoms and manosphere.

This dynamic is exemplified by incels' idolisation of Elliot Rodger on Reddit (Ging, 2019). Incels (involuntary celibates) belong to the manosphere, a network of online spaces where men discuss gender roles through misogyny and violence (Bainotti & Semenzin, 2021). While existing research has examined the manosphere's role in networked harassment (Hartmann, 2021), it has largely overlooked its role in celebrating femicide perpetrators.

The representation of killers as celebrities can be understood through Driessens' concept of celebrification (2013), addressing the transformation of ordinary individuals into celebrities through media rituals, alongside the long-standing sensationalization of crime reporting. Cultural platformization has introduced short-form videos on YouTube and TikTok, blending infotainment with journalism to present crime stories entertainingly (Shejuti, 2023). This trend is furthered by clickbait, algorithmic curation, and virality, intensifying the visibility of crime narratives (Lischka & Garz, 2023; Klinger & Svensson, 2015). Crucially, this process extends beyond traditional and platformized news media and is also evident in what Broll (2020) defines as dark fandoms, communities of fans who identify with or celebrate individuals who have committed heinous acts (Fathallah, 2023). Dark fandoms engage in fan production and online

interactions, warranting critical attention. Fathallah (2023) shows that dark fandoms often appropriate mainstream narratives, and this remains a key strategy, especially when fanworks reinforce dominant discourses, such as the focus on white, heterosexual male serial killers. Existing research has examined dark fandoms' celebration of serial killers, overlooking their role in celebrating femicide perpetrator.

Objectives/Research Questions

This study aims to investigate how online communities contribute to the construction of femicide perpetrators as celebrity-like figures through the circulation and commentary of crime news on social media, and thus to the normalization of femicide as a social fact. This research addresses a significant gap in existing literature, which has focused on the representation of femicide perpetrators in online newspapers articles, overlooking the crucial role played by social media communities in actively amplifying a killer's pseudo-celebrity status. It explores how such glorification contributes to the normalization of femicide as a social fact, as current literature tends to focus on broader forms of online gender-based violence without considering how specific portrayals of perpetrators may relate to femicide itself. Drawing on Durkheim's (1982) concept of social fact, this project treats femicide as a collective mode of action rooted in patriarchal beliefs that assign distinct roles to men and women, viewing women as men's property. These practices, transmitted through institutions, exert an external constraint on individuals by generating social expectations that transcend individual will (Durkheim, 1982).

Combining feminist research on femicide (Taylor & Jasinski, 2011) and practices of celebrification, the project examines the role of manosphere and dark fandoms in transforming femicide perpetrators into celebrity-like figures through the circulation and commentary of crime news on social media, addressing the following research question: How do online communities, fueled by the circulation of crime news on social media, contribute to the construction of a femicide perpetrator's celebrity status and the normalization of femicide as a social fact? The study is part of a broader project comprising four objectives. The present contribution focuses on two of them, presenting initial findings on the digital celebration of femicide perpetrators on social media.

The first objective is to investigate the dissemination of crime news involving femicide perpetrators by influential users on Italian social media, addressing the following sub-research question: How is the femicide perpetrator portrayed in the dissemination of Italian crime news on TikTok and Instagram?

The core of this study is to understand how online communities contribute to the glorification of femicide perpetrators, which addresses the second sub-research

question: What types of content do the manosphere and dark fandoms disseminate on TikTok and Instagram to celebrate the femicide perpetrator?

The social media considered in the study are TikTok and Instagram, as they are the fastest-growing platforms for news consumption globally (Newman et al., 2023), and increasingly rely on short-form video content, which shapes modern media habits (Newman et al., 2024). Moreover, they are also platforms where the manosphere is emerging, moving beyond more closed digital environments.

Methods

The project combines digital methods (Rogers, 2009) within a qualitative framework (Caliandro & Gandini, 2017) and traditional qualitative methods (Cardano, 2011). The present study presents initial findings on the digital celebration of femicide perpetrators on social media.

To answer the main research question, qualitative content analysis (Caliandro & Gandini, 2017) is conducted on Instagram Reels and TikTok videos published from the initial circulation of the news up to two months following the perpetrator's sentencing. Focusing on recent Italian femicide perpetrators - Filippo Turetta, Alessandro Impagnatiello, Stefano Argentino, and Mark Samson - a cross-platform analysis, divided into two phases, is performed on TikTok and Instagram videos.

The first phase explores how mainstream media accounts, digital news outlets, and true crime creators frame killers through narratives and cultural imaginaries. Accounts with more than 500,000 followers were considered, and among those covering these cases, the 24 most-liked videos were selected. An ethnographic content analysis (Altheide, 1987) is performed to generate theoretical insights through iterative reading, constructing categories aligned with the producers' conceptions. The approach examines the meanings users attribute to texts, their practical uses, and the cultural narratives emerging from these interactions. Furthermore, Critical Discourse Analysis is conducted on the 12 most commented videos from the dataset to examine narratives surrounding femicide perpetrators. Following Fairclough's (2013) three-step model, the analysis moves from content classification to pattern identification, reconstructing broader cultural imaginaries. It examines how selected accounts produce cohesive or conflicting representations, identifying hegemonic and counter-narratives. Finally, a narration analysis (Caliandro & Gandini, 2017) examines the comments on these same videos to explore users' positioning toward the femicide perpetrator, identifying comments linked to the manosphere and dark fandoms.

The second phase addresses the second sub-research question, investigating how the manosphere and dark fandoms draw on content analysed in the first phase to produce celebratory representations of perpetrators. Using hashtag-based sampling with perpetrators' and victims' names, the 36 most-liked videos were analyzed with

ethnographic content analysis to identify themes underpinning the celebrification of femicide perpetrators. In addition, a qualitative visual analysis (Caliandro & Gandini, 2017) examines how visual content contributes to celebrating the perpetrators, including video captions. Furthermore, a critical discourse analysis investigates the main narratives around the killers' glorification in the 18 most commented videos.

Finally, a narration analysis examines the comments on these same videos to understand how users situate themselves in relation to the femicide perpetrators' celebrification.

Results

This study provides an original contribution to the understanding of celebrity and fandom dynamics surrounding femicide in Italy by examining how social media communities celebrate perpetrators through the circulation and commentary of crime news, thereby contributing to the normalization of femicide as a social practice.

The first phase of the study demonstrates how crime news dissemination on platforms such as TikTok and Instagram often adopts sensational and dramatized narratives, shaped by algorithmic logics and the pursuit of virality. Ethnographic content analysis and critical discourse analysis reveal dominant portrayals of perpetrators as psychologically unstable, driven by passion, or constructed as monstrous figures, reflecting patterns commonly found in online newspaper reporting. At the same time, the study uncovers counter-narratives portraying the perpetrator as an ordinary individual. The narration analysis of user comments highlights multiple perspectives. Some users offer critical reflections on the perpetrator and broader critiques of patriarchy and institutional failures. Others, particularly within manosphere and dark fandom communities, express support for the perpetrator. This support contributes to the perpetrator's celebrification and the normalization of gender-based violence.

In this regard, the findings of the second phase, which combined ethnographic content analysis, critical discourse analysis, and qualitative visual analysis of user-generated content from these communities, reveal how the perpetrator is represented as both a hero and a video game character, thereby constructing him as a famous figure and a pseudo-celebrity. The analyses also highlight celebrification strategies that rely on dark humour, irony, and memes, which simultaneously mask and reinforce misogynistic attitudes and beliefs. These strategies also emerge from the narration analysis conducted on the comments under the videos, where users frequently employ irony to align with the content presented, thereby reinforcing and further amplifying the representations promoted in the videos.

These findings demonstrate how online communities rework online crime news through everyday digital practices, contributing to the celebrification of femicide perpetrators and the normalization of femicide and gender-based violence.

Future Work

Different directions emerge from this study. First, the research could be extended to YouTube Shorts in order to include another short-form video format similar to TikTok videos and Instagram Reels, thereby providing a more comprehensive overview of the phenomenon. Moreover, in Italy some podcasts have emerged as entertainment products in which participants spread misogynistic and anti-feminist discourse, such as Podcasterone (Semenzin, 2026). Analysing this type of content could help examine whether the celebration of perpetrators also extends to these environments and how it is articulated, as well as whether these dynamics differ from those observed in short-form videos. In addition, this study focuses only on young Italian femicide perpetrators who share a relatively similar social context. Therefore, future research could examine these dynamics across different types of perpetrators, taking into account variables such as ethnicity, social class, and age. Finally, further studies could also investigate the representation of the victim on social media, with particular attention to how the manosphere constructs and frames her within its narratives.

References

- Altheide, D. L. (1987). Reflections: Ethnographic content analysis. *Qualitative sociology*, 10(1), 65-77.
- Bainotti, L., & Semenzin, S. (2021). Donne tutte puttane: revenge porn e maschilità egemone. Durango.
- Broll, R. (2020). Dark Fandoms: An Introduction and Case Study. *Deviant Behavior*, 41(6), 792–804. <https://doi.org/10.1080/01639625.2019.1596453>
- Caliandro, A., Gandini, A. (2017). *Qualitative Research in Digital Environments: A research toolkit*. Routledge.
- Cardano, M. (2011). *La ricerca qualitativa*. Il Mulino.
- Corradi, C., Marcuello-Servós, C., Boira, S., & Weil, S. (2016). Theories of femicide and their significance for social research. *Current Sociology*, 64(7), 975–995. <https://doi.org/10.1177/0011392115622256>
- Dell'Anno, M. (2021). *Parole e pregiudizi: il linguaggio dei giornali italiani nei casi di femminicidio*. Città di Castello: LuoghInteriori.
- Driessens, O. (2013). Celebrity capital: redefining celebrity using field theory. *Theory and Society*, 42(5), 543–560. <https://doi.org/10.1007/s11186-013-9202-3>
- Duffy, B. E., Poell, T., & Nieborg, D. B. (2019). Platform practices in the cultural industries: Creativity, labor, and citizenship. *Social media+ society*, 5(4), 2056305119879672. <https://journals.sagepub.com/doi/full/10.1177/2056305119879672>

- Durkheim, É., Lukes, S., & Halls, W. D. (1982). *The rules of sociological method : and selected texts on sociology and its method*. Macmillan.
- Fairclough, N. & R. Wodak (1997): "Critical Discourse Analysis". In: T. A. van Dijk (ed.) (1997b): 258-284.
- Fathallah, J. (2023). *Killer Fandom: Fan Studies and the Celebrity Serial Killer*. mediastudies. press. <https://library.oapen.org/handle/20.500.12657/79407>
- Ging, D. (2019). Alphas, Betas, and Incels: Theorizing the Masculinities of the Manosphere. *Men and Masculinities*, 22(4), 638–657. <https://doi.org/10.1177/1097184X17706401>
- Hartmann, M. (2021). The totalizing meritocracy of heterosex: Subjectivity in NoFap. *Sexualities*, 24(3), 409-430. <https://doi.org/10.1177/1363460720932387>
- Klinger, U., & Svensson, J. (2015). Network media logic: Some conceptual considerations. In *The Routledge companion to social media and politics* (pp. 23-38). Routledge.
- Lischka, J. A., & Garz, M. (2023). Clickbait news and algorithmic curation: A game theory framework of the relation between journalism, users, and platforms. *New Media & Society*, 25(8), 2073–2094. <https://doi.org/10.1177/14614448211027174>
- Newman, N., Fletcher, R., Eddy., K., Robertson, C., Nielsen, R. (2023) *Digital News Report 2023*. Reuters Institute. <https://www.digitalnewsreport.org/>
- Newman, N., Fletcher, R., Robertson, C. T., Ross Arguedas, A., & Nielsen, R. K. (2024). *Reuters Institute digital news report 2024*. Reuters Institute for the study of Journalism. <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024>
- Morrone, E., & Capecchi, S. (2024). L'approccio giornalistico alla violenza maschile contro le donne in Italia. I casi studentesse-Camuffo/Costa, Pomarelli-Sebastiani, Bagnai-Genovese. *Problemi dell'informazione*, 3, 289–328. <https://www.rivisteweb.it/doi/10.1445/115448>
- Rogers, R. (2009). *The end of the virtual: Digital methods* (Vol. 339). Amsterdam University Press.
- Saccà, F. (2021). Stereotipo e pregiudizio: la rappresentazione giuridica e mediatica della violenza di genere. <https://digital.casalini.it/9788835124474>
- Semenzin, S. (2026). *Internet non è un posto per femmine*. Torino: Einaudi, Opera viva.
- Shejuti, K. K. (2023). Role of Social Media on Deviance and Crime: A Study on Content Creators of Tiktok. *American Journal of Multidisciplinary Research and Innovation*, 2(2), 130-137.

Taylor, R., & Jasinski, J. L. (2011). Femicide and the Feminist Perspective. *Homicide Studies*, 15(4), 341-362. <https://doi.org/10.1177/1088767911424541>

12.27 Trading Grammar for Likes: Linguists and Language Influencers in Czech Language Popularization

Helena Lázníčka (Charles University)

Introduction

In the last decade, a significant transformation has occurred in the landscape of public language discourse, particularly concerning orthography and written communication practices. While traditional linguistic authorities—academic linguists, language institutes, and prescriptive grammarians—have long shaped public discussions of language norms, a new category of actors has emerged within the digital media ecosystem: language influencers. These content creators combine educational objectives with entertainment value, using social media platforms to engage diverse audiences in discussions about language use, correctness, and variation.

The emergence of language influencers has not occurred without tensions. Traditional linguists have responded to this development with skepticism and, in some cases, criticism. Notably, Škrabal and Dufek (2024) introduced the pejorative term "jazykokupčení" (language profiteering) to characterize what they perceive as the commodification of linguistic knowledge by content creators who allegedly exploit language topics for personal gain rather than genuine educational purposes. This conceptualization reveals underlying anxieties about legitimacy, expertise, and the boundaries of linguistic authority in the digital age.

This study integrates language management theory (Neustupný & Nekvapil 2003), language ideology research (Milroy 2004, Kroskrity 2010), and language culture sociology (Kraus 2017) as its theoretical foundation. Language management theory provides tools for understanding how language-related problems are identified and addressed; language ideologies enable examination of beliefs and values that different actors associate with language practices and authority. Additionally, this research employs Ohanian's source credibility model (Ohanian 1990), which posits that communication effectiveness depends on three dimensions: credibility (expertise and trustworthiness), attractiveness (likability), and usefulness (practical value). This framework allows systematic comparison of how traditional and digital actors establish legitimacy through different combinations of these dimensions. This paper presents preliminary results from an ongoing doctoral project.

Objectives/Research Questions

The primary research objectives include: (1) examining how traditional linguists and language influencers conceptualize their respective roles and legitimacy in Czech

language popularization; (2) identifying the goals and intended outcomes that guide their communicative practices; (3) exploring how different actors define and attribute linguistic authority; and (4) analyzing how participants perceive the relationship between traditional and emergent forms of linguistic authority, including tensions such as those manifested in accusations of "jazykokupčení."

Methods

The study is based on semi-structured interviews with actors involved in Czech language popularization, including both academic linguists and language influencers active on social media. At the time of the conference, the analysis will draw on an initial set of interviews conducted within this ongoing doctoral project. Data analysis employs qualitative thematic analysis with systematic coding and identification of recurring themes.

Results

Preliminary findings suggest that whereas traditional linguists emphasize credibility through institutional affiliation and academic credentials, language influencers rely on a combination of perceived usefulness, personal attractiveness, and accessible communication styles that resonate with digital audiences. These findings indicate emerging differences in how linguistic authority is constructed and communicated across institutional and social media environments. The research reveals not merely a coexistence of different authority types, but an ongoing negotiation over the nature and boundaries of linguistic expertise in public discourse.

The study contributes to sociolinguistic discussions of how linguistic authority is being reconfigured in the digital age, offering insights into broader transformations in science communication and the diversification of channels through which expert knowledge reaches public audiences in contemporary Czech society.

Future Work

Future stages of the project will include additional interviews with both linguists and language influencers in order to expand the dataset and refine the analysis. The project will also explore how audiences perceive different forms of linguistic authority in online environments.

References

- KRAUS, Jiří (2017). Jazyková kultura. In CzechEncy–Nový encyklopedický slovník češtiny.
- KROSKRITY, Paul V. (2010). Language ideologies – Evolving perspectives. In: JASPERS, Jürgen, Jan-Ola ÖSTMAN and Jef VERSCHUEREN (eds.). *Society and Language Use*. Amsterdam: John Benjamins, pp. 192–211.

MILROY, Lesley (2004). Language ideologies and linguistic change. *Sociolinguistic variation: Critical reflections*, pp. 161–177.

NEUSTUPNÝ, Jiří V., NEKVAPIL, Jiří (2003). Language management in the Czech Republic. *Current Issues in Language Planning*, 4(3), pp. 181–366.

OHANIAN, Roobina (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising*, 19(3), pp. 39–52.

ŠKRABAL, Michal, DUFEK, Ondřej (2024). O jazykokupčení aneb Limity popularizace kodifikace. *Naše řeč*, 107(4–5), pp. 291–302.

12.28 Ethical Principles for Engineering Digital Services for Children

Beckett LeClair (5Rights Foundation) Maud Stiernet (A Little Lining Comes) Natalia Kucirkova (International Centre for EdTech Impact) Neil Gordon (University of Hull)

Introduction

While the digital environment continues to evolve at pace and scale, children and teens are disproportionately early adopters and increasingly “data subjects”. Digital technologies touch every aspect of life, but children’s rights and needs are often under-considered during design and development, their own views rarely sought or lent due weight.

There are numerous cases where this has led to direct harm to children and teens - from ‘AI’ chatbots providing fatal guidance, to insecure educational technologies leaking sensitive personal data en masse. Harms are currently systemic, though there has been comparatively little work to promote sufficient understanding of child rights by those who design and build these technologies. There are various reasons why harms could go left unmitigated and opportunities could go unexplored. Some organizations may see it as just another financial cost or could have political motivations. Others may have a lack of awareness or be unsure where to start. Many do not have regular contact with key community members such as children themselves, and they typically do not engage learning scientists or childhood studies experts. In most cases, there are likely interlinking reasons.

Meanwhile, the United Nations Convention on the Rights of the Child (United Nations, 1989) is the most widely ratified human rights treaty in the world, and General Comment No. 25 (United Nations, 2021) expands on this to explore how child rights apply in digital environments. In practice, the global landscape of definitions, policy, regulation and legislation is fragmented. The international arenas of both ethics and the technology industry are increasingly politically fraught, and misconceptions have proliferated. Our work rejects the notion that ethical behaviour is a barrier to industrial innovation - rather, it is a lens which can guide innovation in more beneficial directions. We see our definition of ethical principles as one that can exist among a plurality of other definitions, though our work is grounded in core ideas internationally agreed in fundamental rights instruments and related frameworks.

Objectives/Research Questions

The goal of this activity was to co-create a short, simple, and easily applied set of ethics principles for engineers and developers to use when designing and implementing digital

experiences for children, in alignment with the overarching ethos of the IEEE 2089 international standard on Age-Appropriate Design (IEEE, 2021). The working group contained representatives from across industry, academia and civil society.

We used the principles to create an oath usable by organisations and/or individual engineers as a starting point when considering children's rights and well-being in digital spaces, insofar as they can align their engineering practices to the ethical principles, or have an organization-wide commitment to adhere to them.

Our contributions seek to include the voices of a wide variety of impacted stakeholders across industry, academia, civil society and other experts, and especially children themselves.

Methods

Our project was made up of the following stages: a semi-structured literature review, youth engagement, expert engagement, distillation, reporting and reflection.

The literature review resulted in deeper analysis of 28 documents from a variety of sources including civil society reports, international standards, policy guidance, academic articles and so on. From this we extracted key themes common across the broader context and considered any gaps.

We then sought inputs from a wide range of youth, targeting participants aged 13 and up from all around the world. The methodology here was intentionally designed in collaboration with youth engagement experts to avoid leading participants to certain 'desirable' points of view or conclusions. Participants were given project background information and context as a primer, then asked to develop their own set of proposed ethical principles in a group. They were then given a condensed form of principles written by adult experts from the literature. We asked if they would amend any of their own principles in light of this, and how they would critique the perspectives of 'the adults' more generally. This resulted in a final set of principles they wished to move forward with. Participants were given multiple methods to choose from in terms of engagement, with the most popular being an online workshop using a Miro board. A post-session debrief was shared and participants were encouraged to challenge if they thought any conclusions were missing or misrepresented.

The next activity was to consult with a multi-disciplinary set of experts from industry, academia and civil society. We sought to understand their views on both the findings from youth engagement and the literature, and to hear their thoughts on related topics more generally.

Taking the outputs from all of these activities together but weighting the interests of children as our primary consideration, we distilled a finalised list of principles and created a short oath which can be used to commit to them. We remained in close

contact with the youth participants throughout to allow them to see and engage with the fruits of their work.

Results

Literature analysis uncovered a range of principles - many of which overlapped in underlying intent, though there was divergence in exactly how these principles should be expressed. This sometimes manifested in ways impacting operationalisation, such as the desire to protect autonomy manifesting as principles of 'do not engage in behavioural nudging' versus 'engage only in healthy behavioural nudging'. Common recurring themes included the need for organizations to consider child rights impacts, to follow age-appropriate design principles, to properly balance privacy and safety, to promote inclusion, to not manipulate children through design features, and to uphold opportunities for child participation.

Among youth, there was consensus on the importance of key topics such as privacy, inclusivity, age-appropriateness, freedom of expression, good health, safety, and youth participation. Many reflected on real experiences with services where they felt their rights were not respected. Their thoughts were often complex, with frequent recognition that one-size solutions are rarely sensitive to all their contexts or cultures. Participants generally thought services should be differentiated on the basis of user age, proposing principles broadly aligning to safety-by-design theory and to IEEE 2089. The full set of principles co-developed by the youth participants will be included in the final poster.

