

2021

Conceptualising Social Bots for Countering Online Extremist Messages

Kevin M. Blasiak

The University of Queensland, k.blasiak@business.uq.edu.au

Marten Risius

The University of Queensland, m.risius@business.uq.edu.au

Sabine Matook

The University of Queensland, s.matook@business.uq.edu.au

Follow this and additional works at: <https://aisel.aisnet.org/acis2021>

Recommended Citation

Blasiak, Kevin M.; Risius, Marten; and Matook, Sabine, "Conceptualising Social Bots for Countering Online Extremist Messages" (2021). *ACIS 2021 Proceedings*. 81.

<https://aisel.aisnet.org/acis2021/81>

This material is brought to you by the Australasian (ACIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ACIS 2021 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Conceptualising Social Bots for Countering Online Extremist Messages

Research-in-progress

Kevin M. Blasiak

UQ Business School
The University of Queensland
Brisbane, Australia
Email: k.blasiak@business.uq.edu.au

Marten Risius

UQ Business School
The University of Queensland
Brisbane, Australia
Email: m.risius@business.uq.edu.au

Sabine Matook

UQ Business School
The University of Queensland
Brisbane, Australia
Email: s.matook@business.uq.edu.au

Abstract

Online extremism and radicalisation on social media (SM) are significant concerns of governments, SM companies, and society. The 2021 attack on the US Capitol illustrates the severity of extremism fuelled through SM communications. The literature suggests the removal of extremist messages from SM to limit online extremism. However, scholars argue that these interventions are ineffective in containing the threat of extremist messages on SM. This study draws on dual-process theory and reactance theory to conceptualize the factors that contribute to limiting online extremism. Our model proposes cognitive factors and socio-technical factors that impact how SM users respond to online extremist messages. The model is tested with an Artificial Intelligence (AI)-based automated software agents (bots). The research contributes to a novel understanding of bots as social bots that are programmed as interventions to extremism.

Keywords Social media communication, social bot, extremism, dual-process theory, reactance theory

1 Introduction

The threat to society through extremist messages on social media (SM) is a primary concern for governments, SM companies, and societies worldwide. Extremists use the SM environment and its reach to broadcast extremist propaganda. Through this communication, extremists seek to persuade users to accept violence as a necessary means of bringing about change. These calls for violence are undemocratic and threaten democratic values. Australia's Director-General of Security, Mike Burgess, recently highlighted the importance of online counter-extremism, identifying the threatening online environment as a key multiplier for extremism (Australian Security Intelligence Organisation 2021).

The literature suggests removing extremist messages from SM to limit online extremism and using counter-narratives (Winter et al. 2020). However, others have argued that these interventions are inept and impractical to counter extremist messaging (Reed et al. 2017). Indeed, manual or keyword-based removal is coarse and lacks contextual focus (Thomas et al. 2021). Moreover, the interventions are often too slow (Marcellino et al. 2020) to stop the radicalisation process (Youngblood 2020). Moreover, despite current interventions, extremists successfully target susceptible users with subtle radicalisation methods that are difficult to identify and thus rarely removed (Marcellino et al. 2020).

The latest research in Artificial Intelligence (AI) presents social bots as human-designed and developed software that operates as autonomous agents (Maruping and Matook 2020). Social bots are used on SM to undertake communicative actions with users based on AI algorithms. For example, these bots are capable of engaging in dyadic conversations with users (Seering et al. 2019), and like, comment, or create content independently. These abilities enable them to take on essential roles on SM, such as content moderator (Seering et al. 2018; Seering et al. 2019), learning facilitator (Winkler and Söllner 2018), and medical assistants (Dennis et al. 2020).

Social bots can be used for interventions with extremist messages on SM. Bots have shown to be efficient in changing user attitudes (Zarouali et al. 2020) and creating higher persuasiveness due to their non-judgemental nature (Dennis et al. 2020). Their ability to engage with other users in dyadic conversation in comments to postings makes them approachable for users (Seering et al. 2019); even in challenging contexts (Winkler and Söllner 2018). In addition, bots offer scalability, thus having an economic advantage over human intervention (Mittal et al. 2016).