Concerningly, participants felt services often dismissed them, forgot they were there, or made reductive assumptions that worsened their experiences. They disliked what they saw as companies 'self-imposing' solutions without asking for youth input, or making 'hollow' gestures that only served their PR interests. Youth firmly believed they should have more opportunities to exercise their rights, especially their right to participate.

The core of the discussion with adult expert stakeholders focused on what the youth participants had created. The principles youth developed were broadly in line with what experts also advocated for in their work. Experts were pleasantly surprised by the youth's principles, using words like 'profound' and 'thorough' to describe them. The principles were seen to align well with some outcomes the EU seeks to achieve in digital regulation. Participants felt the youth principles were both relevant to engineers and feasible to act upon. There were no youth principles participants outright disagreed with.

From all of these insights we were able to refine the final set of principles and use this to create an oath (to be included in poster).

Future Work

Engineers likely try to follow many of the covered ethical principles to some extent already, but they lack the language or formal processes to articulate this. This could be framed as an incentive to use the outputs of this project, as it could highlight that positive change does not necessarily require complete overhauls of organizational culture. Another proposed framing of this project's outputs was as a tool to use on the route to sustainable innovation and more long-term business prospects more broadly.

Neither ethics nor child rights by design can be condensed into simple check-box exercises. In the same way, simply agreeing to the oath does not solve all existing problems. Sustained efforts are needed to address the pervasive scope and scale of online harms and missed opportunities for children, unconstructive workplace cultures positioning ethics as optional or anti-innovation, the systemic denial of children's right to be heard (among other rights), and the dearth of mandatory training on these topics for technical stakeholders. Many of these efforts will require coordination between a variety of stakeholder groups across industries, geographies and contexts. Notably, where children are impacted by this work directly or indirectly, they must be afforded a proportionate voice in the process.

As for this work in particular, our next steps will be to engage with industry groups, certification bodies, academic institutions and similar, to raise awareness of the principles and oath, to advocate for its inclusion into engineering curricula and industry body codes of conduct, and to ensure it remains a living project that can be maintained by experts (in collaboration with youth) in the future.

References

United Nations. (1989). Convention on the Rights of the Child. Retrieved from OHCHR website: <https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-child>

United Nations. (2021). OHCHR | General comment No. 25 (2021) on children's rights in relation to the digital environment. Retrieved from OHCHR website: <https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation>

IEEE. (2021). IEEE Standard for an Age Appropriate Digital Services Framework Based on the 5Rights Principles for Children. Retrieved from IEEE Standards Association website: <https://standards.ieee.org/ieee/2089/7633/>

12.29 Repair in Online Discourse: A Social Listening and Machine Learning Approach to Understanding Circular Economy Conversations

Barna Lipics, Pálma Mogyorósi, Zoltán Kmetty, Ágnes Buvár (Eötvös Loránd University Human-Environment Interactions Program)

Introduction

Extending product lifecycles through repair is widely recognized as a critical lever for the transition to a circular economy (Godfrey et al., 2022), yet a persistent gap exists between consumers' positive attitudes toward repair and their actual behavior (Jaeger-Erben et al., 2021; Ozanne & Madhuhansi, 2025). The scholarly literature on repair motivations commonly distinguishes between rational appeals—emphasizing functional, economic, or environmental utility—and emotional appeals, which frame objects as carriers of personal identity, memory, or sentimental value (Cherrier & Türe, 2020; Mugge et al., 2008). Recent experimental evidence by our research group (yet to be published) suggests that persuasive messages employing rational appeals can directly increase repair choices without necessarily generating more favorable attitudes toward repair in general. However, most existing studies rely on self-reported survey data or controlled experiments (Roskladka et al., 2023), limiting their ecological validity and capacity to capture the full spectrum of organically expressed motivations. Social media platforms and online communities have become key sites where consumers share repair experiences, seek advice, and articulate their motivations for—or against—repairing products. Online forums, comment sections, and social media groups dedicated to DIY repair, electronics troubleshooting, and sustainable consumption generate rich, naturalistic discourse that reflects genuine consumer decision-making processes unmediated by survey instruments. Unlike controlled experimental settings, these digital spaces capture the full complexity of how individuals negotiate repair decisions in everyday life—including the social dynamics of peer influence, community knowledge-sharing, and collective identity formation around repair practices. Despite this wealth of organic data, relatively few studies have applied computational methods to systematically analyze repair discourse at scale, leaving a significant gap in our understanding of how consumers express and negotiate repair motivations in their own words.

Objectives/Research Questions

This study leverages computational social media analysis to investigate how repair motivations manifest in naturally occurring online discourse. We aim to move beyond the limitations of survey-based and experimental approaches by analyzing large-scale,

multi-platform textual data using a combination of machine learning, large language models (LLMs), and unsupervised clustering techniques.

Our central research questions are:

- (1) What latent thematic structure underlies organic repair-related social media discourse?
- (2) To what extent do established motivational taxonomies—rational versus emotional appeals—reflect the drivers that people spontaneously articulate in online conversations?
- (3) What is the relative frequency and distribution of these motivational drivers across a large-scale, multi-platform corpus?

By addressing these questions, we seek to establish a data-driven taxonomy of repair motivations grounded in organic online discourse, complementing and extending existing classification schemes (Ozanne & Madhuhansi, 2025; Roskladka et al., 2023). We also aim to demonstrate the efficacy of a reproducible, multi-method computational pipeline for deriving psychologically meaningful constructs from large-scale, unstructured social media data, and to offer actionable implications for policy-makers and practitioners seeking to design targeted communication interventions that promote repair behavior through digital channels.

Methods

We collected approximately 30,000 user comments and 7,000 articles from a Hungarian multi-platform content aggregator spanning the past three years. The aggregator indexes content from multiple social media platforms, news sites, blogs, and online forums, providing a comprehensive view of the Hungarian-language digital public sphere. This social listening approach captures a rich diversity of product categories—from consumer electronics and household appliances to clothing and vehicles—as well as varied user demographics and conversational contexts. The dataset includes both long-form discussions and brief comment exchanges. All text analysis tools were customized for Hungarian, a morphologically complex agglutinative language that presents non-trivial challenges for standard NLP pipelines.

Our analytical pipeline integrates classical machine learning, LLM capabilities, and modern dimensionality reduction in a four-stage workflow:

Stage 1 – Relevance filtering: We trained a Support Vector Machine (SVM) classifier on a manually annotated subset of the corpus to pre-filter content, automatically distinguishing repair-relevant from irrelevant documents.

Stage 2 – Prompt optimization: We applied GEPA (Guided Evolution of Prompts and Assessments; Agrawal et al., 2025) to iteratively refine our LLM prompts using a

manually annotated evaluation set. This addressed the challenge of output variability in LLM-based content analysis, yielding substantially higher inter-annotator agreement between human coders and LLM-generated labels compared to naive prompting baselines.

Stage 3 – Motivation extraction: We deployed the optimized prompts within a hybrid framework combining locally hosted open-source and proprietary LLMs to extract specific user motivations from the filtered corpus, producing atomic, clean motivational labels suitable for quantitative analysis.

Stage 4 – Unsupervised clustering: We applied UMAP (McInnes et al., 2018) for dimensionality reduction coupled with HDBSCAN (McInnes et al., 2017) to identify latent thematic clusters, enabling data-driven discovery of motivational topics without imposing a priori categories.

Results

We analyzed 18,670 Hungarian-language online comments discussing mobile phone repair attitudes, collected from technology forums and social media platforms. Using our embedding-based clustering pipeline — comprising Gemini embeddings, PCA–UMAP dimensionality reduction, HDBSCAN clustering, and LLM-guided cluster consolidation — we identified distinct thematic groups across three open-ended response dimensions: 7 repair statement clusters, 10 perceived barrier clusters, and 22 perceived benefit clusters.

Preliminary findings indicate that economic considerations dominate repair-related online discourse. Consumers frequently weigh repair costs against device residual value, with device age and battery degradation serving as key decision factors. This is consistent with prior literature suggesting the primacy of rational appeals in repair decision-making, though the granularity afforded by our clustering approach reveals that "economic motivation" is not monolithic but comprises several distinct sub-themes (e.g., cost–benefit calculation, replacement timing, component-specific depreciation).

Warranty-related skepticism emerged as a pervasive barrier, with commenters expressing distrust toward both policy enforcement and authorized service center quality. This finding highlights a dimension of repair reluctance that is underrepresented in existing survey-based taxonomies, where warranty concerns are typically subsumed under broader institutional trust constructs.

Among perceived benefits, device longevity, warranty coverage, and value-for-money were the most prominent themes, indicating that consumers who do engage with repair primarily frame it through a financial lens. However, the 22-cluster granularity of the

benefits dimension suggests a richer motivational landscape than binary rational–emotional frameworks would predict.

Brand-specific frustrations — particularly regarding warranty denials and poor service communication by a specific manufacturer — constituted a distinct cluster, highlighting the role of manufacturer-level practices in shaping repair attitudes within online discourse. This finding underscores how social media serves as a space where consumers collectively process and amplify grievances against specific corporate practices, a dynamic that traditional survey methods are unlikely to capture.

Future Work

The current pipeline, while functional, has several limitations that point to concrete directions for improvement.

First, our LLM evaluation set was constructed primarily from one content type, yet our corpus comprises comments, forum posts, social media posts, and news articles with substantially different linguistic registers and structures. A more representative evaluation set that reflects this heterogeneity is needed to ensure that prompt optimization generalizes reliably across content types and that inter-annotator agreement metrics are not inflated by over-fitting to a single textual genre.

Second, named entities — brand names, product models, service center names, and geographic references — currently pollute the embedding space, causing semantically distinct motivations to cluster together when they share entity mentions, or splitting otherwise coherent motivational themes along brand-specific lines. Future iterations of the pipeline will need to implement entity masking or normalization prior to embedding generation to produce cleaner, motivation-driven cluster structures.

Third, the narrow topical focus of repair discourse presents a fundamental challenge for unsupervised clustering. Because the entire corpus revolves around a single domain, HDBSCAN tends to either produce overly broad clusters that conflate distinct motivational nuances, or generate highly granular clusters that are difficult to distinguish from one another substantively. Addressing this may require ontological augmentation — enriching the embedding input with structured knowledge about repair motivations, product categories, or consumer decision frameworks — to provide the clustering algorithm with additional semantic dimensions along which to differentiate otherwise similar texts.

Beyond pipeline refinements, we plan to reconcile our computational findings with the experimental results from our research group's ongoing work on persuasive repair messaging, enabling triangulation between what consumers spontaneously express online and how they respond to controlled interventions.

References

- Agrawal, L. A., Tan, S., Soylu, D., Ziemis, N., Khare, R., Opsahl-Ong, K., Singhvi, A., Shandilya, H., Ryan, M. J., Jiang, M., Potts, C., Sen, K., Dimakis, A. G., Stoica, I., Klein, D., Zaharia, M., & Khattab, O. (2025). GEPA: Reflective prompt evolution can outperform reinforcement learning. arXiv preprint arXiv:2507.19457.
- Cherrier, H., & Türe, M. (2020). Value dynamics in ordinary object disposal. *Journal of Business Research*, 116, 221–228.
- Godfrey, D. M., Price, L. L., & Lusch, R. F. (2022). Repair, Consumption, and Sustainability: Fixing Fragile Objects and Maintaining Consumer Practices. *Journal of Consumer Research*, 49(2), 229–251.
- Jaeger-Erben, M., Frick, V., & Hipp, T. (2021). Why do users (not) repair their devices? A study of the predictors of repair practices. *Journal of Cleaner Production*, 286, 125382.
- McCollough, J. (2009). Factors impacting the demand for repair services of household products: The disappearing repair trades and the throwaway society. *International Journal of Consumer Studies*, 33(6), 619–626.
- McInnes, L., Healy, J., & Astels, S. (2017). hdbscan: Hierarchical density based clustering. *Journal of Open Source Software*, 2(11), 205.
- McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for dimension reduction. arXiv preprint arXiv:1802.03426.
- Mugge, R., Schoormans, J. P. L., & Schifferstein, H. N. J. (2008). Product Attachment: Design Strategies to Stimulate the Emotional Bonding to Products. In *Product Experience* (pp. 425–440). Elsevier.
- Ozanne, L. K., & Madhuhansi, W. H. T. (2025). Consumer Repair Decision-Making: A Systematic Literature Review and Future Research Agenda. *International Journal of Consumer Studies*, 49(6), e70127.
- Roskladka, N., Jaegler, A., & Miragliotta, G. (2023). From “right to repair” to “willingness to repair”: Exploring consumer’s perspective to product lifecycle extension. *Journal of Cleaner Production*, 432, 139705.

12.30 “Read Me! Conceptualizing Social Media Attention Accumulation as Discursive Work”

Gaultier Lonqueu, Alette Lambert (University of Bath)

Introduction

Attention has become a central organizing principle of contemporary media environments. As news consumption increasingly shifts toward social media platforms, visibility, influence, and authority are no longer primarily structured by editorial hierarchies. Within this context, attention is commonly described as scarce: individuals face cognitive limits, information is abundant, and platforms are said to allocate visibility through opaque ranking and recommendation systems. Much existing scholarship has therefore explained attention outcomes primarily through platform design, algorithmic governance, or infrastructural bias (Bucher, 2012; Gillespie, 2015; Mears, 2023).

While these perspectives have generated important insights, they tend to attribute attention dynamics to platforms themselves and under-theorize the strategic role of high-visibility actors operating within these environments. Social media attention is not only distributed by algorithms; it is actively shaped by accounts that learn, iterate, and optimize their communicative practices in response to infrastructural signals. In particular, platform-native news actors – such as the depersonalized and highly visible News Aggregator Accounts (NAAs) like The Spectator Index on Twitter/X – have emerged as influential intermediaries in digital news ecosystems. These accounts do not merely compete for a pre-existing scarce resource; they participate in producing the conditions under which attention becomes urgent, repetitive, and normalized over time.

This article advances the argument that attention accumulation on social media should be understood as a discursive and strategic process, rather than as a passive byproduct of platform architecture. Through recurring practices of formatting, epistemic performance, high-velocity production, and audience orientation, large accounts actively shape how attention is captured, sustained, and legitimized. Algorithmic infrastructures do not originate these dynamics but amplify and stabilize them, rewarding practices that sustain speed, repetition, and engagement.

By reframing attention accumulation as the outcome of discursive work enacted at the account level, this article shifts the analytical focus from platform determinism toward the co-production of attention by platforms, content producers, and audiences. Doing so helps explain how a small number of actors come to dominate visibility in moments of crisis, why attention remains persistently unsatisfied despite informational

abundance, and how legitimacy and authority are increasingly performed through visibility rather than substantiated through verification. This perspective offers a framework for understanding attention accumulation as a reproducible process with important implications for public discourse and media power in platform-mediated environments.

Objectives/Research Questions

The central research question guiding the study is: how do accounts strategically accumulate attention through discourse on social media?

We pursue two interrelated objectives. First, we aim to identify the recurring discursive practices through which attention is captured, sustained, and stabilized in high-salience news contexts. Second, we seek to theorize how these practices work together and interact with platform infrastructures. By articulating attention accumulation as the outcome of strategic discursive work rather than platform allocation alone, the article contributes a framework for understanding how visibility, legitimacy, and media power are produced in contemporary social media news environments.

We therefore theorize attention accumulation not as a passive outcome of platform design, but as a strategic and reproducible product of discursive work enacted at the account level within algorithmically mediated environments.

Methods

To address the research question, we chose to focus on a high-salience, time-bounded case: the June 2025 Israel-Iran war. Beyond extensive mainstream coverage, The Spectator Index became a central actor on Twitter/X during this period, posting at very high frequency and achieving unusually high reach, making it a relevant site to examine how attention is accumulated through discourse in a platform setting. All tweets published by The Spectator Index between June 11th and June 29th, 2025, were collected to capture the full “issue-attention cycle” (Downs, 1972) around the event (from pre-conflict build-up to post-ceasefire decline). Data were scraped using Apify (API Dojo actor for Twitter/X scraping), exported to Excel for cleaning and chronological sorting, and filtered to retain only conflict-relevant posts. The final corpus comprised 821 tweets about the Israel-Iran war; engagement metrics reflect values as captured at the time of collection (July 1, 2025).

We conducted a Critical Discourse Analysis (CDA) to examine how linguistic choices relate to power and visibility in political communication, treating tweets as strategic resources deployed to shape salience. Analysis is anchored in Fairclough’s (1995) three-dimensional model of CDA to connect micro-features of language to production/circulation dynamics, and broader power relations. Analytically, the strategy was inductive and iterative. After familiarization and descriptive statistics in

Excel, the corpus was imported into NVivo 14. An initial open-coding phase on 30 tweets informed a textual codebook, which was consolidated into 38 themes, then abstracted into 16 discursive practices, organized in 8 strategies. From these, the three types of work were identified. The codebook was tested on additional batches (including randomly selected tweets) until thematic saturation was reached, with further targeted coding during write-up. Findings were iteratively interpreted to develop a model of how attention can be strategically accumulated through discourse. The analysis focused on written text only (excluding visuals).

Results

Preliminary findings conceptualize attention accumulation as a reproducible outcome of discursive work, with socio-political consequences discussed in our article. Three interrelated forms of work are identified: design work, operational work, and audience work. They are underpinned by a non-discursive dynamic of infrastructural amplification where the platform architecture (e.g., feed design, algorithmic ranking) magnifies the effects of each.

Design work refers to the process through which content is formatted and epistemically optimized to maximize immediate legibility and engagement. First, posts' formats are engineered to be short and predictable, enhancing scannability. These features reduce cognitive effort and allow posts to be processed instantaneously. Predictability further enables audiences to recognize posts from the account, reinforcing habitual attention. Second, epistemic engineering concerns how knowledge claims are performed rather than substantiated. Posts frequently activate affect while simultaneously performing neutrality and authority. Emotional activation increases salience, while neutrality, authority, and institutional performances signal credibility.

Operational work refers to the process of the account's content production. Event sequencing compresses time by breaking developments into discrete, rapidly successive updates, producing a sense of continuous escalation. This is also achieved through quantified continuity where updates are given on numerical details. These temporal dynamics transform the event into an unfolding stream that demands repeated attention. Production efficiency further supports this process. Reusable formats and a practice of a high-velocity output allow the account to respond quickly to developments while maintaining stylistic consistency. Interaction minimization – characterized by a non-dialogic structure – positions the account as a broadcaster rather than a participant, reinforcing authority while avoiding accountability. Verification offloading shifts epistemic responsibility outward. Claims are attributed to unnamed officials, institutions, or other media outlets, allowing the account to circulate information without owning its accuracy and benefitting from others' legitimacy. Some claims are also unsourced, presenting the events as facts to be taken for granted. This allows minimizing friction.

Audience Work captures how discursive practices orient audiences toward particular ways of attending to content over time and how legitimization occurs. Through status quo alignment, the account's reporting implicitly reinforces dominant geopolitical frames by often centering Western actors and institutions. This alignment is not overtly ideological but emerges through patterns of sourcing, framing, and emphasis. In parallel, ideological framing simplifies complex conflicts into binary oppositions, which facilitates comprehension and engagement while narrowing interpretive possibilities.

Future Work

First, the analysis focuses on a single high-salience geopolitical event and one prominent News Aggregator Account: further research could examine the transferability of the proposed model across different types of actors, issues, and platforms. Comparative studies could explore whether similar forms of design, operational, and audience work are enacted by persona-driven newsfluencers, institutional media outlets, or political actors, and how these practices vary across platform architectures such as Instagram, TikTok, or Telegram. Second, the present analysis is limited to textual content. Future work could extend this framework to multimodal communication by incorporating images, videos, and interface elements, which may play a growing role in, for instance, affective activation, scannability, and authority performance. Integrating visual and audiovisual analysis would allow for a more comprehensive account of how attention is accumulated in increasingly image- and video-centric environments. Third, while this article conceptualizes attention accumulation as a discursive and strategic process, future research could examine audience-side dynamics more directly. Combining discourse analysis with audience interviews, surveys, or experimental designs would help assess how audiences interpret, internalize, or resist the discursive patterns identified here, and how habituation to particular attention regimes develops over time. Finally, future work could further develop the policy implications of manufactured attention accumulation. If attention is actively produced and stabilized through strategic practices rather than merely allocated by platforms, interventions aimed at improving the quality of public discourse may need to move beyond platform-level regulation to address the incentive structures and visibility strategies that reward speed, repetition, and epistemic performance. Examining alternative discursive practices that sustain attention without reproducing scarcity may offer productive directions for both scholarly inquiry and media reform.

References

- Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, 14(7), 1164–1180.
<https://doi.org/10.1177/1461444812440159>

Downs, A. (1972). Up and down with ecology: The “issue-attention cycle.” *The Public Interest*, 28, 38-50.

<https://www.nationalaffairs.com/storage/app/uploads/public/58e/1a4/b56/58e1a4b56d25f917699992.pdf>

Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. Longman.

Gillespie, T. (2015). Platforms intervene. *Social Media + Society*, 1(1).

<https://doi.org/10.1177/2056305115580479>

Mears, A. (2023). Bringing Bourdieu to a content farm: Social media production fields and the cultural economy of attention. *Social Media + Society*.

<https://doi.org/10.1177/20563051231193027>

12.31 Understanding Ordinary Users of Algorithmic Recommendation Systems: A Digital Ethnographic Approach

Michelle Gay Nidoy (Queensland University of Technology)

Introduction

Research on social media users has predominantly focused on content creators and how they develop folk theories about algorithms to develop strategies to optimise their visibility (Bishop, 2018; Cotter, 2019). This focus on content creators leaves the vast majority of platform audiences, the ordinary users, largely unexamined. Yet understanding how everyday users experience, make sense of, and negotiate with algorithmic recommendation systems is crucial for examining cultural experiences on platforms such as TikTok, where algorithmic curation mediates cultural consumption.

This gap is particularly significant for understanding user agency in algorithmic recommendation systems because, unlike content creators who actively strategise visibility for monetisation, ordinary users engage with recommendation systems primarily as cultural consumers. Their agency manifests beyond acceptance or resistance, but is a nuanced relationship with the algorithm and moves through different forms of engagement (Siles, 2023).

Studying ordinary users' experiences with algorithmic recommendation systems poses distinct methodological challenges. How can researchers access users' subjective experiences of algorithmic recommendation systems? How do researchers understand agency in contexts where user behavior is constantly shaped by opaque algorithmic systems? Existing methodological approaches face limitations in addressing these challenges. Computational approaches like analysis of platform data could reveal behavioral patterns, but leave out users' subjective understandings, folk theories, or how they perceive their own agency. Relying solely on qualitative approaches, such as interviews or focus groups, would also be insufficient and would not capture how users respond to algorithmic recommendation systems. Traditional app walkthrough (Light et al., 2018) analyses app interface and features, but will not be able to capture the complexity of data flows and algorithmic personalization (Duguay & Gold-Apel, 2023).

Objectives/Research Questions

This study develops and tests a methodological approach for studying ordinary users' experiences with algorithmic recommendation systems. The central question guiding this work is:

How can researchers study user experiences with algorithmic recommendation systems, in relation to their agency and preferences?

The study uses TikTok as a case study to demonstrate this methodological approach, while also generating empirical insights about how ordinary users experience and navigate algorithmic curation.

Methods

An approach to studying users of algorithmic recommendation systems

This ongoing paper proposes a digital ethnographic approach for studying ordinary users' experiences with algorithmic recommendation systems, using TikTok as a case study. The approach responds to methodological challenges and integrates three complementary methods to address the limitations identified above: extended app walkthrough, semi-structured interviews, and participant data donations with visualization-based reflection.

Extended App Walkthrough

Building on Light et al.'s (2018) walkthrough method, this study employs an extended app walkthrough that goes beyond documenting interface features to develop a more grounded understanding of TikTok's recommendation system through sustained engagement. Unlike traditional walkthroughs that focus primarily on registration flows and feature analysis, this extended approach involves deliberate experimentation with the platform over an extended period. This extended app walkthrough runs from December 15, 2025 to March 15, 2026.

The researcher will engage with TikTok's For You Page with varying behaviors identified in prior algorithmic research as potentially significant, such as watch time duration, following a creator, engagement types (liking, sharing, commenting), and observes how the For You Page responds to these variations over time. The extended walkthrough serves as ethnographic preparation to build the grounded platform knowledge and will inform the formulation of questions for the interview about users' experiences with TikTok's recommendation system.

Semi-Structured Interviews

In-depth interviews explore ordinary users' folk theories, negotiation strategies, and sense-making practices about TikTok's recommendation system. Interview questions are informed by insights from the extended walkthrough and probe how users understand their feeds, what strategies they use to shape their experience, moments of surprise or frustration, and perceptions of their agency in shaping their feed. Interviews capture the subjective, interpretive dimension of user experience that behavioral data alone cannot reveal, how users make sense of an algorithmic recommendation system, what agency means to them, and how they navigate the platform in everyday practice.

Participant Data Donations

Participant data donations are integrated directly into interviews through visualization dashboards. Users share their TikTok data and have control over which data to share or delete. During interviews, these data, which may include their interaction patterns (content they watch, creators they follow, videos they posted and commented, etc.), are visualised through a dashboard and discussed alongside the participants. This integration addresses a key limitation in user research: the disconnect between how people think they use platforms and their actual usage patterns.

Results

These preliminary results cover the first twenty days of the extended app walkthrough (December 15-January 6) and reveal patterns about user-algorithm interaction on TikTok with implications for studying user agency.

Rapid but Opaque Algorithmic Learning

TikTok's algorithm learns user preferences remarkably fast, but the logic remains opaque. For instance, after liking a single video from one creator (@havea676) on the first day of the walkthrough, their content appeared 10 times within 30 minutes. While this may suggest that the algorithm heavily weights early engagement, patterns remain inconsistent. Some creators I engaged with appeared frequently while others appeared sporadically. Without clear feedback, users develop their own theories about what's happening. My assumptions evolved throughout, initially attributing patterns to watch time, then likes, then following. It remains unclear which signals most inform the algorithm about user preferences.

Limited and Inconsistent User Control

Users possess limited control over algorithmic curation. The "not interested" button worked inconsistently. It reduced some types of content but had little effect on others, particularly sponsored content. When clicking 'not interested' on ads, I tested the options 'irrelevant', 'excessive', and 'repetitive', yet the algorithm still served sponsored content eight times within 30 minutes in one session.

Whenever there is an attempt to diversify content, the algorithm takes time to adapt. Even when I deliberately searched for and engaged with different types of content, the algorithm kept showing me similar creators and topics I'd previously interacted with. New content would occasionally appear, but the feed quickly reverted to my earlier established preferences. This suggests that initial interaction patterns create constraints that are difficult to escape later.

Dynamic User-Algorithm Relationship Creates Tensions

The relationship between user and algorithm is fluid and contradictory. I enjoy personalization (content matching my preferences), while also seeking diversity and unfamiliar content. The algorithm appeared to prioritize personalization, creating frustration when I wanted discovery. This reflects a broader challenge where users depend on the system to discover content but lack visibility into how curation works or how to shape their experience. My preferences felt very much dependent on the algorithmic systems rather than entirely my own.

Methodological Implications

These preliminary insights illustrate the value of an extended app walkthrough for studying users of algorithmic recommendation systems. By developing grounded experiential knowledge, the walkthrough prepares researchers to ask informed questions and interpret responses, connecting and enriching the interview and data donation phases.

Future Work

Integration and Future Work

The walkthrough continues through March 15, 2026. Following its completion, a comprehensive account of the walkthrough will be developed to inform the development of the interview protocol. Participant recruitment (15-20 ordinary users) will begin after March 2026. The visualization dashboard is being developed for data discussion sessions. Analysis will triangulate walkthrough observations, interview narratives, and behavioral data to examine user agency in algorithmic environments.

References

- Bishop, S. (2019). Managing visibility on YouTube through algorithmic gossip. *New Media & Society*, 21(11–12), 2589–2606.
<https://doi.org/10.1177/1461444819854731>
- Cotter, K. (2019). Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram. *New Media & Society*, 21(4), 895–913.
<https://doi.org/10.1177/1461444818815684>
- Duguay, S., & Gold-Apel, H. (2023). Stumbling Blocks and Alternative Paths: Reconsidering the Walkthrough Method for Analyzing Apps. *Social Media + Society*, 9(1), 20563051231158822.
<https://doi.org/10.1177/20563051231158822>
- Light, B., Burgess, J., & Duguay, S. (2018). The walkthrough method: An approach to the study of apps. *New Media & Society*, 20(3), 881–900.
<https://doi.org/10.1177/1461444816675438>

Siles, I. (2023). Living with Algorithms: Agency and User Culture in Costa Rica. The MIT Press. <https://doi.org/10.7551/mitpress/14966.001.0001>

12.32 Digital Placemaking Between Strategy and Everyday Life: Negotiating Belonging in a Hybrid Urban Space

Giulia Nunes da Costa (Università della Svizzera Italiana)

Introduction

Digital media are increasingly embedded in the everyday production of urban space, shaping how places are navigated, imagined, and governed through platforms such as social media, mapping services, smart city dashboards, and municipal apps (Halegoua, 2020; Basaraba, 2023; Głowczyński, 2024). Scholarship on digital placemaking has expanded across urban studies, media studies, and human geography, but remains fragmented and often treats digital interventions as technical enhancements or design add-ons rather than as socio-spatial processes (Canelas & Hoehnk, 2025; Toland et al., 2020). This technologically-centred perspective overshadows how digital tools actively mediate human agency, social inequalities, and emotional connections to the city. Building on critical spatial theory of Lefebvre's (1991) and Massey's (2005), the paper understands placemaking as an ongoing negotiation of belonging, power, and membership in hybrid physical-digital environments

Digital placemaking can be defined as a mundane civic practice through which diverse residents negotiate belonging, conviviality, and difference in hybrid urban space, moving beyond a narrow focus on either marginalized groups or spectacular smart city projects (Brekhus, 1998; Alencar et al., 2025; Bennett et al., 2023). Research on digital placemaking has been done along two strands. One strand focuses on marginalized or otherwise marked groups and their usage of media (Alencar et al., 2025; Yu & Blain, 2019; Minchilli & Ponzanesi, 2025) to mediate everyday multicultural encounters and neighborhood conviviality (Bennett et al., 2023; Breck et al., 2020). While providing important insights, such studies tend to neglect relations among groups and places and overlook the complex socio-spatial dynamics, power imbalances, and continuous negotiations between diverse residents that shape a shared or divided and contested sense of belonging to a place (Atteneder and Lohmeier, 2025; Chen, 2025).

The other strand of research on digital placemaking explores top down practices such as strategic or property led initiatives driven by governments, planners, and developers using smart city platforms, loyalty apps, and dashboards to optimize services, brand cities, and stabilize coherent destination images. This casts residents as users, consumers, or data points rather than co-creators of urban meanings (Halegoua, 2020; Basaraba, 2021; Nafaji et al., 2024). Furthermore, critics warn that these instruments based approached frequently act as catalysts for gentrification and the commodification of public space, displacing marginalized communities while relying on surveillance and algorithmic control that replace authentic, messy and lively urban

life with sanitized, homogenized environments (Canelas & Hoehnk, 2025; Głowczyński, 2024; te Lintelo et al., 2024)

Objectives/Research Questions

Critics argue that these digital infrastructures privilege conceived space over lived space, curating urban experience through algorithmic visibility while rendering messy or contentious practices difficult to register (Delsante & Orlandi, 2020; Pinkster and Loomans, 2025). In contrast, studies of bottom up and platform native counter mapping highlight how residents appropriate social media, navigation apps, and “city walk” practices to contest official hierarchies, surface omitted spaces, and inscribe alternative narratives of belonging into urban environments (Xu and Chen, 2025). The tension between governance-led digital representations and residents’ practices of place attachment has not been fully explored. In particular, we are only beginning to understand how hybrid urban spaces are co-produced through uneven encounters between strategic logics of efficiency and competitiveness, and everyday efforts to claim, narrate, and inhabit the city (Atteneder & Lohmeier, 2025; Toland et al., 2020). Lastly, most existing research focuses on large metropolitan areas, leaving smaller cities and neighbourhoods comparatively underexamined.

To fully grasp the socio-political consequences of digital placemaking, empirical research must critically interrogate this intersection between how dominant technological frameworks structure spatial access and inequality, and how diverse communities appropriate these tools to negotiate their coexistence and cultural identities. Moving beyond the assumption that digital interventions naturally create and foster community, it is necessary to unpack the friction within this co-production. This project analyzes how top-down practices collide with the mundane ways residents create a sense of togetherness with everyday digital practices.

Further, we examine what and how membership categories and group relations emerge across both strategic and everyday digital interventions in the same urban setting.

To unpack these dual dynamics and address these aims, the study asks:

How do top-down strategic digital placemaking initiatives align with, reshape, or disrupt residents’ mundane digital spatial practices, and what forms of inclusion, exclusion, or relational inequality do they produce? How do everyday digital placemaking practices mediate the negotiation of belonging and conviviality across diverse urban populations, and what categories of membership and difference emerge through these practices?

Methods

The study is based on place ethnography and digital ethnography conducted in a small neighbourhood of a Swiss city. This neighbourhood is the most populous in the city (“Rapporto di Sostenibilità, 2024”) and is characterised by a high diversity of residents, including people with immigration backgrounds and Swiss nationals. The city itself is

located in a region close to the Italian border, which facilitates the presence of cross-border workers. It is also a city with considerable resources, offering well-developed infrastructure and a wide range of social and cultural opportunities.

Furthermore, the city explicitly embeds in the municipality's long-term development strategy digitalization and explicitly states its ambitions of becoming a "model of digital innovation" and a hub for crypto-finance ("Rapporto di Sostenibilità", 2024).

All these elements contribute to making the neighbourhood a unique site for research. Fieldwork combines participant observation in key neighborhood spaces and/or events, semi-structured interviews with residents from different demographic and socio-economic backgrounds, and document and interface analysis of municipal digital initiatives, including the city app, the eGov portal, and emerging projects associated with PlanB and the city's digital twin (Delsante & Orlandi, 2019; Pinkster and Loomans, 2025; Xu and Chen, 2025).

Participant observation in the neighbourhood examines how residents use media in everyday life, but also in site-specific ways, for example, to coordinate activities – mundane and not –, narrate local places, or navigate the neighborhood. This approach helps identify both convivial encounters and the subtle drawing of boundaries within the community. Building on these observations, the interviews that follow with the neighbourhood residents explore their perceptions of belonging, visibility, and recognition, as well as their media usage, and awareness, adoption, or rejection of municipal digital tools. Finally, an analysis of municipal strategies, reports, and digital interfaces traces how the city imagines its audiences, represents neighborhood space, and accounts for particular forms of participation and value in its digital infrastructures. Together, these methods develop a multi-scalar account of digital placemaking that connects everyday practices to institutional imaginaries and governance logics, while accounting for different digital realities.

Results

Preliminary analysis indicates that everyday digital placemaking in Molino Nuovo both fosters convivial encounters and reproduces relational inequalities. Residents use social media groups, messaging channels, and location-based platforms to share recommendations, coordinate childcare, informal services, and share stories, reinforcing familiarity and mutual support (Halegoua, 2020; Bennett et al., 2023). At the same time, these practices create selective digital publics that privilege certain languages, technical capacities, and established social ties, differentiating "insiders" and "outsiders" within the same streets (Brekhus, 1998; Breek et al., 2020). Early findings also show a partial misalignment between residents' mundane digital routines and Lugano's top-down digital placemaking initiatives, which foreground efficiency, innovation, and city branding (Basaraba, 2021). The municipal platforms analyzed, such

as the city app and its related smart city projects, tend to highlight consumption, mobility, and flagship events, stabilizing a coherent image of the city; but this leads to more precarious, contentious, or mundane uses of space less visible (Nafaji et al., 2024). This uneven intersection shapes who feels entitled to claim, narrate, and inhabit urban space, as some residents experience municipal platforms as irrelevant or exclusionary, while others are the ideal, digitally savvy urban subject (Delsante & Orlandi, 2019; Pinkster and Loomans, 2025; Xu and Chen, 2025). The results thus point to layered regimes of visibility, in which strategic and everyday digital placemaking intersect unevenly, and redistribute symbolic and practical access to the neighborhood.

Future Work

Future work will deepen the comparative analysis of resident groups and their positioning within Lugano's evolving digital governance infrastructure and their sense of place and belonging. This involves expanding the interview corpus and refining the coding of everyday practices in relation to municipal interfaces. Additional interface and policy analysis will track how new initiatives reconfigure categories of citizenship, participation, and value ("Rapporto di Sostenibilità", 2024; Nafaji et al., 2024). The paper will further dwell into the real of digital placemaking, detailing analytical dimensions such as visibility, addressability, and entitlement to narrate place that can be mobilized in other urban contexts (Atteneder & Lohmeier, 2025; Toland et al., 2020). Finally, subsequent research will explore potential interventions such as co-design processes, participatory mapping, or platform-native counter-mapping that might better align strategic digital initiatives with residents' everyday digital spatial practices and support more equitable forms of urban belonging (Basaraba, 2023; Xu and Chen, 2025)

References

- Alencar, A., Sarria Sanz, C., & Mawazini, J. (2025). Aspirational place-making and digital practices of refugees in the Netherlands. *Journal of Ethnic and Migration Studies*, 51(14), 3716-3733.
- Atteneder, H., & Lohmeier, C. (2025). Digital placemaking and its discontents: exploring practices, power relations, and socio-spatial dynamics in Salzburg's 'Andräviertel'. *Convergence*, 31(2), 634-653.
- Basaraba, N. (2023). The emergence of creative and digital place-making: A scoping review across disciplines. *New Media & Society*, 25(6), 1470-1497.
- Bennett, K., Gardner, Z., & De Sabbata, S. (2023). Digital geographies of everyday multiculturalism: 'Let's go Nando's!'. *Social & Cultural Geography*, 24(8), 1458-1477.

- Breek, P., Eshuis, J., & Hermes, J. (2021). Sharing feelings about neighborhood transformation on Facebook: online affective placemaking in Amsterdam-Noord. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, 14(2), 145-164.
- Brekhus, W. (1998). A sociology of the unmarked: Redirecting our focus. *Sociological theory*, 16(1), 34-51.
- Canelas, P., & Hoehnk, J. (2025). Practices of digital placemaking: accelerating place attachment. *Urban Research & Practice*, 1-19.
- Città di Lugano. (2024). Rapporto di Sostenibilità.
- Chen, P. (2025). Public spaces in Bologna: mediating social interactions by intercultural placemaking. *European Planning Studies*, 33(7), 1077-1103.
- Delsante, I., & Orlandi, S. (2020). Mapping Uses, People and Places: Towards a Counter-Cartography of Commoning Practices and Spaces for Commons. A Case Study in Pavia, Italy. *European Journal of Creative Practices in Cities and Landscapes*, 2(2), 121-150.
- Głowczyński, A. (2024). Digital representations of places: The role of spatial media in shaping urban experience. *Visual Communication Quarterly*, 32(1), 45-62.
- Halegoua, G. R. (2020). *The Digital City*. New York University Press.
- Lefebvre, H. (1991). The production of space. In *The people, place, and space reader* (pp. 289-293). Routledge.
- Lehner-Mear, R. (2020). Negotiating the ethics of Netnography: developing an ethical approach to an online study of mother perspectives. *International Journal of Social Research Methodology*, 23(2), 123-137.
- Minchilli, C., & Ponzanesi, S. (2025). Digital/postcolonial belongings. Migration, digital practices & the everyday. *Journal of Intercultural Studies*, 1-12
- Najafi, P., Mohammadi, M., Le Blanc, P. M., & Van Wesemael, P. (2024). Insights into placemaking, senior people, and digital technology: a systematic quantitative review. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, 17(4), 525-554.
- Pinkster, F. M., & Loomans, D. (2025). Urban belonging as place-based affect. *Social & Cultural Geography*, 26(4), 421-438.
- Russell, L., & Barley, R. (2020). Ethnography, ethics and ownership of data. *Ethnography*, 21(1), 5-25.

- te Lintelo, D. J., Ip, M. A., Lappi, T. R., Lakshman, R. W. D., Hemmersam, P., Dar, A., & Tervonen, M. (2024). Urban displacement and placemaking in public space for wellbeing: a systematic review of global literature. *Environment & Urbanization*, 36(2), 358-376.
- Toland, A., Foth, M., & Satchell, C. (2020). Digital placemaking: researching the co-production of hybrid urban places. *Media, Culture & Society*, 42(2), 173-191.
- Xu, D., & Chen, Y. (2025). City Walk as Platform-Native Counter-Mapping: Entangled Resistance and Algorithmic Visibility in Chinese Digital Urbanism. *Media and Communication*, 14.
- Yu, H., & Blain, H. (2019). Tongzhi on the move: Digital/social media and placemaking practices among young gay Chinese in Australia. *Media International Australia*, 173(1), 66-80.

12.33 Resisting medical misogynoir through digital counterpublics: Examining how Black Women use TikTok to support wellbeing

Chelsea Peterson-Salahuddin (University of Michigan Ann Arbor), Sadia Cheema (Seton Hall University), Ammina Kothari (Simmons University), Kelli Burns (University of South Florida), Kayla Hopper (Seton Hall University)

Introduction

Due to the history of systemic racism and sexism in the U.S. medical fields, Black women continue to experience adverse impacts on their well-being, or misogynoir in medicine (Agénor et al., 2019; Bailey, 2016; Chinn et al., 2021; S. E. Taylor, 2001). Due to a confluence of systemic cultural, social, and economic factors, Black women disproportionately experience chronic conditions (Banks, 2020; Chinn et al., 2021; Sternthal et al., 2011) and face higher rates of maternal mortality and adverse pregnancy outcomes (Lister et al., 2019; Nash, 2019). Further, experiencing racism and sexism not only negatively impacts Black women's mental health (Jones et al., 2021; Schulz et al., 2006), but Black women struggle to access proper mental health support due to lack of access, cultural stigma, and pervasive raced and gendered stereotypes, such as the trope of the "strong Black woman" that prohibit proper mental health diagnoses (Abrams et al., 2019; Liao et al., 2020; Ward et al., 2009).

Given U.S. Black women's adverse experiences within mainstream medical institutions, social media has become an important space where Black women can engage in discourse around their well-being. While studies have highlighted how social media platforms are still a space where Black women experience misogynoir-based harassment (Musgrave et al., 2022; Onuoha et al., 2024; Peterson-Salahuddin, 2022), these platforms also act as digital counterpublics where Black women support each other and resist racialized sexism in mainstream information sources, particularly through the creation of hashtags counterpublics (Clark, 2025; Peterson-Salahuddin, 2022; Z. A. Taylor & Abidin, 2024). For instance, Black women use Instagram to form online communities rooted in collective healing and push stereotypes that prevent their mental health needs from being met (Cameron, 2019, 2024). However, research on how Black women use digital platforms to support their well-being is still nascent.

Emerging research suggests TikTok may be a particularly apt space to build community and encourage dialogue around well-being. Studies have highlighted how TikTok's affordances, such as interactivity/sociality, algorithmic visibility through the platform's "For YouPage" (FYP) algorithm, and pleasure, support the creation of online communities around mental health information (Mordecai, 2023; Schluchter et al.,

2024; Song et al., 2021). However, no studies have examined how Black women specifically engage the affordances of TikTok to share information and build community around supporting their well-being.

Objectives/Research Questions

This study examines the culturally-specific way Black women engage the affordances of TikTok to build community around and share information about their well-being, to better understand how TikTok can be used to support health equity.

Following Evans et al. (2017), we position affordances as how the design or the perceived design of a technology allows specific types of relationships and interactions between users, as opposed to discrete features or outcomes of technology design.