Research on human thinking suggests that information, as we would find it in messages on SM, is cognitively processed in two ways: Spontaneously implicit and consciously explicit (Wilson et al. 2000). Dual-process theory (DPT) theorizes two systems that interact with each other. DPT suggests that these two systems (System 1 and System 2) are complementary in forming attitudes. The implicit System 1 is receptive to peripheral cues (e.g., message source, visual cues, and social proof), is fast, and requires low cognitive effort. In contrast, System 2 is effortful, slow, and detailed in elaborating message content and context (Kahneman 2011). For instance, messages suggesting a threat are exceptionally efficient in motivating a more elaborate cognitive response (Boss et al. 2015). Reactance theory (RT) suggests that these messages can also lead to attitude change (Dillard and Shen 2005).

This research integrates DPT and RT to conceptualise social bots as an effective cognitive intervention to attitude radicalisation. Current research suggests that deliberate triggering of System 1 and System 2 responses SM is difficult (Moravec et al. 2020). We argue that the dyadic conversation mode and the overall persuasiveness of social bots make them particularly suitable for deliberately triggering distinct System 1 and System 2 responses. Moreover, social bot interventions that use fear appeal messages are arguably particularly suitable to influence attitude change (Boss et al. 2015), foster reactance to extremist messaging, and thus counter its radicalising effects on the user. To explore this assumption and test we address the research question of "*How can social bots counter extremist messaging?*".

The remainder of the paper discusses the present literature on online radicalisation, shortcomings of current extremism interventions, and finally, the potential role of social bots in the unique challenge of interfering with extremist messages in socially complex multiparty online conversations. Furthermore, we propose an empirical research design applying DPT and RT to test the efficacy of social bot interventions to extremist messaging.

2 Background

2.1 Online Radicalisation through Extremist Content

Extremists entice users to increasingly support violence based on undemocratic ideologies (Jasko et al. 2017). Radicalisation is a process and draws on three main determinants (3N): first, an individual's

need and universal desire for personal significance, second, guiding *narratives* for the consequent “quest for significance”, and third, a *network* that validates the collective narrative that justifies violence (Kruglanski et al. 2019).

Extremists use different persuasion tactics (Saleh et al. 2021) in their messages to radicalise SM users and plant attitudes that support the necessity for violence. Extremists apply strategies including persuasive narratives, appeal to emotion and imagery, and fear to convey an extremist message (Braddock 2020; Frischlich 2020). These strategies target the three determinants of extremism to radicalise users and create positive attitudes towards extremist beliefs, behaviours, and emotions (Schmuck et al. 2020). Content that fosters extremist attitudes is particularly present on popular SM platforms like YouTube (Schmitt et al. 2018) and Instagram (Frischlich 2020).

Current Shortcomings and Prospects of Counter Violent Extremism (CVE) Measures

Content intervention to prevent radicalisation is complex due to online moderation's required scale and cost (Jee 2020), but more so due to limited theoretical and empirical foundations (Glazzard 2017). CVE methods like counter-narratives and information campaigns primarily disrupt demand for extremist content but lack empirical evidence of their efficiency (Reed and Ingram 2018). Academics and civic organisations frequently challenge current content moderation techniques for opaque moderation practices (Douek 2020), human rights violations (Human Rights Watch 2020), risks of collateral damage through misclassified content removal (The Economist 2020). Alternative approaches that mitigate the risk of extremist message exposure, like warning labels, are still scarce and limited in their effectiveness to prevent radicalisation (Pennycook et al. 2020).