Following previous research, we focus on how Black on TikTok engage affordances that have found to support online discourse about well-being: a) Sociality, or how digital features such as likes and comments enable social support, connection, and relationship building; b) Algorithmic visibility, or how digital features such as hashtags and features theorized by users to be favored by FYP, such as engaging with platform norms such as trends and captions (Herman, 2023; Schellewald, 2021) to enable information sharing; and c) Expressiveness, or how digital features such as emojis, editing tools, and modes of address allowed on TikTok user to express their identity.

To examine how Black women discuss well-being, we apply the optimal functioning in society (OFIS) model developed by Vallerand et al. (2013). Aligning with the World Health Organization's (1948) definition of overall well-being, the OFIS model encompasses five dimensions of well-being: (a) psychological well-being, (b) physical well-being, (c) social well-being, (d) performance in one's field, and (e) contributions to one's community and society. We use the OFIS framework to explore how Black women engage TikTok to discuss the multiple dimensions of well-being, focusing especially on how they leverage the platform to express culturally situated understanding.

The following research questions guided our analysis:

1. What TikTok features do U.S. Black women use to afford sociality, visibility, and expressiveness when discussing their well-being?
2. How do U.S. Black women on TikTok express their well-being through the dimensions of OFIS?

Methods

Sampling & Data Collection. We adopted a purposive sampling technique to collect TikTok videos created by Black women about well-being in January 2026. In Stage 1, the research assistant used the TikTok search features to find videos using the term 'Black women wellbeing,' and manually scraped the data for the first 50 videos, including the video link and pertinent metadata, such as the caption, creator nationality/place of

origin, and gender. This process was repeated for each of the search terms 'Black women mental health,' 'Black women physical health,' and 'Black women social health,' three dimensions of the OFIS framework, resulting in an initial sample of N=200 videos. In Stage 2, the research team developed an inclusion criterion to focus on unique videos that were a) in English, b) created by Black women or Black women-led organizations, and c) discussed well-being amongst U.S. Black women; after exclusions and duplicates, we had a final sample of N = 189 unique videos.

To analyze the data, we engaged in a Critical Technocultural Discourse Analysis (CDTA) that combined quantitative deductive and qualitative inductive analysis (Brock, 2018). CDTA aims to situate user practices and engagements with technological objects within the material design of technology and the socio-culturally mediated use of said technologies, making it particularly apt for thinking about the culturally specific way U.S. Black women use TikTok to discuss OFIS dimensions of well-being (Brock, 2018).

Analysis occurred in three rounds. In round one, each member of the research team independently reviewed each video to consider key affordances and themes in the data. Then the first and second authors developed a code book consisting of TikTok features used in videos that may be linked to the expression of one of the three primary affordances being examined. Codes were also deductively developed for well-being based on the OFIS framework. In round two, the first and second authors used the codebook to manually deductively code for what features creators used to engage the affordances (R1) and inductive thematic coding to examine how well-being was discussed in each video (R2) (Braun & Clarke, 2021). In round three, the first and second authors will engage in axial coding to examine the relationship between the affordances used and how Black women discuss the dimensions of well-being.

Results

Our preliminary results are based on an initial analysis of N=134 videos corresponding to the search terms 'Black women wellbeing,' 'Black women physical health,' and 'Black women social health.'

R1: To build online community around well-being, creators mainly engaged in sociality through liking (n = 105) and responding to comments (n = 89) on their videos to connect with their audience. Creators frequently used hashtags (n=119) to make their experiences algorithmically visible on the platform, helping to cultivate counterpublic communities centered on Black women's well-being. Finally, creators most commonly utilize the built-in record feature on TikTok to record a direct address-style video (N=39) or a video of events in their life narrated through a voice-over track (n=25), often paired with TikTok editing like text overlay (n=103) and emojis (n=64) to express their thoughts and point of view on well-being.

R2: Discussions of mental well-being often took the form of wellness advice to other Black women, often through direct address to camera videos. These videos focused on encouraging Black women to reject the idea that they needed to be “strong,” focusing on the need for Black women to rest, heal, and engage in self-love. Through the direct address video format, Black women were also able to engage the affordance of intimacy, creating the perception of a safe, online community for the Black women to discuss their mental health needs and challenges. When discussing physical well-being, creators engaged TikTok trends such as “get ready/unready with me” and “day in the life” to share their own fitness, clean eating, and relaxation routines, to encourage a healthy lifestyle amongst other Black women. Videos related to social health highlighted the need for Black women to support each other, both socially and professionally, and featured videos from Black women social workers and therapists with specific advice on how they could come together to support their mental health.

Future Work

Preliminary findings from this study highlight how Black women on TikTok draw on hashtags, comments, and comment liking to socialize, interact, and facilitate intimate discussions about Black women’s wellbeing on TikTok. Given that this study explicitly focused on U.S. Black women and TikTok, future research should examine how these findings map onto other platforms and Black diasporic populations. Additionally, as a content analysis, our findings are unable to make claims about how Black women experience engaging in these communities; future research should triangulate these findings using alternate methods that can speak to user experiences and perceptions.

References

- Abrams, J. A., Hill, A., & Maxwell, M. (2019). Underneath the mask of the strong black woman schema: Disentangling influences of strength and self-silencing on depressive symptoms among U.S. Black Women. *Sex Roles*, 80(9–10), 517–526. <https://doi.org/10.1007/s11199-018-0956-y>
- Agénor, M., Gill, T. G. R., Bailey, M., Benjamin, R., Charles, N., Dryden, O., Edu, U., Hammonds, E., Rosenkrantz, T. B. G., Harvey, S., Judd, B., Miles, A. L., Mobley, I. A., Oni-Orisan, A., Peoples, W., Roberts, D., Weiss, G. A., Schalk, S., & Washington, H. (2019). Open letter to editors of the *Journal of the National Medical Association* from the black feminist health science studies collective. *Journal of the National Medical Association*, 111(5), 573–575. <https://doi.org/10.1016/j.jnma.2019.04.003>
- Bailey, M. (2016). Misogynoir in medical media: On Caster Semenya and R. Kelly. *Catalyst: Feminism, Theory, Technoscience*, 2(2), 1–31. <https://doi.org/10.28968/cftt.v2i2.28800>

- Banks, N. (2020). Black Women in the United States and unpaid collective work: Theorizing the community as a site of production. *The Review of Black Political Economy*, 47(4), 343–362. <https://doi.org/10.1177/0034644620962811>
- Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE.
- Brock, A. (2018). Critical technocultural discourse analysis. *New Media & Society*, 20(3), 1012–1030. <https://doi.org/10.1177/1461444816677532>
- Cameron, S. J. (2019). Be still, Be present: Black girl yoga and digital counter spaces. *Race and Yoga*, 4(1). <https://doi.org/10.5070/R341042908>
- Cameron, S. J. (2024). “We out here”: Black Women, well-being, and #blackgirlsrun on Instagram. *Journal of Black Studies*, 55(5), 418–436. <https://doi.org/10.1177/00219347241244539>
- Chinn, J. J., Martin, I. K., & Redmond, N. (2021). Health equity among Black Women in the United States. *Journal of Women’s Health*, 30(2), 212–219. <https://doi.org/10.1089/jwh.2020.8868>
- Clark, M. D. (2025). Black women’s work. In *We Tried to Tell Y’All: Black Twitter and the Rise of Digital Counternarratives*. Oxford University Press.
- Evans, S. K., Pearce, K. E., Vitak, J., & Treem, J. W. (2017). Explicating affordances: A conceptual framework for understanding affordances in communication research. *Journal of Computer-Mediated Communication*, 22(1), 35–52. <https://doi.org/10.1111/jcc4.12180>
- Herman, L. (2023, October 9). For who page? TikTok creators’ algorithmic dependencies. IASDR 2023: Life-Changing Design. IASDR 2023: Life-Changing Design. <https://doi.org/10.21606/iasdr.2023.576>
- Jones, M. S., Womack, V., Jérémie-Brink, G., & Dickens, D. D. (2021). Gendered racism and mental health among young adult U.S. Black Women: The moderating roles of gendered racial identity centrality and identity shifting. *Sex Roles*, 85(3), 221–231. <https://doi.org/10.1007/s11199-020-01214-1>
- Liao, K. Y.-H., Wei, M., & Yin, M. (2020). The misunderstood schema of the strong Black Woman: Exploring its mental health consequences and coping responses among African American women. *Psychology of Women Quarterly*, 44(1), 84–104. <https://doi.org/10.1177/0361684319883198>
- Lister, R. L., Drake, W., Scott, B. H., & Graves, C. (2019). Black maternal mortality-the elephant in the room. *World Journal of Gynecology & Womens Health*, 3(1), 10.33552/wjgwh.2019.03.000555. <https://doi.org/10.33552/wjgwh.2019.03.000555>

- Mordecai, C. (2023). #anxiety: A multimodal discourse analysis of narrations of anxiety on TikTok. *Computers and Composition*, 67, 102763.
<https://doi.org/10.1016/j.compcom.2023.102763>
- Musgrave, T., Cummings, A., & Schoenebeck, S. (2022). Experiences of harm, healing, and joy among black women and femmes on social media. *CHI Conference on Human Factors in Computing Systems*, 1–17. <https://doi.org/10.1145/3491102.3517608>
- Nash, J. C. (2019). Birthing Black mothers: Birth work and the making of black maternal political subjects. *WSQ: Women's Studies Quarterly*, 47(3–4), 29–50.
<https://doi.org/10.1353/wsq.2019.0054>
- Onuoha, A. C., Matsuzaka, S., Stanton, A. G., Volpe, V. V., & Avery, L. R. (2024). Young Black American Women's social media use and online victimization. *Social Media + Society*, 10(3), 20563051241277607. <https://doi.org/10.1177/20563051241277607>
- Peterson-Salahuddin, C. (2022). Posting back: Exploring platformed Black feminist communities on Twitter and Instagram. *Social Media + Society*, 8(1), 20563051211069051. <https://doi.org/10.1177/20563051211069051>
- Pretorius, C., McCashin, D., & Coyle, D. (2022). Mental health professionals as influencers on TikTok and Instagram: What role do they play in mental health literacy and help-seeking? *Internet Interventions*, 30, 100591.
<https://doi.org/10.1016/j.invent.2022.100591>
- Schellewald, A. (2021). Communicative forms on TikTok: perspectives from digital ethnography. *International Journal of Communication*, 15(0), Article 0
- Schluchter, T., Ahmed, N., & Huang, E. (2024). “It Actually Affected My Relationship”: A qualitative analysis of affordances and attitudes towards mental health content on TikTok. *Proceedings of Mensch Und Computer 2024, MuC '24*, 491–507.
<https://doi.org/10.1145/3670653.3677501>
- Schulz, A. J., Gravlee, C. C., Williams, D. R., Israel, B. A., Mentz, G., & Rowe, Z. (2006). Discrimination, symptoms of depression, and self-rated health among African American women in Detroit: Results from a longitudinal analysis. *American Journal of Public Health*, 96(7), 1265–1270. <https://doi.org/10.2105/AJPH.2005.064543>
- Song, S., Zhao, Y. C., Yao, X., Ba, Z., & Zhu, Q. (2021). Short video apps as a health information source: An investigation of affordances, user experience, and users' intention to continue the use of TikTok. *Internet Research*, 31(6), 2120–2142.
<https://doi.org/10.1108/INTR-10-2020-0593>
- Sternthal, M. J., Slopen, N., & Williams, D. R. (2011). Racial disparities in health: How much does stress really matter? *Du Bois Review: Social Science Research on Race*, 8(1), 95–113. <https://doi.org/10.1017/S1742058X11000087>

Taylor, S. E. (2001). Health status of black women. In R. L. Braithwaite & S. E. Taylor (Eds.), *Health issues in the black community* (2nd ed., pp. 44–61). Jossey-Bass Publishers.

Taylor, Z. A., & Abidin, C. (2024). Where are all the Black girls on TikTok?: Exploring in-group community and (in)visibility through #BlackGirlTikTok. *International Journal of Cultural Studies*, 13678779241276751. <https://doi.org/10.1177/13678779241276751>

Ward, E. C., Clark, L. O., & Heidrich, S. (2009). African American Women's beliefs, coping behaviors, and barriers to seeking mental health services. *Qualitative Health Research*, 19(11), 1589–1601. <https://doi.org/10.1177/1049732309350686>

World Health Organization. (1948). *Constitution of the World Health Organization*. World Health Organization

12.34 From Embodied Ritual to Platform Performance: The Digital Transformation of Nuosu-Yi Bimo Ethnic Culture

Aguo Qubi (Southwest Minzu University) Hao Chen (University of Glasgow) Zhiwei Wang (Southwest Minzu University)

Introduction

This study investigates the digital transformation of Nuosu-Yi bimo practices during their migration to short-video platforms, focusing on how traditional embodied rituals are structurally reorganized within a platformized context. In the mountainous villages of Southwest China, the bimo serves as a ritual specialist and a mediator between humans and the spiritual realm. Traditionally, their authority is established upon family lineage, ritual efficacy, and local community recognition. As a form of folk religious practice, the rituals of the Nuosu-Yi bimo rely on sacred physical spaces, embodied co-presence, material mediums, and collective participation, constituting a specific cultural communication phenomenon.

However, as short-video platforms such as Kuaishou enter rural society, these ritual practices extend beyond their original physical boundaries into an algorithmically structured communicative environment. In this digital context, the visibility, interaction modalities, and dissemination pathways of rituals are influenced by platform recommendation systems, engagement metrics, and platform economic mechanisms. These platform infrastructures reorganize the temporal structure of the rituals, their sensory presentation, and the forms of audience participation, thereby generating a digital platform soundscape.

Mediatization theory posits that media logic systematically restructures religious institutions and ritual structures (Hjarvard, 2008). In the digital platform environment, this restructuring manifests specifically as the reallocation of visibility and authority governed by algorithmic logic. Relevant studies indicate that digital platforms are shifting religious legitimacy from traditional institutional accreditation to online visibility and interaction metrics (Campbell, 2021). Concurrently, sensory theory emphasizes that the sacred is constructed and practiced through sensory forms such as sound and vision (Meyer, 2006; 2020). We argue that when traditional rituals enter algorithm-driven short-video platforms, their sonic and visual forms are fragmented and recombined, forming a new "algorithmic soundscape" (Schafer, 1977), which subsequently transforms the sacred experience. We argue that this mediatization prompts a structural adaptation: complex physical rituals are encoded into streamlined auditory and visual performances, fundamentally reshaping the communicative role of the religious practitioner.

This study frames cultural digital transformation as structural reorganization, not mere preservation or dissolution. Focusing on Nuosu-Yi bimo rituals, it analyzes ritual functional shifts, practitioners' identity adaptation, and indigenous agency under platforms. Drawing on mediatization, sensory and platform theories, it argues digital platforms reconfigure ethnic ritual publicness.

Objectives/Research Questions

The bimo are ritual professionals who guide the spiritual life of the Nuosu community. In the Nuosu language, "bi" means to chant scriptures or host rituals, and "mo" refers to a person of knowledge or accomplishment. The bimo preside over ancestor worship, exorcisms, healing rituals, and various ceremonies. Traditionally, these practices are embedded in the daily life of the village and rely on embodied co-presence.

In the contemporary mediatized context, an increasing number of traditional practices rooted in everyday life are being digitally presented and recontextualized. This study focuses on the migration of bimo practices onto the Kuaishou platform, analyzing the transition from embodied rituals to platform performances. Specifically, this study addresses two research questions:

- RQ1: During digital platform performances, how are the communicative functions of rituals reconfigured across online and offline spaces?
- RQ2: How do Generation-Z bimo construct multidimensional digital identities amidst algorithmic visibility and context collapse?

By detailing how practitioners negotiate with platform infrastructures to maintain traditional authority and community connection, this research provides an empirical lens to understand how traditional ethnic cultures sustain cultural resilience in the digital age.

Methods

This study employed a one-year digital ethnography and a mixed-methods design to capture diverse bimo practices. The dataset includes approximately 60 short videos and 25 live streams, covering Yi autonomous counties and related urban contexts. Platform metadata (view counts, comments, likes, virtual gifts, shares, and follower growth trajectories) were collected to analyze visibility mechanisms and algorithmic circulation patterns. Screen recordings were utilized to analyze video formats, durations, audio-visual presentations, and interactive designs.

Fieldwork included participant observation during live streams and reflexive field diaries detailing researcher positionality and ethical considerations. Fifteen semi-structured interviews were completed, encompassing traditional elders, digitally active Gen-Z bimo, and young Yi users. Interviews were conducted in Nuosu languages, with

translation when necessary. Ethical procedures included informed consent, protection of sacred content, anonymization, and encrypted data storage.

Data analysis followed a three-stage coding process. Open coding identified key phenomena; axial coding linked ritual functions, sensory forms, and digital labor; and selective coding constructed an integrative narrative centered on spatial bifurcation and identity reorganization. NVivo was used for coding, and intercoder reliability tests showed high consistency (Cohen's kappa~0.7–0.8). Explanatory validity was enhanced through the triangulation of videos, live streams, comments, and interview data, with careful attention to platform-specific mechanisms and contextual limitations.

Results

Drawing on triangulated data from a one-year digital ethnography—encompassing the analysis of platform audio-visual content, interaction metadata, and multi-generational interviews—this study finds that the digital migration of bimo practices involves a spatial reconfiguration of ritual functions and a deliberate construction of multidimensional digital identities.

1. Spatial Reconfiguration and the Digital Soundscape

Platform mechanisms facilitate a functional division of rituals based on spatial and physical requirements. Interview data indicates a consensus among both senior and Gen-Z practitioners: online spaces are strictly utilized for auditory and consultative practices, such as divination (Kuisse). As a 63-year-old traditional practitioner explicitly noted, while complex rituals lose their efficacy without physical co-presence, divination remains viable through voice and video interactions. By foregrounding vocal consultation over physical manipulation, practitioners generate a specific digital platform soundscape.

This mediated soundscape transcends geographical limitations. A Gen-Z practitioner highlighted that online consultations provide accessible services to a dispersed Yi population, including migrant workers and urban college students, who can no longer access village-based rituals easily. Conversely, core rituals requiring material interaction remain anchored in offline environments. Consequently, the ritual is not digitized as a whole; rather, its communicative functions are spatially distributed based on platform affordances and traditional orthodox requirements.

2. Multidimensional Identity Construction

Building on this spatial division, practitioners must navigate the tension between algorithmic visibility and context collapse. Gen-Z bimo construct a composite digital identity to maintain legitimacy and cultural boundaries:

- **Traditional Anchor:** Algorithm-related metrics do not determine religious authority, which is largely grounded in family lineage and offline practice. To avoid context

collapse and secularization, practitioners highlight their ancestral backgrounds and exhibit ritual artifacts during live streams. Younger generations also avoid mainstream music and visual filters to maintain the solemnity of the practice.

- **Normative Mediator:** Practitioners balance platform monetization with community norms. While some older practitioners express concerns about commercialization, Gen-Z practitioners interpret virtual gifts as modern forms of ritual offerings. Online income is often reinvested in offline cultural inheritance such as scriptures and ritual instruments.
- **Cultural Transmitter:** The digital shift shows intergenerational change in practice orientation. Unlike elders serving local communities, younger practitioners use short videos and live streams for cultural education, including Yi language and ritual knowledge, to support cultural identity among urban Yi youth and promote wider cultural transmission.

Future Work

Future research will further explore how online visibility impacts offline authority structures, analyzing whether algorithmic traffic alters the hierarchies of legitimacy within villages. Cross-platform comparisons (e.g., Douyin, WeChat) will be utilized to study the differential impacts of various recommendation mechanisms and platform architectures on ritual transformation.

Longitudinal studies will focus on intergenerational shifts, examining how digital-native youth internalize and reinterpret ethnic ritual practices. Furthermore, the tension between spectacle-driven visibility and ritual integrity will be analyzed to construct a more systematic theoretical framework for understanding the negotiation pathways of indigenous cultures in the global digital platform environment.

References

- Hjarvard, S. (2008). The mediatization of religion: A theory of the media as agents of religious change. *Northern Lights: Film & Media Studies Yearbook*, 6(1), 9–26.
https://doi.org/10.1386/nl.6.1.9_1
- Campbell, H. A., & Tsuria, R. (2021). *Digital religion: Understanding religious practice in digital media*. Routledge.
- Meyer, B. (2009). *Aesthetic formations: Media, religion, and the senses*. Springer.
<https://doi.org/10.1057/9780230221374>
- Meyer, B. (2013). Mediation and immediacy: Sensational forms, semiotic ideologies, and the question of the medium. In J. S. Bielo (Ed.), *A companion to the anthropology of religion* (pp. 309–326). John Wiley & Sons, Ltd.

Schafer, R. M. (1977). The soundscape: Our sonic environment and the tuning of the world. Destiny Books.

12.35 Researching suicide and social media after the Online Safety Act

Aguo Qubi (Southwest Minzu University) Hao Chen (University of Glasgow) Zhiwei Wang (Southwest Minzu University)

Introduction

The UK's Online Safety Act (OSA), passed in October 2023, introduced age verification requirements for individuals seeking access to adult content online. This legislation extends to online content related to pornography, suicide, and self-harm, and requires social media platforms such as Reddit and X to implement age verification mechanisms - often involving the submission of photographs or government ID, such as passports and driving licenses. Social media platforms often state that verification data will be deleted after a short retention period - for example, Reddit relies on a US-based verification provider that promises to discard data within a week. However, recent security breaches with social media channels such as Discord have raised concerns about the reliability and safety of third-party age verification systems, which have already been shown to exhibit racial and gender biases (Stardust et al., 2024). These incidents highlight the risks involved in normalising identity verification as a prerequisite for accessing sensitive content online.

These issues are also particularly important for users participating in communities that discuss suicide and self-harm. Ofcom, the UK's communications regulator, has stated that the relevant provisions of the OSA are intended to restrict access to content that "encourages, promotes, or provides instructions" for suicide and self-harm (UK Government, 2025). However, it remains unclear how personal narratives of self-harm or suicidal ideation are addressed under this framework - particularly if the distinction between "harmful" and "helpful" content is constructed by regulatory bodies with limited or no input from those with lived experience. As Marsh et al. (2022) have argued, suicide prevention initiatives have frequently conflated "pro-choice" suicide forums with content that actively promotes suicide. In practice, these spaces often function as communities of support, particularly for users who have found mainstream healthcare services or charity helplines inaccessible or ineffective. Enforcing a rigid, pre-determined binary between "harmful" and "helpful" content (Wadsworth, 2025), and restricting access to one side of that divide at the expense of user privacy, risks silencing perspectives that challenge mainstream suicide prevention narratives and denies access to peer support.

This poster addresses challenges created by the OSA for researchers examining social media and suicide, drawing on challenges encountered on the ‘Discovering Liveability’ project.

Objectives/Research Questions

Work Package 3 of ‘Discovering Liveability’ aims to examine suicide and activism by analysing campaign materials from 20 UK-based public campaigns that seek to ‘make change’ relating to suicidality and ‘liveability’. This poster will discuss practical and methodological issues associated with this type of research. It will also outline the decision-making processes involved in searching for campaigns and materials, illustrating how the OSA may influence the type of data collected, and reflecting on challenges for researchers exploring social media, especially in sensitive contexts such as suicide.

Methods

The sample for the campaign analysis was identified through three approaches. First, based on the research team’s knowledge of prominent and less prominent groups/activists in the suicide prevention and mental health domain, we identified potentially relevant campaigns. We also consulted members of the Work Package 3 project advisory board who made additional suggestions. Second, we conducted keyword searches for campaigns on suicide prevention and/or liveability via a search engine and social media websites X, Reddit, Instagram, TikTok and BlueSky. Third, we followed up campaigns identified through the campaign analysis itself.

The enforcement of the OSA resulted in two practical issues for the campaign analysis. First, due to the institutional Computing Acceptable Use Policy, we could not use a VPN to bypass age restricted content on social media while using a university-managed device. Due to privacy concerns about third-party verification of personal ID, a decision was made to not use team members’ ID to verify age. Second, even though the Online Safety Act became enforceable in July 2025, not all content that could be considered sensitive was immediately age-restricted. Webpages and social media companies introduced restrictions gradually as content was flagged, which meant that some content was visible without age verification for some time after the OSA was enacted. Therefore, as a precautionary measure, we downloaded campaign material from campaigns that may be at risk of becoming age-restricted at a later date – this included content from the groups Rethink Mental Illness, StopSIM, and Stop Oxevision.

Results

We highlight two main challenges associated with searching for campaign materials online.

The first involves the practical problems of working with social media data following the enforcement of the OSA. Although efforts were made to identify both prominent and less visible campaigns to capture a range of views for this part of the project, restrictions that arose due to the OSA may have resulted in the omission of campaigns or campaign materials that are age-restricted. To comply with the OSA, some social media channels have restricted search results in such a way that only registered and age-verified users can see content that may be marked as sensitive due to references to suicide and self-harm. It is possible that these materials may have been missed, as we did not use an age-verified account to browse campaign materials.

It is also difficult to assess the extent to which lived experience of suicidality is amplified or obscured by age-restriction. For example, Reddit allows users to view thread titles but requires an age-verified account to access full discussions, whereas Discord prevents users from seeing age-restricted communities until verification is completed. cursory Google searches retrieved Reddit posts that illustrated inconsistencies: results of preliminary searches on suicide prevention on Reddit revealed that posts asking what to expect when calling a suicide prevention hotline were accessible, as were discussions of suicide in fiction, or debates about the ethics of suicide. However, comments/posts with titles challenging dominant suicide prevention narratives were all inaccessible without an age-verified account. Inaccessible posts included critiques of suicide prevention as virtue signalling, recommendations of texts in critical suicidology, and personal accounts of negative experiences with suicide helplines.

The second challenge was with using social media and search engines to identify and collect data. It is possible that the search strategy used for this campaign analysis may have missed potentially relevant results. For example, activists may not always describe their work as a “campaign”. For example, Faith Action describe their suicide prevention work as a resource rather than a campaign, so the suicide prevention work conducted by this group did not show up in searches. Additionally, it is possible that more disruptive or radical activists may choose to block their campaigns from appearing in Google search results due to privacy or safety concerns. Moreover, search engines such as Google consider various factors such as location and search history to provide customised results. Although a new Google account was created to limit the influence of these factors, it is possible that potentially relevant results may have been missed as they were not prominently displayed in the search results.

Future Work

Activists and those with lived experience of suicide and self-harm frequently use blogs, zines, artwork, short-form media, and other non-traditional forms of communication, such as forums and discussions threads on social media, to articulate experiences that are left out of mainstream suicide prevention discourse (Anderson and Spandler, 2025,

Lundeen et al. 2024). For academic research in particular, such content can help mitigate epistemic injustice and disrupt the closed circulation of knowledge within academia (Walker, 2025). By placing restrictions both on the content of these discussions and the type of people who can participate in these conversations, these new regulations may risk the production and circulation of only ‘approved’ perspectives about suicide and self-harm, both within academic literature and on social media.

Although research has shown that age verification systems can be, and are, bypassed (Woodley et al. 2025), academic researchers must work within legal, institutional, and ethical frameworks. The challenges encountered during the search for campaign materials in this study indicates that with the enforcement of the OSA, researchers may be left with a limited set of options: either submit personal identification for professional purposes, trusting that third-party verification providers will handle this data responsibly, or avoid age verification altogether, and risk losing access to valuable content about lived experience that fall outside what regulators consider acceptable.

The long-term impact of the OSA on researchers handling sensitive social media data remain uncertain. As social media platforms and academic institutions update their policies, a lack of clear guidance creates significant methodological and practical challenges. This uncertainty impacts not only lived experience communities and support spaces, but also the types of knowledge about suicide that are produced and circulated.

References

- Anderson, J., & Spandler, H. (2025). Mad Zine pedagogy: using zines in critical mental health learning and education. *Social Work Education*, 1-19.
<https://doi.org/10.1080/02615479.2025.2469586>
- Lundeen, L. A., McCall, J. R., Bradshaw, A. S., LeBlanc, E. L., & Humphrey, L. (2024). Digital catharsis or harmful exposure? A thematic analysis of self-directed violence Reddit posts. *Social Media+ Society*, 10(3), 20563051241263562.
<https://doi.org/10.1177/20563051241263562>
- Marsh, I., Winter, R., & Marzano, L. (2022). Representing suicide: Giving voice to a desire to die?. *Health*, 26(1), 10-26.
<https://doi.org/10.1177/13634593211046843>
- Stardust, Z., Obeid, A., McKee, A., & Angus, D. (2024). Mandatory age verification for pornography access: Why it can't and won't ‘save the children’. *Big data & society*, 11(2), 20539517241252129.
<https://doi.org/10.1177/20539517241252129>

UK Government. (2025, July 23). What's changing for children on social media from 25 July 2025. GOV.UK. <https://www.gov.uk/government/news/whats-changing-for-children-on-social-media-from-25-july-2025>

Wadsworth, T. (2025). "Care" Beyond Reducing Harm: (Non-)Recovery and the Regulation of Online Self Harm Content. Gold.ac.uk. https://research.gold.ac.uk/id/eprint/39193/1/SOC_thesis_WadhamT_2025.pdf

Walker, T. (2025). Zine and Heard?: mental health survivors, zines and epistemic justice. <http://doi.org/10.17030/uclan.thesis.00057508>

Woodley, G., See, H.W., O'Neill, B., Green, L., Staksrud, E. and Haskell-Dowland, P., 2025. Australian Teen Voices on Age Verification and Age Assurance Measures. Policy & Internet, 17(4), p.e70019. <https://doi.org/10.1002/poi3.70019>

12.36 Blondies, Botox, and Bird Legs: Analyzing Gender Humor in TikTok Videos Criticizing Republican Women

Jerrica Rowlett (Bryant University)

Introduction

Michelle Obama, former First Lady of the United States, famously said in 2016, “when they go low, we go high” (The White House, 2016). This was a call for US citizens on the left side of the political spectrum not to be vengeful when facing a challenge, but to approach it with openness to find a solution (Bruner, 2018). After years of feeling defeated by taking the high road, many have said, “sorry, Michelle, we’re going low.” Using stereotyped jokes about Republican politicians, individuals participate in resistance to the current administration through online bullying and trolling. Gender humor has been used to criticize the women who work within the current Trump administration. Dumb blonde jokes have been made about Press Secretary Karoline Leavitt and US Attorney General Pam Bondi. Secretary of Homeland Security Kristi Noem has been mocked for her dramatic appearance change, with the term “Mar-A-Lago Face” being coined to make fun of the distinct style of plastic surgery. Liberal women started a trend on TikTok called “Republican makeup,” in which they apply heavy amounts of contour and blush poorly to mock the makeup skills of conservative women.

Women in the political field have routinely been critiqued for their physical appearance (Campus, 2013; Hayews et al., 2016). When Kim Davis, a county court clerk, refused to issue marriage licenses to LGBTQIA+ couples following the federal legalization of gay marriage in the US, meme creators online resorted to mocking her for her physical appearance and slut-shaming her rather than discussing LGBTQIA+ rights (Harlow et al., 2018). Humor can aid in the democratic process when it brings light to issues, but it can also perpetuate harmful stereotypes (Peifer, 2012). This research explores the use of gender humor in online discourse about Republican women on TikTok.

Objectives/Research Questions

Gender norms are behaviors, feelings, or thoughts that one is ‘supposed to’ partake in based on their gender (Alberts, 1992). Norms assigned to women include femininity, attractiveness, niceness, morality, sexual loyalty, and submissiveness to men (Crawford & Unger, 2000; Gilbert & Scher, 1999). With this, jokes that affirm these stereotypes will often resort to mocking the appearance of a woman (Spradley & Mann, 1975), objectification (Bergmann, 1986), and sexual shaming (Sykes, 1966). Ridiculing appearance can manifest in the form of mocking a woman’s body, face, clothing, and beyond. Objectification primarily takes form in the reduction of a woman to a sexual

object. Sexual shaming can take many forms, including highlighting a woman's premarital relationships, suggesting sexual infidelity, slut-shaming, and beyond. Building on the previous research, the first proposed research question is:

RQ1: What types of gender humor do TikTok creators utilize when criticizing Republican women?

Accounting for the role of human interaction with technology, social media affordances can be defined as “the perceived actual or imagined properties of social media, emerging through the relation of technological, social, and contextual, that enable and constrain specific uses of the platforms” (Ronzhyn et al., 2023, p. 3178). For creators on TikTok, the ability to add pre-existing sound allows participation in viral trends, remixing content, and telling their story (Vizcaíno-Verdú & Abidin, 2022). Filters allow creators to digitally transform the appearance of their face. The beauty filter is a popular filter that transforms creators by slimming the face, lifting the brow, and applying makeup. This filter has been found to perpetuate beauty stereotypes and increase the focus on women's appearances (Kenalemang-Palm & Eriksson, 2026). Another filter titled “Latina makeup” perpetuates racism through caricaturized versions of Latin features and makeup in an easily shareable template, encouraging participation (Farias, 2026). It is apparent that the use of TikTok filters can perpetuate stereotypes. With this, the following question is posed:

RQ2: How do TikTok creators utilize the technological affordances of the application to criticize Republican women while using gender humor?

Methods

To answer the posed questions, video data was collected using the Zeeschuimer (Peeters, 2026) Firefox browser extension. Using a new TikTok account, the names Karoline Leavitt, Pam Bondi, Kristi Noem, and Marjorie Taylor Greene (MTG) were individually searched with the addition of the words humor, parody, comedy, and funny (e.g., “Karoline Leavitt humor”) using the search feature on the TikTok website. These four women were selected because they are well-known names associated with the Republican Party in the United States. The initial data set included 350 videos. The researcher reviewed the data for relevance and removed irrelevant videos. For example, several videos in the data set were from Karoline Leavitt's official TikTok account. Additionally, videos from news sources and videos that were not humorous were removed. This left a total of 308 videos for analysis.

A textual and visual qualitative analysis (Frey et al., 1999; Hsieh & Shannon, 2005) of the data was performed, examining both the video and their associated captions. Additionally, the metadata of effects used by creators to alter their appearance while creating content was analyzed.

The qualitative analysis has been completed at this time. A quantitative content analysis is underway (currently establishing intercoder reliability). The codebook for the quantitative analysis can be found attached as the supplementary file.

Results

Following the release of photos from a Vanity Fair photoshoot that featured an extreme close-up of Leavitt, creators were quick to comment on the quality of her skin, makeup skills, and noticeable dots on her lips that were presumed to be lip filler injection marks. Many creators drew exaggerated dots on their own lips to mock Leavitt while creating parodies of her. One creator starts with the close-up photo on the screen and states, “alright, if you’re a feminist and you're sharing this photo and making fun of it, then I’ve got some hard truths for you and I need you to watch this video all the way to the end.” The creator goes on to discuss how Leavitt is held to the same beauty standards we all are. In the last seconds, the camera zooms in on the creator to reveal “JK JK SHE SUCKS” written on her lips

Aside from jokes about the physical appearance of the Republican women, they were also slut-shamed at times. MTG called for a “national divorce” between red states and blue states because of the “sick and disgusting woke culture issues shoved down our throats.” In response to this, one creator discussed the moment stating, “ironically, for the same reasons her husband filed for divorce from her, things being shoved down her throat.” This statement references a rumored affair the congresswoman was engaged in, but rather than simply stating she was having an affair, the creator created the image of MTG performing oral sex on a man. There were also videos that insinuated each of the women have sexual relations with Trump.

TikTok allows users the ability to create filters that can be used by anyone on the app. When analyzing the effects used by creators, effects titled “Marjorie” and “KAROLINE L” stood out and were frequently used. The Marjorie TikTok filter widens the cheeks and narrows the eyes of the creator, mocking by over-exaggerating the natural features of MTG. The KAROLINE L filter places dark, contoured makeup on the creator. Many creators also used the “small lips” filter when parodying Leavitt.

In addition to using the filters feature to mock the appearance of the women, users created sounds that could be used by anyone on the app. The song Roses by Outkast, featuring the lyrics, “Caroline, see, she's the reason for the word 'bitch,’” was frequently used over edits of Leavitt. One popular sound was a song with the lyrics “Kristi Noem is a birdlegged hoe” on repeat before ending with, “that hoe, that hoe has got to go.” This sound mocks her physical appearance by emphasizing the proportions of her legs to the rest of her body, while also calling her a hoe, slang that is used in a derogatory manner toward women who are perceived to be promiscuous. Another sound featured in the data set was a song created with the lyrics, “Pam Bondi is a dumb

blondie.” Bondi and Leavitt were both referred to as dumb blondes or blondies, a term used to insinuate that blonde women are less intelligent.

Future Work

In answering RQ1, many creators resort to gender tropes that reinforce stereotypes about women, particularly by fortifying beauty standards, perpetuating blonde stereotypes, slut-shaming, and sexualizing the women. While previous research found that political issues were often ignored in favor of mocking the gender identity of the politician (Harlow et al., 2018), this qualitative analysis shows a different trend. There were many instances of the creators purely mocking the Republican women for not conforming to gender norms, but several videos blended gender humor with mocking direct actions or policies of the women (e.g., MTG calling for national divorce). Humor can be used to perpetuate stereotypes, but it can also provide insight into a situation and strengthen democracy (Peifer, 2012). The quantitative content analysis will allow deeper insight into the messages being spread on TikTok to determine the prevalence of gender humor and political humor that can be insightful.

Answering RQ2, creators take advantage of both the visual and audio affordances of TikTok. Visually, the creators dramatically emphasize the style choices of the Republican women’s makeup, botox, and lip filler. Beyond cosplaying the characters, some have taken the time to create filters that morph the facial features of creators to resemble a caricaturized version of the Republican women. The creation of filters allows for easier and quicker mocking of their physical appearance by a larger audience through an easily transferable filter. Audibly, the creation and use of popular sounds creates a soundscape to the online resistance, affording creators another low-resistance path to participating in the online discourse. Yet, many of the popular sounds reinforce gender norms. As the research continues, TikTok affordances utilized by the creators will be coded in the quantitative analysis, revealing the frequency of use of these customized filters and sounds, as well as the use of other TikTok affordances. A chi-square analysis will also bring insight into the use of affordances with specific gender humor types.

It is arguable that creators of content mocking Republican women would be left-leaning in their political ideologies - a side that typically fights for gender equality. When debating the ethics of political humor, Peifer states, “if the attitude that underlies a joke is sexist...than the laughter evoked by such humor is a morally deficient endorsement of sexism” (2012, p. 266). With this, creators and enjoyers of the content would be participating in sexism while simultaneously voting and arguing for women’s rights. Thus, it would be worthwhile to study if creators and audience members experience cognitive dissonance. Additionally, cross-cultural analysis would reveal if this is a unique experience to the US political system or if citizens of other countries also utilize gender norms when criticizing conservative women.

References

- Alberts, J. K. (1992). Teasing and sexual harassment: Double-bind communication in the workplace. In L. A. Perry, L. H. Turner & H. M. Sterk (Eds.), *Constructing and reconstructing gender: The links among communication, language, and gender* (pp. 185–196). Albany: State University of New York Press.
- Bergmann, M. (1986). How many feminists does it take to make a joke? Sexist humor and what's wrong with it. *Hypatia*, 1(1), 63–82.
- Bruner, R. (2018, November 20). Michelle Obama explains what “going high” really means. *Time*. <https://time.com/5459984/michelle-obama-go-high/>
- Campus, D. (2013). The appearance of power. In *Women Political Leaders and the Media* (pp. 73-93). London: Palgrave Macmillan UK.
- Crawford, M., & Unger, R. (2000). *Women and gender: A feminist psychology*. Boston, MA: McGraw-Hill.
- Farías, C. (2026). Filtered Latinidad: Ethnoracial abstraction and everyday resistance in TikTok beauty filters. *Information, Communication & Society*, 1-18.
- Frey, L., Botan, C., & Kreps, G. (1999). *Investigating communication: An introduction to research methods* (2nd ed.). Boston, MA: Allyn and Bacon.
- Gilbert, L. A., & Scher, M. (1999). *Gender and sex in counseling and psychotherapy*. Boston, MA: Allyn & Bacon.
- Harlow, S., Rowlett, J. T., & Gonyea, L. K. (2018). 'Kim Davis be like...': A feminist critique of gender humor in online political memes. *Information, Communication, & Society*, 23(7), 1057-1073. DOI: 10.1080/1369118X.2018.1550524
- Hayes, D., Lawless, J. L., & Baitinger, G. (2014). Who cares what they wear? Media, gender, and the influence of candidate appearance. *Social Science Quarterly*, 95(5), 1194-1212.
- Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288.
- Kenalemang-Palm, L. M., & Eriksson, G. (2026). The “perfect” filtered look? A multimodal critical discourse analysis of TikTok #beautyfilters. *Critical Studies in Media Communication*, 43(1), 73-87.
- Peeters, S. (2026). Zeeschuimer (v1.13.6). Zenodo. <https://doi.org/10.5281/zenodo.18670373>

- Peifer, J. (2012). Can we be funny? The social responsibility of political humor. *Journal of Mass Media Ethics*, 27(4), 263–276. <https://doi-org.bryant.idm.oclc.org/10.1080/08900523.2012.746110>
- Ronzhyn A., Cardenal A.S., & Batlle Rubio, A. (2023). Defining affordances in social media research: A literature review. *New Media & Society*, 25(11): 3165-3188.
- Spradley, J. P., & Mann, B. (1975). *The cocktail waitress*. New York: Wiley.
- Sykes, A. J. M. (1966). Joking relationships in an industrial setting. *American Anthropologist*, 68, 188–193.
- The White House. (2016, July 25). Remarks by the first lady at the Democratic National Convention. National Archives and Records Administration. <https://obamawhitehouse.archives.gov/the-press-office/2016/07/25/remarks-first-lady-democratic-national-convention>

12.37 Vine Compilations on YouTube: Social Media Archiving, Digital Nostalgia, and the Preservation of User-Generated Content

Joelle Saunders (Oxford Internet Institute)

Introduction

In *Cultural Analytics*, Manovich (2020) characterizes cultural production in the digital age by invoking the idea of billions of images posted, songs shared, or interactions documented across digital media every day. This invocation is not hyperbole - on TikTok alone, approximately 34 million videos are uploaded worldwide each day (Shepherd, 2025). The content that is circulated on these platforms largely falls into the category of user-generated content (UGC): photos, blogs, short and long videos, comments, and any other content made by the users of a platform or service. UGC contributes new cultural content rapidly and at scale (Manovich). While most UGC disappears into the digital ether, some gain substantial attention, going viral and sometimes becoming memes. Thus, we are in an era of cultural production in which UGC is abundant and culturally significant, shaping shared cultural references and memories, in particular examples and categorically. But the social media platforms on which UGC is created, published, and circulated do not remain the same or last forever - they are unstable entities, subject to technical-material, economic, political, or social shifts (Poell et al., 2022). Platforms can evolve or disappear entirely, threatening access to and preservation of content. This renders UGC also categorically vulnerable. These dynamics have prompted projects like Archiving Black Twitter, which seeks to preserve not only content but its significance for internet history and culture (Summers, 2023). Scholars in digital and social media archiving have stressed the continual disappearance of web content and the need for archiving; have differentiated between types of digital material in the archiving process; and have conceptualized social media archiving practices as more open, collaborative, and idiosyncratic, though research remains limited. Platform studies emphasize platformization and precarity, but there are fewer theories and explorations of platform 'death' and the afterlives of content. Finally, digital nostalgia is not just a current buzzword and trend but a burgeoning interdisciplinary field of study where distinctions between types of expressions and practices of nostalgia encourage an agenda for more analytical inquiry and empirical research. This study provides a preliminary outline and execution of this agenda, ultimately weaving together all three fields.