This criticism reveals that CVE measures online cannot rely solely on technical approaches to limiting the supply of extremist messages. Radicalisation intervention has social components. This includes the user demand for radical content, the individual context, susceptibility to external intervention (Miron and Brehm 2006), and the emotional state of users and associated cognitive processes during content exposure. These components add to the complexity of interventions. Cognitive psychology research recognises the potential of cognitive interventions to address the shortcomings of technical interventions (Baugut and Neumann 2020; Nuraniyah 2018). Attitude inoculation to extremist persuasion messages is an emerging approach to prevent the radicalising effects of extremist online propaganda (Saleh et al. 2021). Other research explores the preventative effects of visual and textual cues in other users' comments and messages and users' literate ability to process information online (Vraga et al. 2021). This stream of research holds valuable practical implications for the requirements of forthcoming CVE intervention methods.

Social Bot Interventions as Opportunity in CVE

Our research focuses on social bots as an intervention approach that seems particularly promising to actively counter extremist messaging. Bots are automated computer algorithms that produce content and interact autonomously with human SM users in dyadic conversations (Seering et al. 2019). Social bots are particularly suitable for changing users' perceptions, overcoming polarising content bubbles (Zarouali et al. 2020), and lowering users' threshold for engaging with opposing or conflicting views (Dingler et al. 2018). Thus, AI bots encourage reflection on held attitudinal positions. These aspects of AI bots highlight the technology's value to engage in complex cognition processes, such as attitude changes and radicalisation (Blasiak et al. 2021). The possibility to foster lasting resilience to extremist online propaganda is currently not addressed by the predominantly retroactive CVE measures online.

Social bots have attributes that make them particularly suitable for content interventions on SM. These attributes include context and emotional awareness (Winkler and Söllner 2018), the capability of direct communication with human users, perceived higher persuasiveness (Dennis et al. 2020), and perceived value neutrality. The latter mitigates some of the complexity of users' intervention susceptibility and their predisposition to avoid disclosing socially undesirable views to other people (Nadarzynski et al. 2020). In addition, social bots are cost-efficient due to their potential for scalability, thus, having an economic advantage over human intervention (Mittal et al. 2016). While social bot interventions are promising and presumably feasible (Marcellino et al. 2020), fundamental deficits in our understanding of social bots and online interventions prevent their safe use in CVE.

3 Research Design – A Cognitive Perspective on Attitude Change

Based on Moravec et al. (2020) suggestions about how to apply dual-process theory to the design of SM interventions, our research adopts dual-process theories to understand the effect of social bot interventions on System 1 versus System 2 cognition and attitude change. Dual-process theories

propose that people process messages through two fundamentally different cognitive processes (Kahneman 2011). A primary low effort system 1 that is experiential, associative, and unconscious, and a separate conscious system 2 (System 2), which is capable of rational thoughts based on high effort cognition and increased level of message elaboration (Petty and Cacioppo 1986).

System 1 and System 2 processing are complementary cognitive systems. Low effort System 1 processing is the default system that occurs first, continuously, and unconsciously. System 1 primarily responds to visual or peripheral cues (Petty and Cacioppo 1986) and is judgmental, mainly relying on existing attitudes and beliefs (Moravec et al. 2020). Extremists often exploit this inelaborate, judgmental and emotional response to induce radicalisation (Braddock 2019; Saleh et al. 2021). In contrast, System 1 may trigger consequent System 2 cognition if sufficiently motivated and cognitively able to spend the required elaboration effort (Petty and Cacioppo 1986). The more effortful and deliberate elaborations of System 2 (Kahneman 2011) is typically a response to cues picked up by the instinctive System 1. For instance, fear is a strong motivator (Boss et al. 2015) to assess information more deliberately.

Messages that contain threats and appeal to fear can induce reactance in users. Reactance is the tendency to respond with anger and counterarguing when exposed to threatening messages (Dillard and Shen 2005). A message that highlights the extremist threat and consequences to the individual user may cause anger and counterarguing, triggering user’s reactance and ultimately leads to more resilience to extremist persuasion (Braddock 2019). This extremist threat resilience and subsequent attitude change is likely the result of System 2 cognition (Kahneman 2011).