Objectives/Research Questions

Drawing on concepts from extant scholarship on web and social media archiving, platform studies, and digital nostalgia, this study begins to fill gaps in these fields by asking: What happens to user-generated content when platforms ‘die’? How does an internet culture preserve itself?

The objective of this study is to explore the inherent tension between the scale of cultural production, the meaning of digital internet cultural artifacts to internet users, and the ability to preserve both, given the infrastructure and affordances of the social web. Additionally, through the chosen case study, this study is in itself an effort at remembering a platform that ‘died’: Vine, an exploration of YouTube’s affordances through a particular lens, and a documentation and appraisal of internet users’ responses to these two actualities, both in their words and actions. Lastly, it is a study that is particularly interested in theoretical and empirical development in the study of platform evolution’s effect on UGC and digital nostalgia, especially as the latter is a more novel field of study and a major cultural trend of recent months and years.

Methods

To explore these questions, this study builds a qualitative case study around Vine compilations on YouTube. First, a historical review traces the life and ‘death’ of Vine drawing on extant research on the platform and popular sources. A similar review is executed for YouTube, with necessarily more emphasis on its evolution and affordances for UGC. Second, a descriptive content analysis examines a small sample of Vine compilations on YouTube, making note of their visual presentations and other descriptive characteristics, such as observed similarities and differences in the creator or subject. Third, a preliminary thematic analysis of viewer comments on Vine compilations identifies expressions of digital nostalgia in how internet users, acting as audiences of the compilations, engage with and narrate their experience with them. Given the exploratory and preliminary nature of the study, the sample collection is more purposive, and the analyses are primarily descriptive. Future iterations would benefit from a more systematic and expansive sample collection and analysis, especially as it pertains to the content and thematic analyses of the compilations and comments, where more patterns, themes, and theories can be drawn out, interpreted, and developed.

Results

A summary of the preliminary findings are generally stated as follows: (1) informal archiving practices for UGC (such as vines) can emerge organically from internet users when platforms ‘die’; (2) these practices can be intertwined with expressions and practices of digital nostalgia, extending years beyond the platform’s ‘death’; (3) platforms afford archiving practices and related expressions of digital nostalgia, while ironically structuring the need for both.

To explain further: the existence of and engagement with these compilations illustrate the connections between preserving UGC and expressions of digital nostalgia, with the nature and affordances of platforms playing a crucial role in both. The "reborn-digital" artifacts of Vine compilations link the practices of "informal archiving" and "digital nostalgizing" while still inadvertently emphasizing the platform instability that renders user-generated content vulnerable. These re-presentations of Vine and its content can only approximate the original context. As newer UGC, the compilations are a method of preserving content after platform 'death' on a different, still-active platform, but they are ultimately subject to the same vulnerabilities. YouTube may not face imminent shutdown, but platforms continually evolve, and content and access are always at risk of changing with them. Still, YouTube becomes a space for informal archival practices and for digital nostalgizing. Users comment not only to reflect their personal feelings, but also their collective, communal memory of experiencing what life, culture, and the internet were like when Vine was active.

Attending to the study's core questions, there are two preliminary observations. In the face of the loss and constant threat of loss of UGC due to platform evolution, users can exercise a sort of autonomy in preserving and re-presenting content - exemplified by the organic emergence of compilations (especially compared to platform-led archives that were ultimately discontinued). Additionally, an internet culture preserves itself by going through processes of remembering itself - digitally and discursively bringing up the past as content multiplies ever more in the present and into the future.

Future Work

As mentioned in the methods section, future iterations of this study would benefit from more systematic and expansive sample collection and analysis.

As this work gestures at gaps in three different areas of scholarship and speaks to instances of their collision, future work could address each separately or with respect to their various intersections. For example, future research should examine informal, collective, and/or amateur archiving practices that exist across the social web, particularly the motivations and processes of practitioners. Future work in platform studies could find it fruitful to further develop and empirically explore theories of platform 'death' and the afterlives of content. This work could have implications and insights for areas such as socioeconomics of platformization, personal data rights, interoperability, and digital rights management. The agenda for digital nostalgia proposed in this work would highlight the further theorizing of digital nostalgia as affective experiences, discourses, and practices, and the ways that these empirics and theories would open up avenues to better understand the present and future of social media by better understanding our relationship to the past. This work could prove insightful for users, for those in industry/working at social media platforms, and even for policy stakeholders, as regulating social media platforms should and implicitly does

draw on understanding the sociocultural expectations for our experiences with/on social media.

Overall, this study adds to and encourages more work on the linkages between collective experiences and memories of social media platforms and content, the spectrum of activities and actors that are included in current digital and web archiving, and how platforms perform as sociocultural infrastructure in the digital era of cultural production.

References

Manovich, L. (2020). Cultural Analytics. MIT Press.

Poell, T., Nieborg, D. B., & Duffy, B. E. (2022). Platforms and Cultural Production. Polity Press.

Shepherd, J. (2025). 25 Essential TikTok Statistics You Need to Know in 2025. The Social Shepherd; The Social Shepherd. <https://thesocialshepherd.com/blog/tiktok-statistics>.

Summers, J. (2023). She's trying to archive Black Twitter. It's a delicate and imperfect task. NPR. <https://www.npr.org/2023/05/26/1178262041/archiving-black-twitter-elon-musk-meredith-clark>

12.37 Who Is Afraid Of Fat Bodies? Fatphobic Representations On Social Media: An Analysis

Lia Selvi (Sapienza Università di Roma)

Introduction

Fatphobia can be understood as a product of the cultural habitus (Bourdieu, 2014) underlying contemporary social systems (Strings, 2022). Its development is linked to the transition from modern society to a liquid or hypermodern condition (Bauman, 2008), characterized by fragmented identities (Maffessoli, 2017), existential insecurity and the erosion of stable ideological frameworks, that once structured panoptic modernity (Bentham, 1997). Within this context, contemporary culture increasingly operates through aestheticized and “showcase” logics (Codeluppi, 2024), shaped by media and technological innovations such as photography, cinema and television (Sontag, 2004). These processes have produced a social grammar dominated by visual paradigms (Lipovetsky & Serroy, 2017), in which reality is simplified and transformed into images in order to become socially intelligible (Bellino, 2010; Simmel, 2020).

At the same time, the rise of consumer society and the rationalization of everyday life (Weber, 1991) have made identity increasingly measurable through productivist and performative logics. The body therefore becomes a central identity device, no longer primarily experienced as a relational interface with the world (Bauman, 2008) but as a project to be optimized and displayed. In this sense, the body functions as a simulacrum (Baudrillard, 2022) and a key site of self-construction (Codeluppi, 2024; Bordo, 1997).

Fatphobia should therefore be interpreted not simply as aesthetic discrimination but as a social technology of power (Foucault, 1992), reproduced through media and digital environments (Jiménez-García et al., 2025). In platform society (Van Dijck, Poell & De Waal, 2019), interactions between users and digital infrastructures generate feedback loops between social habitus and machine habitus (Airoldi, 2024), shaping cultural norms and representations (Oswald et al., 2025).

Despite the centrality of the body in sociological inquiry (Gailey, 2025; Gaspar et al., 2025), fatphobia remains underexplored, particularly in the Italian context. This research aims to address this gap by analysing the representation of fat bodies within social media from a sociological perspective.

Objectives/Research Questions

Building on the theoretical frameworks outlined above – critical sociology, media studies, and gender studies – this research investigates the representation of fat female

bodies on social media, with particular attention to platforms frequented by pre-adolescents and adolescents. The analysis centres primarily on TikTok, a platform that remains underexplored in academic literature, while also encompassing Instagram and YouTube, which serve broader demographic ranges and have received greater scholarly attention.

Three interconnected objectives structure the research.

The first is to identify and critically interpret the narratives associated with female fatness in digital environments, examining their ethical, symbolic, and ideological dimensions. Drawing on discourse analysis and visual sociology, this phase aims to map how fat bodies are framed, stigmatized, or reclaimed within platform-specific logics, and to assess how these representations intersect with broader systems of gender, class, and power.

The second objective shifts focus to reception: the research will examine how users, particularly young women and adolescents, perceive, internalize, or resist these representations, and how exposure to fatphobic content may shape body image, self-definition, and identity construction. This phase draws on the concept of machine habitus (Airoidi, 2024) to explore how algorithmic curation amplifies or constrains specific body norms, producing feedback loops between platform architecture and user subjectivity.

The third objective is interpretive: findings will be situated within dominant cultural and social frameworks to assess the structural weight of fatphobic stigma in collective representations and the social imaginary. This step aims to connect micro-level digital interactions to macro-level processes of normalization and symbolic violence.

Taken together, these objectives lead to the following research question: How are fat female bodies represented on TikTok, Instagram, and YouTube, and what role do platform logics and algorithmic infrastructures play in reproducing or challenging fatphobic stigma among female users in the Italian context?

Methods

This research adopts a mixed-methods approach with a predominantly qualitative orientation, structured across three sequential phases.

The first phase consists of an analysis of media content drawn from TikTok, Instagram, and YouTube. Content will be examined through Greimasian semiotic analysis, aimed at identifying the signs, symbols, and narrative structures underlying digital representations of fatphobia, including the construction of semiotic squares to map the logical articulation of the communicative paradigms identified. Content analysis and discourse analysis will be employed as complementary tools to detect recurring visual, syntactic, and semantic elements.

The second phase investigates the reception and internalization of these contents among female-identifying individuals aged 10 to 35. The sample will be divided into two sub-groups according to age. The first, comprising pre-adolescents and adolescents (10–18), will be recruited through schools, within the framework of existing programmes on nutrition education or eating disorder prevention. This sub-group will be further split through randomization into an experimental group, exposed to a selected corpus of fatphobia-related media content before completing a task, and a control group, who will complete only the task. Assessment tools under consideration include the Antifat Attitudes Questionnaire (Lewis et al., 1997) and the Weight Implicit Association Test (Weight-IAT), both suitable for younger populations. Ethical safeguards, including age-appropriate informed consent and careful media selection to avoid negative emotional impact, will be established prior to data collection. The second sub-group, young adults (18–35), will be recruited through snowball sampling and engaged through in-depth interviews, following the same experimental/control structure.

The third phase focuses on the embodiment of fatphobic social paradigms among adults (25–35), demographic already accustomed to prolonged media exposure and equipped with a broader sociocultural background. Focus groups will be used to surface both individual perceptions and collective symbolic constructions, examining the extent to which prolonged exposure to fatphobic media environments shapes body image and identity over time.

Results

This research is conceived with full awareness of the complexity and scope of the phenomena under investigation, as well as the ambition of the proposed methodological framework, which may be appropriately adapted as the study develops.

The analysis pursues three main outcomes.

The first is to define and categorize the principal media narratives associated with fat female bodies across TikTok, Instagram, and YouTube, identifying their core constitutive elements. The second is to analyze how these digital contents are perceived and internalized across different age groups of female-identifying individuals, in order to understand the dynamics of bodily and self-identity construction in relation to fatphobic stigma. The third is to situate fatphobia within contemporary culture, offering interpretive frameworks that may support a critical overcoming of discriminatory representations.

This final point also points toward an applied dimension of the research that extends beyond strictly sociological inquiry. By orienting the communicative landscape toward a media literacy perspective, the study aims to contribute to the development of socially responsible and inclusive communication models, relevant to the fields of

economics and marketing, capable of combining economic effectiveness with social sensitivity, and of promoting more conscious and respectful representations of bodily diversity without compromising the clarity and impact of messaging.

Future Work

Several directions emerge as promising extensions of this research. First, it would be valuable to examine the reproduction of body stigmatization within artificial intelligence systems, with particular attention to how AI-generated images reflect and reinforce weight bias. As generative visual tools become increasingly integrated into media and communication environments, understanding how thinness norms are encoded in their outputs represents a critical and largely unexplored area of inquiry.

Second, the Italian context would benefit from the development of dedicated psychometric instruments, such as IAT-based tools or similar implicit association measures, calibrated to contemporary cultural norms around body image. Existing instruments have been predominantly developed and validated in Anglo-American contexts, and their adaptation to the Italian landscape remains limited.

Both directions are made particularly urgent by a broader cultural moment in which the obsession with thinness appears to be undergoing a significant resurgence. The widespread and increasingly normalized use of weight-loss drugs, such as GLP-1 receptor agonists, signals a renewed cultural investment in body optimization that intersects with media logics, platform cultures, and medical discourse in ways that sociological research has yet to fully address.

References

- Airoidi, M. (2024). *Machine Habitus. Sociologia degli algoritmi*. Luiss University Press.
- Baudrillard, J. (2022). *Simulacri e impostura. Bestie, Beaubourg, apparenze e altri oggetti* (M. G. Brega, Ed.). Pgreco.
- Bauman, Z. (2008). *Vita liquida*. Laterza.
- Bauman, Z. (2014). *La società dell'incertezza*. Il Mulino.
- Bellino, F. (2010). *Per un'etica della comunicazione*. Mondadori Bruno.
- Bentham, J. (1997). *Panopticon ovvero la casa d'ispezione*. Marsilio.
- Bordo, S. (1997). *Il peso del corpo*. Feltrinelli.
- Bourdieu, P. (2014). *Il dominio maschile*. Feltrinelli.
- Codeluppi, V. (2024). *La morte della cultura di massa*. Carocci.
- Dijck, J. van, Poell, T., & Waal, M. de (2019). *Platform Society. Valori pubblici e società connessa*. Guerini Scientifica.

Emerson, R. W. (2012). *L'anima suprema, l'amore, l'amicizia, la politica*. Ortica Editrice.

Foucault, M. (1992). *Tecnologie del sé*. Bollati Boringhieri.

Gailey, J. A. (2025). 'I don't eat that much': Fatphobia and women baby boomers' fear of weight gain. *Journal of Gender Studies*.

<https://doi.org/10.1080/09589236.2025.2554215>

Gaspar, M. C., Verthein, Ú., & de Moraes Sato, P. (2025). 'Medical fatphobia is not something we invent': Weight stigma in healthcare experienced by women in the antifatphobia movement in Brazil and Spain. *Sociology of Health & Illness*, 47, e70062.

<https://doi.org/10.1111/1467-9566.70062>

Jiménez-García, A. M., Arias, N., Hontanaya, E. P., et al. (2025). Impact of body-positive social media content on body image perception. *Journal of Eating Disorders*, 13, 153.

<https://doi.org/10.1186/s40337-025-01286-y>

Lipovetsky, G., & Serroy, J. (2017). *L'Estetizzazione del mondo*. Sellerio Editore Palermo.

Lewis, R. J., Cash, T. F., Jacobi, L., & Bubb-Lewis, C. (1997). Prejudice toward fat people: The development and validation of the antifat attitudes test. *Obesity Research*, 5(4), 297–307. <https://doi.org/10.1002/j.1550-8528.1997.tb00555>

Maffesoli, M. (2017). *Nel vuoto delle apparenze. Per un'etica dell'estetica*. Edizioni Estemporanee.

Oswald, F., Pham, M. D., Harr, R. J., Garr-Schultz, A., & Chaney, K. E. (2025). "They are fat and want special treatment for being fat": Backlash to and lay theories of fat activism. *Analyses of Social Issues and Public Policy*, 25, e70006.

<https://doi.org/10.1111/asap.70006>

Simmel, G. (2020). *Stile moderno. Saggi di estetica sociale*. Einaudi.

Sontag, S. (2004). *Sulla fotografia. Realtà e immagine nella nostra società*. Einaudi.

Strings, S. (2022). *Fat Phobia. Mar dei Sargassi*.

Weber, M. (1991). *L'etica protestante e lo spirito del capitalismo*. Rizzoli.

12.38 The Body in Algorithmic Colonialism: Artistic Inquiry into Emotion-AI and Embodied Affective Relations

Chunyi Shen (UAL)

Introduction

Emotion-AI, a subset of artificial intelligence that translates human facial expressions and physiological signals into standardised emotional categories, is rapidly embedded across social media platforms, healthcare systems, educational technologies, and immersive environments. As these systems proliferate, they do not merely observe affect; they actively shape it. Drawing on Couldry and Mejias's (2019) framework of data colonialism, this research positions Emotion-AI as a contemporary apparatus of algorithmic colonialism: a system that imposes culturally dominant affective norms onto bodies, extending colonial logics of universality and legibility into digital infrastructure. When Emotion-AI translates the complexity of human affect into narrow computational categories, it reproduces culturally specific biases as universal truths, forming what Glickman and Sharot (2025) describe as recursive human-AI feedback loops, in which algorithmic judgement reshapes human perception and behaviour over time.

Critical algorithm studies have extensively theorised the political dimensions of AI systems, yet a significant methodological gap remains: purely textual or theoretical analysis rarely captures the subtle, embodied, and affective consequences of these systems as they are lived in the body (Bulley & Şahin, 2021). This practice-based PhD research addresses this gap by positioning artistic experimentation as a necessary, non-representational mode of inquiry, one capable of making algorithmic affect empirically tangible and critically reconfigurable (Rajcic & McCormack, 2020).

Objectives/Research Questions

This research is guided by two primary questions:

How does Emotion-AI enact algorithmic colonialism through the physiological and psychological discipline of bodies?

How can artistic inquiry function as a research method to investigate this embodied process and prototype alternative modes of human-emotion-algorithm relations?

The study aims to reveal and reconfigure the embodied impact of Emotion-AI through a series of artistic experiments, making the reach of algorithmic colonialism perceptible through the body rather than through abstraction alone.

Methods

This research adopts a practice-as-research methodology as part of an ongoing PhD project in its first year at the London College of Communication, University of the Arts London, positioning creative artistic experimentation as the primary site of knowledge production. The project develops three interconnected artistic artefacts, each corresponding to a different module of emotional algorithms. The first artefact currently at prototype stage is an AI-driven interactive installation that uses real-time facial expression recognition to modulate an on-screen narrative. Visitors' facial gestures are processed by an Emotion-AI system; the resulting emotional classifications trigger textual and visual changes, branching the narrative according to the algorithm's reading of the participant's affect.

Auto-ethnographic logging and phenomenological reflection constitute the core methods for this early research phase, situating the artist-researcher as both designer of and witness to the body-algorithm encounter. This approach is grounded in methodological pluralism: art, technology, affect, and body are treated as mutually constitutive modes of inquiry. Subsequent artefacts will integrate wearable physiological sensors within a VR environment (Artefact 2) and develop an immersive spatial installation responsive to collective emotional states (Artefact 3), employing semi-structured interviews, focus groups, and affective analysis alongside embodied and sensory methodologies.

Results

The first artefact has been realised as a functional interactive prototype and presented in a public exhibition (Unfolding Narratives) at the London College of Communication, University of the Arts London (2025). The installation was built using TouchDesigner with MediaPipe, a real-time facial landmark detection framework, to process visitors' facial gestures and trigger on-screen narrative responses. Auto-ethnographic observation during the exhibition revealed a pattern of immediate significance to the research's theoretical concerns: upon approaching the installation, participants encountered an on-screen prompt "please smile" issued by the system. Participants consistently complied, producing a smile that triggered the narrative response "I see you." Crucially, participants were observed subsequently adjusting and calibrating the degree of their smile widening or relaxing it in response to the system's real-time recognition feedback, displayed as a confidence percentage ("person with XX% happy"). This behaviour enacts precisely the recursive body-algorithm-emotion loop the research theorises: the body does not simply express emotion, but disciplines its affective display in order to become legible to the algorithm. Emotional expression becomes a performance optimised for computational recognition rather than interpersonal communication. The instruction to smile, and the body's compliance and self-adjustment, materially demonstrate how Emotion-AI functions as an apparatus of

affective discipline, a form of algorithmic colonialism operating at the level of facial muscle and felt experience.

Future Work

Following ethics approval from the University of the Arts London, the research will proceed to systematic participant engagement with Artefact 1, incorporating semi-structured interviews and phenomenological reflection to examine how experiences of affective misrecognition and algorithmic compliance vary across cultural backgrounds. Artefact 2 will extend the inquiry into physiological registers through a wearable-responsive VR environment, while Artefact 3 will investigate collective affective states within an immersive spatial installation. Comparative analysis across all three artefacts will develop the study's theoretical contribution: a critical framework for understanding algorithmic colonialism's affective reach, and a set of prototyped, symbiotic human-algorithm relational models that resist the disciplinary logic of emotional legibility. The full research will culminate in three publicly presentable artistic works and a 50,000–60,000-word thesis.

References

- Ahmed, S. (2004). *The cultural politics of emotion*. Edinburgh University Press.
- Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.
- Bulley, J., & Şahin, Ö. (2021). Report 1: What is practice research? British Library.
<https://bl.iro.bl.uk/concern/reports/b51c0f52-9801-49d9-9f00-cca89741091b>
- Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.
- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Foucault, M. (1977). *Discipline and punish: The birth of the prison*. Pantheon Books.
- Glickman, M., & Sharot, T. (2025). How human–AI feedback loops alter human perceptual, emotional and social judgements. *Nature Human Behaviour*, 9(2), 345–359. <https://doi.org/10.1038/s41562-024-02077-2>
- McStay, A. (2018). *Emotional AI: The rise of empathic media*. Sage.
- Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, 33(4), 659–684. <https://doi.org/10.1007/s13347-020-00405-8>
- Puar, J. (2017). *The right to maim: Debility, capacity, disability*. Duke University Press.

Rajcic, N., & McCormack, J. (2020). Mirror Ritual: An affective interface for emotional self-reflection. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1–14). ACM.

12.39 Parents Under Pressure: Social Media's Impact on Parental Wellbeing

Ali Shull (University of Nebraska-Lincoln)

Introduction

Social media platforms increasingly shape how everyday roles and identities are performed and evaluated, including the role of parenting. What began with early online forums has evolved into a dense ecosystem of algorithmically curated feeds, short-form videos, influencer advice, and peer commentary. Parents now encounter a constant stream of content about sleep training, discipline, emotional regulation, child development, and family routines. As a result, social media has become not only a source of advice but also a cultural environment where parenting norms are constructed, negotiated, and publicly evaluated (Amaro et al., 2019; Mertens et al., 2024).

At the same time, many parents are raising children with fewer traditional support structures than previous generations. With families living farther from extended relatives and entering parenthood later in life, access to intergenerational caregiving support has declined (Osterman et al., 2025). In this context, social media increasingly functions as a digital “village,” providing access to information, emotional reassurance, and communities of caregivers navigating similar challenges. Research suggests that online parenting spaces can reduce feelings of isolation and foster connection during the demanding early years of childrearing (Amaro et al., 2019; Glatz et al., 2023; Moujaes & Verrier, 2021).

However, these digital environments also expose parents to highly curated portrayals of family life and competing standards of what “good parenting” should look like. Parenting influencers play a particularly influential role in shaping these norms, presenting seemingly relatable scripts for feeding, discipline, developmental activities, and household organization (Ruiz-Gómez, 2019). Such portrayals invite social comparison, as parents evaluate their own caregiving practices against idealized depictions of attentive parenting and well-regulated children (Coyne et al., 2017; Ouvrein, 2024).

While parenting content can provide inspiration and practical ideas, it can also generate pressure or feelings of inadequacy when parents perceive gaps between their experiences and those displayed online. Emerging research links exposure to idealized parenting content with increased anxiety, lower parenting satisfaction, and diminished perceptions of competence (Germic et al., 2021; Wagstaff et al., 2024). At the same time, parents often report validation, reassurance, and connection through online

parenting communities. Together, these findings suggest that social media functions as a complex environment where support, comparison, and identity negotiation intersect in contemporary parenting.

Objectives/Research Questions

Little is known about how parents respond to parenting-related social media content in real time or how these reactions influence their sense of competence. To address this gap, this study examines the relationships among parenting-related social media exposure, physiological stress responses, emotional reactions, and parenting self-efficacy. A mixed-methods design integrates experimental, physiological, and qualitative approaches to capture both measurable stress responses and parents' interpretations of their experiences. Physiological responses are assessed through heart rate variability (HRV) while participants view social media content, alongside post-exposure measures of negative affect.

The experimental component compares parents exposed to parenting-related social media content with those viewing neutral crafting content. It is hypothesized that parents exposed to parenting content will exhibit stronger emotional and physiological stress responses, reflected in lower HRV and higher negative affect. The study also examines how these responses relate to parenting self-efficacy—parents' beliefs about their ability to effectively manage caregiving tasks and support their child's development. Drawing on social comparison and self-efficacy theory, exposure to idealized parenting portrayals may shape parents' self-evaluations by providing models against which they assess their own performance.

In addition, qualitative inquiry explores how parents interpret parenting content encountered online. Participants reflect on how social media influences their emotions, perceptions of other families, and interactions with their children. Thematic analysis identifies recurring patterns such as inspiration, validation, guilt, pressure, and reassurance. By integrating physiological data, emotional responses, and qualitative narratives, this study examines how digital parenting content shapes immediate stress reactions as well as broader interpretations of parenting identity and experience.

Methods

This study used a mixed-methods convergent design to examine how exposure to parenting-related social media content influenced parents' stress responses and parenting self-efficacy. Participants included 83 parents with at least one child ages 2–10 who reported using social media at least once per week. Parents were recruited through community outreach, including schools, childcare centers, pediatric clinics, community organizations, and targeted social media posts. All sessions were conducted in a lab at a Midwest university in the United States.

After providing consent, participants completed a pre-video survey assessing demographics, parenting stress, and state anxiety. Participants were then fitted with an Actiheart device to measure heart rate variability (HRV). A five-minute baseline HRV recording was collected while participants viewed neutral travel-related content. Participants were randomly assigned to one of two conditions: an experimental group that viewed parenting-related social media videos or a control group that viewed neutral crafting videos. Each group then viewed a 15-minute compilation of short-form social media content designed to resemble typical platform feeds.

Physiological stress responses were measured using HRV, specifically the Root Mean Square of Successive Differences (RMSSD), a time-domain measure sensitive to parasympathetic nervous system activity. Following the viewing task, participants completed a post-video survey measuring negative affect and parenting self-efficacy and participated in a semi-structured interview exploring their social media use, the parenting content they encountered online, and how it influenced their stress, confidence, and perceptions of parent–child interactions.

Quantitative analyses examined differences in physiological and emotional responses between conditions and associations between HRV and parenting self-efficacy. Repeated-measures ANOVA assessed changes in HRV across conditions, while correlations examined relationships among HRV, parenting stress, and self-efficacy. Regression analyses tested whether baseline parenting stress or social comparison orientation moderated these relationships. Qualitative data were analyzed using thematic analysis to identify patterns in parents' descriptions of how social media influenced their emotions, parenting beliefs, and interactions with their children. Quantitative and qualitative findings were then integrated using joint displays to identify points of convergence and divergence across the two strands.

Results

Quantitative results indicated that exposure to parenting-related social media content produced meaningful emotional responses but did not produce distinct physiological stress patterns. Parents who viewed parenting-related content reported significantly higher levels of post-exposure negative affect compared with those who viewed neutral crafting content, suggesting that parenting content activated emotionally salient self-evaluative processes. However, physiological stress responses, measured through heart rate variability (HRV), did not significantly differ between experimental and control conditions. Although HRV decreased during video exposure overall, this pattern occurred across both conditions, indicating that short-form video engagement may elicit mild autonomic changes regardless of content type.

Contrary to the original hypothesis, physiological reactivity was not significantly associated with parenting self-efficacy. Lower HRV responses did not predict lower

confidence in parenting ability. Interestingly, negative affect showed a small positive association with parenting self-efficacy, suggesting that emotional activation may reflect engagement with personally meaningful parenting standards rather than diminished confidence. Parents who reported stronger emotional responses may have been those most invested in their parenting role.

Qualitative findings helped contextualize these patterns. Parents frequently described experiencing upward social comparison, guilt related to time, patience, and presence with their children, and pressure to meet highly visible standards of “good parenting.” At the same time, participants also reported moments of validation, inspiration, and intentional generational change when engaging with parenting-related content. Many described using social media as a space to reflect on their own childhood experiences and make deliberate efforts to “parent differently” than they were raised. In this way, digital parenting content often prompted reflection on parenting values, encouraged curation of the content they consumed, and provided reassurance from other parents navigating similar challenges.

Taken together, the findings suggest that parenting-related social media content operates less as a straightforward stressor and more as an emotionally activating environment where identity, comparison, and cultural expectations intersect. Rather than directly undermining parenting competence, digital parenting content appears to trigger reflective processes about what kind of parent individuals want to be, producing emotional responses that coexist with sustained confidence in their parenting abilities.

Future Work

Future research should examine parenting-related social media engagement in more naturalistic contexts. While the present study identified emotional responses to brief laboratory exposure, everyday social media use is fluid, self-directed, and embedded within the rhythms of family life. Ecological momentary assessment combined with passive digital tracking and wearable physiological monitoring could allow researchers to observe how parents respond to content during real-world scrolling episodes and whether emotional or physiological responses accumulate or recover over time.

Future studies should also examine interactive platform features. Many parents described using comment sections, peer responses, and credibility cues when evaluating parenting advice online. Experimental designs that incorporate interactive feeds or comment threads could clarify how social validation and community discourse shape emotional reactions to parenting content.

Expanding the diversity of caregiving populations is also critical. The present sample consisted primarily of White, highly educated mothers, which limits generalizability. Including fathers, single parents, lower-income caregivers, and racially and culturally

diverse families will help clarify how parenting expectations, digital access, and social comparison processes shape experiences with online parenting content.

These findings also highlight the need for greater awareness of how digital parenting environments shape parents' emotional experiences. Rather than framing social media as inherently harmful, future work should focus on how parents can engage with parenting content in more intentional ways while maintaining the informational and relational benefits of online communities. Researchers, practitioners, and platform designers may also consider how features such as infinite scroll, rapid clip sequencing, and algorithmic amplification influence users' regulatory capacity. Greater transparency around algorithmic content delivery and reflective design features—such as usage dashboards or pause prompts—may help mitigate potential strain.

Together, these directions point toward a more nuanced research agenda that moves beyond asking whether social media is beneficial or harmful. Instead, future work should examine when, how, and for whom digital parenting content amplifies stress, encourages reflection, or provides meaningful validation within contemporary parenting environments.

References

Amaro, L. M., Joseph, N. T., & de los Santos, T. M. (2019). Relationships of Online Social Comparison and Parenting Satisfaction among New Mothers: The Mediating Roles of Belonging and Emotion. *Journal of Family Communication*, 19(2), 144–156.

<https://doi.org/10.1080/15267431.2019.1586711>

Conde, R., & Casais, B. (2023). Micro, macro and mega-influencers on instagram: The power of persuasion via the parasocial relationship. *Journal of Business Research*, 158, 113708. <https://doi.org/10.1016/j.jbusres.2023.113708>

Coyne, S. M., McDaniel, B. T., & Stockdale, L. A. (2017). “Do you dare to compare?” Associations between maternal social comparisons on social networking sites and parenting, mental health, and romantic relationship outcomes. *Computers in Human Behavior*, 70, 335–340. <https://doi.org/10.1016/j.chb.2016.12.081>

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research*. SAGE Publications.

Germic, E. R., Eckert, S., & Vultee, F. (2021). The Impact of Instagram Mommy Blogger Content on the Perceived Self-Efficacy of Mothers. *Social Media + Society*, 7(3), 20563051211041649. <https://doi.org/10.1177/20563051211041649>

Glatz, T., Daneback, K., Alsarve, J., & Sorbring, E. (2023). Parents' feelings, distress, and self-efficacy in response to social comparisons on social media. *Journal of Child and Family Studies*, 32(8), 2453–2464. APA PsycInfo (2023-84763-001).

<https://doi.org/10.1007/s10826-023-02611-2>

Hurst, K. (2022, May 18). More than half of Americans live within an hour of extended family. Pew Research Center. <https://www.pewresearch.org/short-reads/2022/05/18/more-than-half-of-americans-live-within-an-hour-of-extended-family/>

Jones, T. L., & Prinz, R. J. (2005). Potential roles of parental self-efficacy in parent and child adjustment: A review. *Clinical Psychology Review*, 25(3), 341–363. <https://doi.org/10.1016/j.cpr.2004.12.004>

Kirkpatrick, C. E., & Lee, S. (2022). Comparisons to picture-perfect motherhood: How Instagram’s idealized portrayals of motherhood affect new mothers’ well-being. *Computers in Human Behavior*, 137, 107417. <https://doi.org/10.1016/j.chb.2022.107417>

Laborde, S., Mosley, E., & Thayer, J. F. (2017). Heart Rate Variability and Cardiac Vagal Tone in Psychophysiological Research—Recommendations for Experiment Planning, Data Analysis, and Data Reporting. *Frontiers in Psychology*, 8, 213. <https://doi.org/10.3389/fpsyg.2017.00213>

Mertens, E., ye, G., Beuckels, E., & Hudders, L. (2024). Parenting Information on Social Media: A Systematic Literature Review. *JMIR Pediatrics and Parenting*, 7. <https://doi.org/10.2196/55372>

Moujaes, M., & Verrier, D. (2021). Instagram Use, InstaMums, and Anxiety in Mothers of Young Children. *Journal of Media Psychology*, 33(2), 72–81. <https://doi.org/10.1027/1864-1105/a000282>

Ouvrein, G. (2024). Mommy influencers: Helpful or harmful? The relationship between exposure to mommy influencers and perceived parental self-efficacy among mothers and primigravida. *New Media & Society*, 26(4), 2295–2314. <https://doi.org/10.1177/14614448221086296>

Ruiz-Gómez, A. (2019). Digital Fame and Fortune in the age of Social Media: A Classification of social media influencers. *aDResearch ESIC*, 19(19), 8–29.

12.40 “Healing Your Heart”: The Production and Commodification of Emotions on Chinese Social Media

Congman Zhang (University of Macau)

Introduction

Against the background of economic pressure, employment instability and information overload, the public's attention to mental health and emotional health is increasing, and emotional consumption has become a necessity for people. This growing demand for mental health has promoted the rise and development of the "healing economy" (GWI, 2024). "Healing" refers to the social practice of individuals seeking calmness, certainty and ideal self through mediated and symbolic practices in the face of emotional distress and social anxiety (Yuan & Huang, 2023). Importantly, healing differs from therapy, which is grounded in professional psychological knowledge and expert governance (Madsen, 2014; Foster, 2016). Healing culture shifts emotional repair from clinical intervention to daily routines, personal experience, and cultural practices (Yuan, 2025), and digital platforms have accelerated its circulation and commercialization.

On Chinese social media, healing culture is increasingly produced through influencer aesthetics and everyday interactions. Influencers share emotional healing methods through visual symbols and gentle narratives, while users gain emotional identity and mental comfort through imitation and interaction. Industry analytics suggest that on REDNote (Xiaohongshu), posts and searches related to “healing” grew substantially in 2025, indicating the platform’s centrality in the dissemination of healing culture (Qiangua, 2025).

Despite the growing interest in the emergence and spread of healing culture on social media, academic research on this phenomenon remains limited. Most studies have focused on specific healing texts (Liang, 2022), subcultural practices (Floyd, 2023; Teng & Li, 2023), or user motivations for wellness applications (Gál, É et al., 2020; Yang, 2024), with limited attention to the interrelationships between affective logics and social mechanisms. Few studies have elucidated how healing transforms from a private behavior into a replicable and recyclable social practice intertwined with platform algorithms and consumerist logics.

Therefore, this study investigates digital healing culture on REDNote as a platformed emotional practice that is produced through influencer–user interaction, amplified by algorithmic circulation and market logics. It based on theory of emotional capitalism (Illouz, 2008; 2017) and affective practices (Wetherell, 2012; 2014) and employed digital ethnography and semi-structured interviews to examine how healing content

and practices emerge, circulate, and are experienced in everyday settings. One key contribution of this study is to bring different literature bodies into conversation, from political economy, emotional capitalism to Confucianism to examine the way emotion has been reconfigured from a private experience to a social behavior.

Objectives/Research Questions

This study aims to theorize and empirically examine digital healing culture as an affective-commercial formation on Chinese social media. Conceptually, it asks how healing is produced and circulated as a socially recognizable emotional ideal (e.g., calmness, peace, happiness), and how such ideals become commodified through influencer labor, platform interactions, and consumer markets. Empirically, it investigates how healing influencers and users co-construct healing as an affective practice on REDNote, and how this practice is shaped by cultural-political discourses in China, including Confucian self-cultivation and the state discourse of “positive energy.” This study will address the following questions:

RQ1: How do healing influencers on REDNote construct and commodify ideal emotional states through visual and narrative strategies?

RQ2: How do platform interactions transform individual emotional repair into collective affective practices?

RQ3: How do users negotiate the meanings of “healing” under the interplay among consumerist logic, Confucian culture, and the state discourse of “positive energy”?

This study builds on Illouz's theory of emotional capitalism (2017) and the concept of affective practices by Weatherall (2012) to explore the production, circulation and commercialization of healing culture on Chinese social media platforms. On social media platforms, emotional healing is constantly integrated into the logic of consumerism. Healing influencers set specific emotional states such as peace, calmness and happiness as ideal emotional states, and promote "healing necessities" such as healing crystals, aromatherapy and online courses, thereby transform "healing" into consumable emotional commodities. Additionally, based on the perspective of "affective practice" (Wetherell, 2012), emotion is no longer understood as a private inner experience, but a social behavior, which is generated, disseminated and reproduced through interaction and collective participation (Wetherell, 2014; Lehto, 2021). On the REDNote, influencers and users co-construct the healing culture through sharing, commenting and liking, transforming healing from personal emotional repair to an emotional behavior of social consultation and collective practice.

In the Chinese context, healing culture intersects with the Confucian tradition of “xiuxin yangxing” (修心养性) and the national discourse of “positive energy,” forming a complex practice that links individual well-being to collective harmony. “Xiuxin

yangxing” emphasizes the achievement of physical and mental balance through self-cultivation and inner cultivation (Yang, 2024), while “positive energy”, as a national discourse, emphasizes the dissemination of optimistic emotions and messages for social cohesion (Chen and Wang, 2020). Together with the neoliberal logic of self-governance, healing not only an individual's personal emotional management, but is also intertwined with broader social and ideological structures.

Methods

The study employs digital ethnography combining textual analysis and semi-structured interviews. To ensure conceptual precision and sampling rigor, it advances an operational definition of “healing influencers” as creators whose content systematically: (a) stages “healing” as an aspirational emotional state (e.g., peaceful, happiness); (b) offers repeatable repertoires of affect regulation (routines, guided practices, symbolic objects); and (c) is recognized and circulated as “healing” through platform vernaculars (hashtags, captions, and recurring audience responses). This definition does not rely on self-identification; rather, it captures how “healing” is performed and socially validated within the platform ecology.

Based on an initial ethnographic mapping of REDNote’s healing discourse, we identified five prominent healing influencers that repeatedly appear across healing-related searches, recommendations, and user discussions. These influencers represent five recurring styles, meditative, lifestyle-oriented, aesthetic, spiritual-oriented, and community-driven, which function as comparative cultural cases. We collected posts and associated comment threads from these accounts (approximately ~100 posts total, March–September 2025) to examine variation in emotional scripts, aesthetic strategies, and commodification routes, and to explore how distinct healing practices become aligned with consumer value, moral repertoires, or normative affect (Loveday, 2016).

We then conducted thematic analysis (Braun & Clarke, 2006) to code how ideal emotional states and feeling rules are articulated, how visual-symbolic strategies materialize calmness (e.g., cozy domestic scenes, slow-life routines, sensory objects), and how commercialization is embedded through product placements and paid courses framed as self-care. Because platform algorithms are not directly observable, we analyze qualitative traces of visibility incentives, including standardized “how-to” templates, strategic hashtag, call-to-action phrasing, and user talk about being recommended or discovered, as well as the extent to which certain emotional scripts become repeatable and widely imitated.

In addition, the project is conducting approximately ten interviews with active users who engage healing content to understand motivations, perceived emotional effects, attitudes toward commercialization, and how “healing” is interpreted in relation to

Confucian self-cultivation and the discourse of “positive energy.” Interview transcripts will be thematically analyzed and triangulated with post/comment analysis. All identifiable information is anonymized and reported in harm-minimizing ways.

Results

Preliminary findings show that healing on REDNote works as a repeatable genre. Influencers present peaceful and stability as goals. They use warm home scenes, slow routines, and sensory objects (e.g., crystal, candles, aromatherapy). Posts often mix short personal stories with gentle “how-to” tips. These frames healing as daily self-work.

Additionally, comments turn healing into a shared practice. Users ask for comfort, give advice, and report small changes in mood. Many repeat the same words and rules (e.g., “be gentle,” “avoid negativity”). Some users copy routines and styles in their own posts. In this way, comfort is made through interaction and attention.

Finally, commercial content is common but not always direct. Products and paid courses are framed as “healing tools.” Buying is described as caring for the self and doing the work. At the same time, some users doubt the value of paid items and argue about what “real healing” means.

Future Work

We will extend this work in three directions. First, we will complete semi-structured interviews with active users to understand their emotional experiences and motivations. We will ask how they define “healing,” how they engage with healing content, and how it affects their daily emotional life. Second, we will further analyze the five influencer styles and their different focuses. From each style, we will choose one representative influencer for closer tracking. We will collect high-engagement posts and their comment threads and replies. We will also save screenshots and short video clips when needed, to support analysis of multimodal healing aesthetics. Finally, we will conceptualize the conclusions by refining “platformed healing” and its key features, and by building a brief typology of healing styles. We will explain how everyday participation turns healing into a shared practice, and how commercial value becomes attached to emotional relief in a China-specific moral setting shaped by self-cultivation and relational harmony.

References

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
<https://doi.org/10.1191/1478088706qp063oa>

- Chen, Z., & Wang, C. Y. (2020). The Discipline of Happiness: The Foucauldian Use of the “Positive Energy” Discourse in China’s Ideological Works. *Journal of Current Chinese Affairs*, 48(2), 201–225. <https://doi.org/10.1177/1868102619899409>
- Floyd, RJ (2023). Understanding the Psychological Significance of Astrology in Millennial Women’s Lives. Doctoral Dissertation, Duquesne University, Pittsburgh, PA.
- Foster, R. (2016). Therapeutic culture, authenticity and neo-liberalism. *History of the Human Sciences*, 29(1), 99–116. <https://doi.org/10.1177/0952695115617384>
- Gál, É., Ștefan, S., & Cristea, I. A. (2020). The efficacy of mindfulness meditation apps in enhancing users’ well-being and mental health related outcomes: a meta-analysis of randomized controlled trials. *Journal of Affective Disorders*, 279(1), 131–142. <https://doi.org/10.1016/j.jad.2020.09.134>
- GWJ (2024). Global Wellness Economy Monitor. Available at: <https://globalwellnessinstitute.org/industry-research/2024-global-wellnesseconomy-monitor/>
- Hochschild, A. R. (2012). *The Managed Heart: Commercialization of Human Feeling*. University of California Press.
- Illouz, E. (2017). *Saving the modern soul: Therapy, emotions, and the culture of selfhelp*. University of California Press.
- Lehto, M. (2021). Ambivalent influencers: Feeling rules and the affective practice of anxiety in social media influencer work. *European Journal of Cultural Studies*, 25(1), 201–216. <https://doi.org/10.1177/1367549421988958>
- Liang, L. (2022). Consuming the Pastoral Desire: Li Ziqi, Food Vlogging and the Structure of Feeling in the Era of Microcelebrity. *Winter 2021*, 1(2), 7–39. <https://doi.org/10.3998/gs.1020>
- Madsen, O. J. (2014). *The therapeutic turn: How psychology altered Western culture*. Routledge.
- QianGua Data (2025). “Interest-based lifestyle” data insight report (Xiaohongshu platform). Available at: <https://www.qian-gua.com/information/detail/3197>
- Teng, L., & Li, Y. (2023). Emotional cocoons and self-resistance: A qualitative study on media consumption of young female astrology enthusiasts. *News and Writing*, (8), 83–95. <https://doi.org/CNKI:SUN:XWXZ.0.2023-08-009> (in Chinese)
- Wetherell M (2012) *Affect and Emotion: A New Social Science Understanding*. London: Sage.