We consider both, dual-process cognition, and reactance to fear-inducing intervention messages. We presuppose that (H1) there is a relationship between System 1 cues that suggest a threat to users and subsequent reactance (i.e., feelings of anger and counterarguing). Reactance theory suggests that threat-induced fear leads to expressions of anger and counterarguing in users, ultimately leading to a shift in attitude and behaviour (Dillard and Shen 2005). Attitudes form as a result of System 2 (Petty and Cacioppo 1986). Hence, we propose a relationship between System 2 (fear) appeal messages of social bot interventions and subsequent attitude change as a response to fear appeal messages (H2).

We propose a research model (Figure 1) that captures dual-process cognitions in System 1 and System 2 as a response to social bot interventions. We observe the attitude-changing qualities of social bot interventions as a response to exposure to extremist content on SM. The social bot consequently interferes with the dual-process outcome using peripheral cues (i.e., source credibility, profile information, social proof) and message attributes that use fear appeal to motivate message engagement, as well as supplementary information that informs a user’s ability to recognize the threat of extremist online messages. Hence, we observe potential changes in attitude and user reactance.

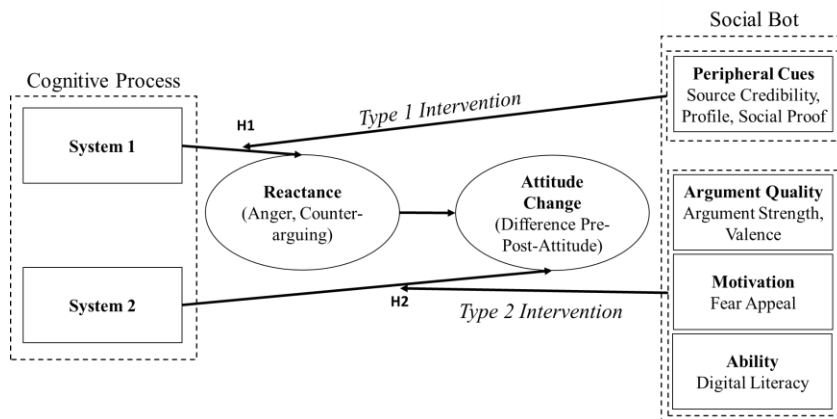


Figure 1 Research model to study the social bot interventions to extremist messaging

An attitude is a complex mental state that encompasses beliefs, feelings, values, and dispositions to act in specific ways and encompasses users' evaluations of 'objects' (Chaiken et al. 1995, p. 387). These objects may be anti-immigration or traditional Western values in the case of Australian extremism. A change in attitude constitutes a shift in a user’s evaluation of an attitude object featured in the content on SM. In practice, this could be the level of agreement or disagreement to statements relating to anti-immigration or traditional Western values.

In a dual process, attitude change is subject to how information messages are received and cognitively processed. Low elaboration of the factual elements of a message is the rudimentary state of message processing (Kahneman 2011), moderated by peripheral cues associated with, for instance: source

credibility of the bot and visual cues: bot-name, avatar, and bot details (i.e., social proof) (Petty and Cacioppo 1986). Most users on SM are in a hedonic mindset. Consequently, users are reluctant to more deliberate and effortful processing (Moravec et al. 2020) because it requires more cognitive load and increased elaboration activity. For effortful processing in system 2 to occur, a potential intervention needs to motivate the user to undertake the more effortful evaluation (Petty and Cacioppo 1986). Furthermore, cognitive biases, including conformity and confirmation biases, influence attitude-changing qualities of a message (Kahneman 2011) that evokes a system 2 response.

4 Research Method

To test how specific social bot attributes can influence attitude change and trigger user reactance to combat extremist messages, we propose a qualitative pilot study to inform a between-subjects 2 x 3 factorial experiment. Different social bot peripheral cues and message attributes are included.

In a pilot study, we initially test the efficacy of social bot interventions by assessing participant responses to a structured interview. We will also interview users of online groups that are associated with far-right sentiment and attitude objects (i.e., anti-immigration, discrimination of minorities) to ensure attitude variation. We ask participants about their trust in government intervention versus platform provider intervention using social bots and test the design of our subsequent online experiment intervention design.

In the main experiment, we present participants with public SM posts containing conservative views on far-right attitude objects. The comment section of the posts will be visible to the participants. Within the comment section, comments by other users and intervention messages identifiable as social bots through their account attributes: account name, avatar, and profile page will be visible and constitute our treatments. The messages commented by social bots aim to increase user awareness of the threats of extremist content. Our interventions aim to trigger System 1 and System 2 processing, observing potential differences in attitude change between respondents.

A Likert scale survey is used to determine user attitudes pre-and post- interventions. Similarly, we will assess anger and counterarguing to observe reactance (Dillard and Shen 2005).

5 Discussion and Conclusion

Our research builds on previous work in the field of AI. Our contribution is to introduce social bots as dual-process interventions to the CVE context while considering the efficacy of specific social bot attributes in intervening with extremist propaganda and user attitudes. We follow recent calls to action (Moravec et al. 2020) and contribute to SM intervention research by applying System 1 and System 2 processing, identifying potential mechanisms to trigger System 1 or System 2 cognition in users deliberately. Theoretical insights into how intervention mechanisms and cues can be used to deliberately trigger attitude change and reactance in users are invaluable to the design of interventions. We call for interventions that go beyond the technical removal of dangerous content and rather foster more cautious and deliberative user engagement with content online. Thus, it is crucial to understand the specific effects of intervention cues on a user's cognition on SM. These findings' implications go beyond the extremism context and, for instance, may apply to IS security research and the design of interventions to phishing messages.

Government agencies and tech companies worldwide are highly interested in interventions against online radicalisation and the prevention of extremism, even more so considering recent events. In our empirical study, we propose social bots as a scalable, fast, and economical alternative to current radicalisation intervention. We hope that the results of our research contribute to safer online environments and stimulate more explorative research that addresses current challenges in the design of our SM landscape. We anticipate that future research will consider the ethics of live implementation of interventive bots on SM and potential inter-user effects of social bot exposure.

6 References

- Australian Security Intelligence Organisation. 2021. "Director-General's Annual Threat Assessment." (<https://www.asio.gov.au/publications/speeches-and-statements/director-generals-annual-threat-assessment-2021.html>, accessed May 1, 2021)
- Baugut, P., and Neumann, K. 2020. "Online Propaganda Use During Islamist Radicalization," *Information, Communication & Society* (23:11), pp. 1570-1592.