Wetherell, M. (2014). Trends in the Turn to Affect: A Social Psychological Critique: A Social Psychological Critique. *Body & Society*, 21(2), 139-166.

<https://doi.org/10.1177/1357034X14539020>

Yuan, G. (2025). “Take care of yourself”: Digital healing culture as a technology of the self. *Nanjing Social Sciences*, (5), 98–106.

<https://doi.org/10.15937/j.cnki.issn1001-8263.2025.05.010> (in Chinese)

12.41 Whose Voice? Power, Geopolitics, and Narrative Dynamics in the #ChineseVirus Discourse

Qian Zhao (University of Strathclyde)

Introduction

Existing research mostly understands #ChineseVirus as a racialised naming practice pushed within the US political context and containing anti-Asian stigma and hate expressions. The focus usually revolves around political elites such as Trump and their diffusion effects on Twitter/X (Budhwani & Sun, 2020; Hswen et al., 2021; Chong & Chen, 2021). However, analysis of 238,687 tweets collected via the Twitter API reveals that the transnational use of this hashtag is far more active than existing literature discloses. India-related narratives occupy an unusually prominent position in the overall discourse structure, a transnational dimension that has received almost no systematic attention. Indian discourse under #ChineseVirus operates with its own independent narrative logics, not as a follower of the US framework. Yet West-centric presuppositions in much existing research obscure the true face of geopolitical competition within Asia (Cooper & Lampropoulou, 2024).

Beyond asking what this label does, there is a more fundamental structural question: who is truly controlling this discourse? In a space with over 70,000 participants, visibility and narrative power are distributed unequally. Even four years on, the pattern of rapidly directing public crises toward state confrontation is not unique to COVID-19. In discussions of the Russia-Ukraine war, Israel-Palestine conflict, and the current U.S.-Israeli war on Iran (Reuters, 2026), people quickly take opposing sides along ideological lines (Loru et al., 2025). #ChineseVirus is not just a pandemic-era label, it reveals how the "state accountability" framework in crisis discourse is reinforced and serves political agendas across different national contexts.

Objectives/Research Questions

This study examines the transnational discursive structure and power distribution of #ChineseVirus on Twitter/X, guided by two core research questions.

RQ1: Which recurring narrative frames are consistently mobilised in #ChineseVirus discourse, and through what recognisable forms of expression are they constructed?

RQ2: Who controls discursive power within the #ChineseVirus discourse space, and how is power distributed across different user types and national contexts?

Methods

The dataset comprises 238,687 #ChineseVirus tweets collected via the Twitter API (March 2020–January 2021), retaining tweet text, timestamps, user identifiers, and like counts. Basic cleaning included URL removal, lowercasing, and noise filtering.

A mixed-methods strategy combines quantitative pattern recognition with qualitative contextual interpretation. Word co-occurrence analysis extracts high-frequency word pairs to identify recurring narrative combinations. Comparative text versions—retaining and removing #ChineseVirus—assess the hashtag's structural effect on narrative focus. Country visibility is estimated via place name mentions in tweet text, with variant forms standardised (e.g., "U.S."/"USA"/"America" → "United States").

For RQ2, discursive power is measured by like counts (Gillespie, 2014) as a proxy for visibility and audience interaction. The top 20 users by total likes were analysed across verification status, follower count, user type (politicians, media, cultural influencers, activists), and national background. High-engagement tweets were subject to critical close reading to examine discursive strategy differences between US and Indian dominant actors.

Results

Three key findings emerge. First, comparing retained vs. removed #ChineseVirus shows the hashtag is not a neutral marker but restructures discourse. With the hashtag retained, "Stigma/Political attack" word pairs account for ~54% and "Geopolitical" ~12%. When removed, "Stigma/Political attack" drops to ~46% while "Geopolitical" rises to ~30%. The hashtag compresses complex geopolitical discussion into a direct blame logic.

Second, place mention analysis identifies a China–India–US triangular reference structure. China appears most frequently (27.69%), followed by India (6.54%) and the US (5.87%)—India's visibility exceeds the US, a pattern largely absent from existing literature. India-related discussion peaked in June (13.69%), September (12.03%), and November 2020 (11.84%), aligning with the Galwan Valley clash and Chinese app bans, confirming event-driven, geopolitically independent participation.

Third, discursive power is extremely unequal: 20 users (0.027% of 74,124 participants) captured ~40% of total likes. Verified accounts receive 21.8x more interactions per tweet than unverified accounts, suggesting verification functions as an algorithmic advantage. The 20 most-liked tweets came from just 9 users; @ReaJamesWoods contributed five, @ShefVaidya four. Dominant voices came almost exclusively from right-wing and nationalist actors in the US and India. US elites weaponised #ChineseVirus as a partisan tool centred on Trump; Indian elites embedded it in border security and Hindu nationalist frameworks. Notably, China, the discourse's core subject, is almost entirely absent as a producer despite being its most-mentioned

country, reflecting combined effects of platform blocking, language barriers, and Twitter's institutional labelling of Chinese state media.

Future Work

This study focuses on publicly accessible tweet text and excludes retweets, quote tweets, and reply chains. Future research integrating these interaction layers would help track how accountability narratives are amplified or challenged over time, and how different user categories—politicians, media professionals, activists—gain influence across platform positions. More broadly, this study raises a question worth further investigation: whether platform verification mechanisms and algorithmic amplification are structurally biased toward specific geopolitical voices, and whether similar patterns of power concentration appear in other transnational crisis hashtags.

References

- Budhwani, H., & Sun, R. (2020). Creating COVID-19 stigma by referencing the novel coronavirus as the "Chinese virus" on Twitter: Quantitative analysis of social media data. *Journal of Medical Internet Research*, 22(5), e19301. <https://doi.org/10.2196/19301>
- Chong, M., & Chen, H. (2021). Racist framing through stigmatized naming: A topical and geo-locational analysis of #Chinavirus and #Chinesevirus on Twitter. *Proceedings of the Association for Information Science and Technology*, 58(1), 70–79. <https://doi.org/10.1002/pr2.437>
- Cooper, P., & Lampropoulou, S. (2024). China virus, Kung Flu, and MAGA: Countervailing values and sociological fractionation on Twitter as evidenced by pro- and anti-Trump discourses in relation to Covid-19. *Discourse, Context & Media*, 57, 100758. <https://doi.org/10.1016/j.dcm.2024.100758>
- Hswen, Y., Xu, X., Hing, A., Hawkins, J. B., Brownstein, J. S., & Gee, G. C. (2021). Association of "#covid19" versus "#chinesevirus" with anti-Asian sentiments on Twitter: March 9–23, 2020. *American Journal of Public Health*, 111(5), 956–964. <https://doi.org/10.2105/AJPH.2021.306154>
- Loru, E., Galeazzi, A., Bonetti, A., et al. (2025). Ideology and polarization set the agenda on social media. *Scientific Reports*, 15, 35816. <https://doi.org/10.1038/s41598-025-19776-z>
- Reuters. (2026, March 13). Both sides trade strikes and threats as the Iran war approaches the two-week mark. Reuters. <https://www.reuters.com>

12.42 (Re)listening to a Better Day on YouTube: Memory, Affect, and Hope in Global Charity Singles from Record to AI

Jingyu Zhu (Nanjing University)

Introduction

Synthetic Hope by AI?

In 2026, searching for We Are The World on YouTube reveals a series of digital remediation: world leaders singing together in harmony. These AI-generated videos, as described by their creators, have sparked heated discussions ranging from nostalgic praise and peaceful wishes to ethical critique. When opposing politicians appear to chorus for a “better day”, the song transcends its origin as a charity single, serving as a form of political irony, or what might be termed technological synthetic hope.

YouTube has been the primary stage for this song’s social life, from the original 1985 recording and the 2010 WATW 25 for Haiti, to the AI-generated versions of 2025. Global users do not merely consume a song, but react, comment, and interpret. They transform the listening into a memorial and affective action. The phenomenon is deeply rooted in the digital platforms and sociocultural background.

YouTube, Music and Platform Affordances

Initially designed for musical content, YouTube has become one of the most influential platforms for music consumption. It continues to be a major force for music while providing updated content, user interaction (through comments and likes...), political participation, and forms of remediation and rescripting (Rogers et al., 2023; Burgess & Green, 2018; Georgakopoulou, 2015). The platform and its interfaces can be regarded as multi-authored and multimodal texts (Benson, 2016). On this stage, users create various types of music videos including traditional, user-appropriated, derivative videos and reaction videos, etc. (Liikkanen & Salovaara, 2015; McDaniel, 2021). Their practices, along with those of platforms and industries, co-contribute to global mobility (Mjos, 2013), optimization of culture (Morris, 2020) and decontextualization (Sadler, 2022).

Charity Singles and Discourses

Charity singles are intended to raise funds for people in need, but they actually exist within much more complex contexts. They can serve as a response to crisis while providing individualized solutions based on political, economic and moral situations (Robinson, 2018). From a global perspective, the growth of charity singles has been associated with disasters such as the famine in Ethiopia and the earthquake in Haiti. Since 1984, public figures from Europe and North America have led initiatives like Band

Aid, Live Aid for Africa. These actions were discourse practices drawing praise and criticism. As criticism, the singles exhibited constitutional discourses, and use of a collective “we” would overlook other cultures and their own efforts (Howes, 1990; Yúdice, 1992). The notions seemingly imply that the global north is superior to the south, “otherizing” and decontextualizing developing countries (Grant, 2015; Walonen, 2022). Furthermore, some local residents join these artworks to ensure their voices are not erased by Global North celebrities (Boutros, 2021).

Objectives/Research Questions

Literature Review and Analytical Framework: Memory, Affect, and Hope Musical Memory and Digital Media

Pop music can be the object of memory or a vehicle for both individual and collective memory (Van Dijck, 2009; Bennett & Janssen, 2016). Moreover, social media plays a role in digital archives for constructed memory and discourses, reconstructing the memorial resources (Makhortykh, 2020; Smit, Heinrich & Broersma, 2017). Music in the digital media contributes to a double mediation of memory. Consequently, platforms like YouTube hosting music videos create complex spaces for transforming private and public memory, fostering “connective memory” (Khiun, 2014; Hoskins, 2011; Bar-Gil, 2025). The forty-year span since the original release links multiple layers of memory to these songs.

Listening and Affect

Listening is not only about the sounds and lyrics, it also produces affective responses (Kassabian, 2013). Listening to music simultaneously engages with the representational meanings while sensing nonrepresentational embodied vibrations. Affect is expressed and spread through music, offering a critical lens into the political projects and ideologies (Gray, 2020; Desai-Stephens & Reissour, 2020). And music could be regarded as the historical archives to be felt (Long, Baker, Istvandity & Collins, 2017). In addition, the publics connected by digitally-mediated feelings should be considered (Papacharissi, 2016). In this context, listening to charity singles becomes a deeply affective practice.

Politics of Hope and Memes

Hope promises something better in the future. Politics can operate by building up and offering hope to someone to tolerate the present, reconfiguring despair and expectation (Lindroth & Sinevaara-Niskanen, 2019; Warren, 2015). The Internet has become a field where politics plays a role and people voice their opinions. Take memes for instance, they are a rush of affect and playful performance of humour and politics (Griffin, 2021; Mortensen & Neumayer, 2021). It is a means through which ordinary people define meaning and join in political action during unstable times (Bouchard & Antsybor, 2025).

AI-generated charity singles calling for a better world show politics of hope within meme culture.

Case Study: We Are The World (original and remediated versions) on YouTube

Research Questions:

RQ1: How do the different versions of WATW and user practices (e.g., comments, reaction videos) on YouTube construct collective and individual memory from original release to AI versions?

RQ2: How do global listeners manifest affective responses to these singles on YouTube, and how do these engagements constitute or challenge the formation of “affective publics”?

RQ3: How have the discourses of global hope in WATW transformed over 40 years, particularly, what kind of political work does the “hope” performed by its AI-generated versions do?

Methods

Methodology: Multimodal Critical Discourse Analysis

The research is primarily based on multimodal critical discourse analysis (MCDA), supplemented by digital ethnography to capture the nuanced dynamics.

MCDA is the extension of the critical discourse analysis tradition of Fairclough and Van Dijk, beyond linguistic texts to multimodal resources. This study adopts four key analytical dimensions: visual analysis, sonic analysis, embodied modes and social media roles (Machin & Mayr, 2023; Machin & Richardson, 2012; Teo, 2021; Kopf, 2025). For charity single videos, the analysis, mainly from audiovisual dimensions, focuses on the images, visual grammar, acoustic features, lyrics, etc. For reaction videos, I examine the embodied responses of the creators. User comments are situated within their social media contexts.

Digital ethnography (from Nov. 2025 to Jan. 2026) serves as a complementary approach to unpack the holistic contextualisation and rethink the digital environment for the operation of discourses (Miller et al., 2016; Pink, et al., 2016). YouTube is treated as a virtual field where users engage with videos, comments, forums, etc.

Data collection

The research employs purposive sampling, including:

Music Videos (n=10): the 1985 original recording, the 2010 Haiti version, and 8 AI-generated versions (by “Most Liked”).

Reaction Videos (n=10): Almost all available search results for AI-version reactions (n=2), the 8 videos for the original recording.

Comments (n=1429): 200 comments each are sampled from the 1985 and 2010 versions (100 by “Most Liked” and 100 by “Newest first”). 100 “Most Liked” comments are collected from each of the 8 AI-generated videos. 30 comments are collected from each of the 10 reaction videos. (Estimated to be 1500, but eventually 1429 valid samples, due to limited data availability till Jan. 2026)

Results

Current general results are as follows:

In Search of Lost Time: Pop Stars, Golden Age and Personal Youth

The singles and comments function as sites of memory for past eras and generational icons like Michael Jackson, Celine Dion. Listeners evoke the 1980s/90s, mourning an age when their loved ones were still alive. They revisit these recordings to leave future-oriented digital signatures, for instance, “Here in January 2026” marking their presence in the new year. Crucially, private narratives are shared by those who witnessed the original events. One comment reads “I was that starving child in 1984”, recalling lived experience.

Touched by Spirits of Care and Myths of “Great Unity”

Moved feelings stem from inspiring rhythm and lyrics, the spirits of caring for all human beings and myths of great global unity, or for some, a nostalgia for American cultural leadership. The affect is transmitted through diverse accents, singing styles, multiethnic imagery, yet it does not form a united public. While 1985 and 2010 versions centered on optimistic globalization and a “kind America”, the AI versions receive complex resonance: users laughed, prayed, praised, cried, and were shocked by AI. Notably, these emotional reactions are sometimes an affective performance within the YouTube context.

For Whom Does Hope Sing? Others, Ourselves, One Unreal Day

Hope in earlier versions offered specific individual solutions to support the people under natural disasters. It reinforced a belief that “tomorrow will be better” and people could make it over decades. However, this hope remains the “golden age”. In AI-generated world, Trump and Putin sing “When the world must come together as one.” in unison, here “words” replaces “worlds”, and arranged lyrics contradict the politicians’ actual deeds. Peace and reconciliation are symbolized. “Best use of AI of all times.” heals the hearts in 2026, but “you begin to feel sad when realizing this is not possible to be real”.

Future Work

Future research involves a deeper multimodal analysis of the resources, supplemented by historical archives. Ultimately, it reflects a scholarly and personal wish for a “better day” beyond the interfaces.

References

- Bar-Gil, O. (2025). Holocaust remembrance in the digital age: The transformative influence of technology, digital archives, and connective memory. *Memory Studies*, 17506980241312341.
- Bennett, A., & Janssen, S. (2016). Popular music, cultural memory, and heritage. *Popular music and Society*, 39(1), 1-7.
- Benson, P. (2016). *The discourse of YouTube: Multimodal text in a global context*. Routledge.
- Bouchard, M., & Antsybor, D. (2025). Memeing war: the use of humor for hope, resistance, and forging the nation. *National Identities*, 1-34.
- Boutros, A. (2021). “Sounds like Haiti”: Haiti as Muse in Canadian Popular Music. *Popular Music and Society*, 44(5), 504-522.
- Burgess, J., & Green, J. (2018). *YouTube: Online video and participatory culture*. John Wiley & Sons.
- Desai-Stephens, A., & Reissour, N. (2020). Musical feelings and affective politics. *Culture, Theory and Critique*, 61(2-3), 99-111.
- Georgakopoulou, A. (2015). Sharing as rescripting: Place manipulations on YouTube between narrative and social media affordances. *Discourse, Context & Media*, 9, 64-72.
- Grant, J. (2015). Live Aid/8: perpetuating the superiority myth. *Critical Arts*, 29(3), 310-326.
- Gray, L. E. (2020). Listening for affect: musical ethnography and the challenge of/to affect. *Culture, Theory and Critique*, 61(2-3), 319-337.
- Griffin, H. (2021). Living through it: Anger, laughter, and internet memes in dark times. *International Journal of Cultural Studies*, 24(3), 381-397.
- Hoskins, A. (2011). 7/7 and connective memory: Interactional trajectories of remembering in post-scarcity culture. *Memory studies*, 4(3), 269-280.
- Howes, D. (1990). “We are the world” and its counterparts: Popular song as constitutional discourse. *International Journal of Politics, Culture, and Society*, 3(3), 315-339.

- Kassabian, A. (2013). *Ubiquitous listening: Affect, attention, and distributed subjectivity*. Univ of California Press.
- Khiun, L. K. (2014). Rewind and recollect: Activating dormant memories and politics in Teresa Teng's music videos uploaded on YouTube. *International Journal of Cultural Studies*, 17(5), 503-515.
- Kopf, S. (2025). Unravelling social media critical discourse studies (SM-CDS)–four approaches to studying social media through the critical lens. *Critical Discourse Studies*, 1-18.
- Liikkanen, L. A., & Salovaara, A. (2015). Music on YouTube: User engagement with traditional, user-appropriated and derivative videos. *Computers in human behavior*, 50, 108-124.
- Lindroth, M., & Sinevaara-Niskanen, H. (2019). Politics of hope. *Globalizations*, 16(5), 644-648.
- Long, P., Baker, S., Istvandy, L., & Collins, J. (2017). A labour of love: the affective archives of popular music culture. *Archives and Records*, 38(1), 61-79.
- Machin, D., & Mayr, A. (2023). *How to do critical discourse analysis: A multimodal introduction*. SAGE Publications.
- Machin, D., & Richardson, J. E. (2012). Discourses of unity and purpose in the sounds of fascist music: A multimodal approach. *Critical Discourse Studies*, 9(4), 329-345.
- Makhortykh, M. (2020). Remediating the past: YouTube and second world war memory in Ukraine and Russia. *Memory Studies*, 13(2), 146-161.
- McDaniel, B. (2021). Popular music reaction videos: Reactivity, creator labor, and the performance of listening online. *New Media & Society*, 23(6), 1624-1641.
- Miller, D., Costa, E., Haynes, N., McDonald, T., Nicolescu, R., Sinanan, J., Spyer, J., Venkatraman, S., & Wang, X. (2016). *How the World Changed Social Media*. London, UK. UCL Press.
- Mjos, O. J. (2013). *Music, social media and global mobility: MySpace, Facebook, YouTube*. Routledge.
- Morris, J. W. (2020). Music platforms and the optimization of culture. *Social Media+ Society*, 6(3), 2056305120940690.
- Mortensen, M., & Neumayer, C. (2021). The playful politics of memes. *Information, Communication & Society*, 24(16), 2367-2377.
- Papacharissi, Z. (2016). Affective publics and structures of storytelling: Sentiment, events and mediality. *Information, communication & society*, 19(3), 307-324.

- Pink, S., Horst, H., Postill, J., Hjorth, L., Lewis, T., & Tacchi, J. (2016). *Digital Ethnography: Principles and Practice*. SAGE Publications Ltd.
- Robinson, L. (2018). Putting the charity back into charity singles: Charity singles in Britain 1984–1995. In *Youth Culture, Popular Music and the End of 'Consensus'* (pp. 141-161). Routledge.
- Rogers, H., Freitas, J., & Porfírio, J. F. (2023). *Youtube and music: Online culture and everyday life*. Bloomsbury Academic.
- Sadler, O. (2022). Defiant amplification or decontextualized commercialization? Protest music, TikTok, and social movements. *Social Media+ Society*, 8(2), 20563051221094769.
- Smit, R., Heinrich, A., & Broersma, M. (2017). Witnessing in the new memory ecology: Memory construction of the Syrian conflict on YouTube. *New media & society*, 19(2), 289-307.
- Teo, P. (2021). 'It all begins with a teacher': A multimodal critical discourse analysis of Singapore's teacher recruitment videos. *Discourse & Communication*, 15(3), 330-348.
- Van Dijck, J. (2009). Remembering songs through telling stories: pop music as a resource for memory. In K. Bijsterveld, & J. van Dijck (Eds.), *Sound souvenirs: audio technologies, memory and cultural practices* (pp. 107-119). (Transformations in art and culture). Amsterdam University Press.
- Walonen, M. K. (2022). An Idea of Africa in "We Are the World": Decontextualization, Celebrity Savior Complex, Race, and Neoliberal Charity. *The Journal of Popular Culture*, 55(3), 561-574.
- Warren, C. L. (2015). Black nihilism and the politics of hope. *CR: The New Centennial Review*, 15(1), 215-248.
- Yúdice, G. (1992). We are not the World. *Social Text*, (31/32), 202-216.