- Blasiak, K. M., Risius, M., and Matoon, S. 2021. "“Social Bots for Peace”: A Dual-Process Perspective to Counter Online Extremist Messaging," in: *International Conference on Information Systems (ICIS)*. Austin, Texas.
- Boss, S. R., Galletta, D. F., Lowry, P. B., Moody, G. D., and Polak, P. 2015. "What Do Systems Users Have to Fear? Using Fear Appeals to Engender Threats and Fear That Motivate Protective Security Behaviors," *MIS Q.* (39:4), pp. 837–864.
- Braddock, K. 2019. "Vaccinating against Hate: Using Attitudinal Inoculation to Confer Resistance to Persuasion by Extremist Propaganda," *Terrorism and Political Violence*, pp. 1-23.
- Braddock, K. 2020. "Clearing the Smoke and Breaking the Mirrors: Using Attitudinal Inoculation to Challenge Online Disinformation by Extremists," in *Radicalization and Counter-Radicalization*, D.M.D. Silva and M. Deflem (eds.). Emerald Publishing Limited, pp. 155-166.
- Chaiken, S., Pomerantz, E. M., and Giner-Sorolla, R. 1995. "Structural Consistency and Attitude Strength," in *Attitude Strength: Antecedents and Consequences*. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc, pp. 387-412.
- Dennis, A. R., Kim, A., Rahimi, M., and Ayabakan, S. 2020. "User Reactions to Covid-19 Screening Chatbots from Reputable Providers," *Journal of the American Medical Informatics Association* (27:11), pp. 1727-1731.
- Dillard, J. P., and Shen, L. 2005. "On the Nature of Reactance and Its Role in Persuasive Health Communication," *Communication Monographs* (72:2), pp. 144-168.
- Dingler, T., Choudhury, A., and Kostakos, V. 2018. "Biased Bots: Conversational Agents to Overcome Polarization," *ACM International Joint Conference: ACM*, pp. 1664-1668.
- Douek, E. 2020. "The Rise of Content Cartels." (<https://knightcolumbia.org/content/the-rise-of-content-cartels>, accessed March 20, 2021)
- Frischlich, L. 2020. "#Dark Inspiration: Eudaimonic Entertainment in Extremist Instagram Posts," *New Media & Society* (23:3), pp. 554-577.
- Glazzard, A. 2017. "Losing the Plot: Narrative, Counter-Narrative and Violent Extremism," (<https://icct.nl/publication/losing-the-plot-narrative-counter-narrative-and-violent-extremism>, accessed March 20, 2021)
- Human Rights Watch. 2020. "Tech Firms' Counterterrorism Forum Threatens Rights." (<https://www.hrw.org/news/2020/07/30/tech-firms-counterterrorism-forum-threatens-rights#>, accessed April 2, 2021)
- Jasko, K., LaFree, G., and Kruglanski, A. 2017. "Quest for Significance and Violent Extremism: The Case of Domestic Radicalization," *Political Psychology* (38:5), pp. 815-831.
- Jee, C. 2020. "Facebook Needs 30,000 of Its Own Content Moderators, Says a New Report." *MIT Technology Review* (<https://www.technologyreview.com/2020/06/08/1002894/facebook-needs-30000-of-its-own-content-moderators-says-a-new-report/>, accessed April 2, 2021)
- Kahneman, D. 2011. *Thinking, Fast and Slow*. New York: Macmillan.
- Kruglanski, A., Bélanger, J., and Gunaratna, R. 2019. *The Three Pillars of Radicalization: Needs, Narratives, and Networks*. New York: Oxford University Press.
- Marcellino, W., Magnuson, M., Stickells, A., Bourdreaux, B., Helmus, T. C., Geist, E., and Winkelman, Z. 2020. "Counter-Radicalization Bot Research. Using Social Bots to Fight Violent Extremism," *RAND*, Santa Monica, Calif., pp. 1-155.
- Maruping, L., and Matoon, S. 2020. "The Evolution of Software Development Orchestration: Current State and an Agenda for Future Research," *European Journal of Information Systems* (29:5), pp. 443-457.
- Miron, A. M., and Brehm, J. W. 2006. "Reactance Theory - 40 Years Later," *Zeitschrift für Sozialpsychologie* (37:1), pp. 9-18.
- Mittal, A., Agrawal, A., Chouksey, A., Shriwas, R., and Agrawal, S. 2016. "A Comparative Study of Chatbots and Humans," *International Journal of Advanced Research in Computer and Communication Engineering* (5:3), pp. 1055 - 1057.

- Moravec, P. L., Kim, A., and Dennis, A. R. 2020. "Appealing to Sense and Sensibility: System 1 and System 2 Interventions for Fake News on Social Media," *Information Systems Research* (31:3), pp. 987-1006.
- Nadarzynski, T., Bayley, J., Llewellyn, C., Kidsley, S., and Graham, C. A. 2020. "Acceptability of Artificial Intelligence (AI)-Enabled Chatbots, Video Consultations and Live Webchats as Online Platforms for Sexual Health Advice," *BMJ Sexual & Reproductive Health* (46:3), pp. 210-217.
- Nuraniyah, N. 2018. "Not Just Brainwashed: Understanding the Radicalization of Indonesian Female Supporters of the Islamic State," *Terrorism and Political Violence* (30:6), pp. 890-910.
- Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., and Rand, D. G. 2020. "Fighting Covid-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention," *Psychological Science* (31:7), pp. 770-780.
- Petty, R., and Cacioppo, J. 1986. "The Elaboration Likelihood Model of Persuasion," *Advances in Experimental Social Psychology* (19), pp. 123-205.
- Reed, A., and Ingram, H. 2018. "A Practical Guide to the First Rule of Ct-Cve Messaging. Do Violent Extremists No Favours," (<https://www.europol.europa.eu/publications-documents/practical-guide-to-first-rule-of-ctcve>, accessed May 05, 2021)
- Reed, A., Ingram, H. J. I., and Whittaker, J. 2017. "Countering Terrorist Narratives," *EU Policy Department for Citizen's Rights and Constitutional Affairs*, pp. 1-57.
- Saleh, N. F., Roozenbeek, J., Makki, F. A., McClanahan, W. P., and Van Der Linden, S. 2021. "Active Inoculation Boosts Attitudinal Resistance against Extremist Persuasion Techniques: A Novel Approach Towards the Prevention of Violent Extremism," *Behavioural Public Policy*, pp. 1-24.
- Schmitt, J. B., Rieger, D., Rutkowski, O., and Ernst, J. 2018. "Counter-Messages as Prevention or Promotion of Extremism?! The Potential Role of Youtube," *Journal of Communication* (68:4), pp. 780-808.
- Schmuck, D., Heiss, R., and Matthes, J. 2020. "Drifting Further Apart? How Exposure to Media Portrayals of Muslims Affects Attitude Polarization," *Political Psychology* (41:6), pp. 1-18.
- Seering, J., Flores, J. P., Savage, S., and Hammer, J. 2018. "The Social Roles of Bots," *Proceedings of the ACM on Human-Computer Interaction* (2:CSCW), pp. 1-29.
- Seering, J., Luria, M., Kaufman, G., and Hammer, J. 2019. "Beyond Dyadic Interactions," *CHI Conference on Human Factors in Computing Systems: ACM*, pp. 1-13.
- The Economist. 2020. "Social-Media Platforms Are Destroying Evidence of War Crimes." (<https://www.economist.com/international/2020/09/21/social-media-platforms-are-destroying-evidence-of-war-crimes>, accessed May 05, 2021)
- Thomas, K., Akhawe, D., Bailey, M., Boneh, D., Bursztein, E., Consolvo, S., Dell, N., Durumeric, Z., Kelley, P. G., and Kumar, D. 2021. "Sok: Hate, Harassment, and the Changing Landscape of Online Abuse." (<https://research.google/pubs/pub49786>, accessed May 05, 2021)
- Vraga, E. K., Bode, L., and Tully, M. 2021. "The Effects of a News Literacy Video and Real-Time Corrections to Video Misinformation Related to Sunscreen and Skin Cancer," *Health Communication*, pp. 1-9.
- Wilson, T. D., Lindsey, S., and Schooler, T. Y. 2000. "A Model of Dual Attitudes," *Psychological Review* (107:1), pp. 101-126.
- Winkler, R., and Söllner, M. 2018. "Unleashing the Potential of Chatbots in Education: A State-of-the-Art Analysis," *Academy of Management Proceedings* (2018), pp. 1-40.
- Winter, C., Neumann, P., Meleagrou-Hitchens, A., Ranstorp, M., Vidino, L., and Fürst, J. 2020. "Online Extremism: Research Trends in Internet Activism, Radicalization, and Counter-Strategies," *International Journal of Conflict and Violence (IJCV)* (14:2), pp. 1-20.
- Zarouali, B., Makhortykh, M., Bastian, M., and Araujo, T. 2020. "Overcoming Polarization with Chatbot News? Investigating the Impact of News Content Containing Opposing Views on Agreement and Credibility," *European Journal of Communication* (36:1), pp. 53 -68.

Acknowledgements

Financial support for the project from The University of Queensland School of Business in the Research Start-up Support Funding is gratefully acknowledged.

Copyright

Copyright © 2021 authors. This is an open-access article licensed under a [Creative Commons Attribution-NonCommercial 3.0 Australia License](https://creativecommons.org/licenses/by-nc/3.0/au/), which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and ACIS are credited.