



Selected Papers of #AoIR2020:
The 22nd Annual Conference of the
Association of Internet Researchers
Virtual Event / 13-16 Oct 2021

SOCIAL MEDIA USE, TRUST AND TECHNOLOGY ACCEPTANCE: INVESTIGATING THE EFFECTIVENESS OF A CO-CREATED BROWSER PLUGIN IN MITIGATING THE SPREAD OF MISINFORMATION ON SOCIAL MEDIA

Eleni A. Kyza

Department of Communication and Internet Studies, Cyprus University of Technology

Christiana Varda

Department of Communication and Internet Studies, Cyprus University of Technology

Loukas Konstantinou

Department of Communication and Internet Studies, Cyprus University of Technology

Evangelos Karapanos

Department of Communication and Internet Studies, Cyprus University of Technology

Serena Coppolino Perfumi

EGov Lab, Stockholm University

Mattias Svahn

EGov Lab, Stockholm University

Yiannis Georgiou

Department of Communication and Internet Studies, Cyprus University of Technology

Introduction

The viral spread of misinformation on social media, be it by algorithms or humans, is a threat to democracy and individual autonomy (Bastick, 2021), and can heighten the risk for physical and psychological harm. As traditional gatekeeping is replaced online by new models of information control, this work investigated how some form of power over the control of information can be renegotiated back to the everyday user. In response to this year's conference theme, we present a multi-disciplinary effort to build capacity to support laypersons (citizens) in identifying and critically reviewing misinformation on

Kyza, E. A., Varda, C., Konstantinou, L., Karapanos, E., Perfumi, S. C., Svahn, M. & Georgiou, Y. (2021, October). *Social media use, trust, and technology acceptance: Investigating the effectiveness of a co-created browser plugin in mitigating the spread of misinformation on social media*. Paper presented at AoIR 2021: The 22nd Annual Conference of the Association of Internet Researchers. Virtual Event: AoIR. Retrieved from <http://spir.aoir.org>.

social media. This effort adopted a sociotechnical approach to online engagement with (mis)information on social media platforms, using Twitter as a context. Using participatory methodologies, we co-created a technological solution (a web browser plugin) that can support citizens in assuming an agentic relationship with technology and supporting them in resisting the spread of misinformation. We report on an empirical study that investigated the effectiveness of the strategies that were adopted as part of this work.

Power on social media

The technologically mediated landscape of social media presents opportunities for increased engagement in democratic discussions (Loader & Mercea, 2011); researchers also argue that it may contribute toward misleading the user through technological and other means (Kozyreva et al., 2020). We take a human-centered approach to the role of technology in combating misinformation; using a social constructivist lens we approach technology as something that can be shaped by humans and which can support human values and interests. While experts acknowledge the power struggles in relation to misinformation online, a Pew Research Center study with 1116 experts and stakeholders (Anderson & Rainie, 2017) suggested that there is no clear consensus on whether society will be able to address misinformation on social media, or what the optimal solution might be. Our study explores this vast, multi-disciplinary and complex topic, and asks whether power over the control of information can be redistributed to human actors, supported by co-created artificial intelligence-driven misinformation tools.

The role of technology in combating misinformation

We view technology as value-laden, but also as a force that can be shaped by humans. We view humans as having reflective agency which, according to John Dewey, can lead to democratic engagement (Whipple, 2005). Such agency is mitigated by individual, situational and technological factors (Tandoc, 2019), which is the focus of this study. In this context we organized nine co-creation workshops (three at each site) in Austria, Greece and Sweden to explore citizens' challenges in identifying and addressing misinformative content on social media and to guide our agile technology development efforts. The end result was a browser plugin prototype, designed to work with Twitter, which combined artificial intelligence models and human input. The plugin provides support in addressing misinformation by examining tweets through AI and a rule engine to determine credibility, presents the credibility metrics and explanations to the user, and employs a nudging mechanism ('blurry'), which can be turned off by the user, to deter engagement with non-credible posts.

Methods

We conducted a quasi-experimental study across four European sites to understand whether the use of a co-created technological intervention on Twitter could lead to more misinformation-resilient behavior, as indicated by participants' intention to avoid liking or sharing misinformative posts. We adopted a mixed-methods research design. Participants (n=80) were assigned to an experimental (n=40) or control condition (n=40). Each group viewed a curated Twitter timeline with or without the browser plugin. The timeline included both accurate and misinformative science-related posts, derived from a dataset of 4,550 fact-checked social media posts.

Participants first completed questionnaires on individual characteristics (digital literacy, media literacy, epistemic beliefs, social media use, trust in science, and digital citizenship) and were then presented with the Twitter feed, which featured credible and non-credible posts. Following a semi-structured data collection protocol, they were asked to think aloud as they decided which actions they would take for each post. At the end of each intervention, we conducted semi-structured interviews to obtain the participants' perspective of the process of evaluating post credibility. The participants' actions with the Twitter timeline and their externalized reflections were videotaped and transcribed verbatim. Participants in the experimental condition also completed a post-intervention questionnaire relating to trust in the technology and technology acceptance. To answer our research question we used multinomial regressions, bivariate analysis and Fisher's exact test in SPSS, followed by qualitative analyses using NVivo 12 to provide further insights.

Findings

The aim of this study was to investigate whether a co-created software solution could support citizens' resilience to misinformation on social media, and understand the role of individual, situational and technological characteristics in the citizens' engagement with misinformation.

Individual characteristics: A series of multinomial regressions indicated that all but one of the individual characteristics assessed were not statistically significant predictors of user's trust/distrust profile toward the plugin: only social media use was a statistically significant predictor of users' trust and distrust towards the plugin. A multinomial logistic regression analysis indicated that social media use predicted the participants' profile (trust/distrust). Participants who reported using social media less, also indicated less trust and less acceptance of the technological intervention ($\beta=-1.061$, $OR=0.346$, $p=0.041$).

Situational characteristics: A bivariate analysis indicated that trust in technology and technology acceptance were strongly positively correlated ($r=0.85$, $p<0.01$). We found a significant relationship between the participants' profile (trust / distrust) and "sharing" misinformation; 'trust' profile participants were less likely to share misinformation ($p<0.05$). No relationship was found between trust / distrust profile and "liking".

Technological characteristics: A Fisher's exact test indicated a significant relationship between the condition (plugin/ no_plugin) and endorsing ("liking") the misinformative posts. The participants in the control condition (no_plugin) were more likely to "like" a misinformative post than the participants in the experimental condition ($p < 0.05$).

Discussion and Implications

The findings of this study confirm that individual, situational and technological characteristics impact a person's engagement with misinformative posts. Results suggest that the presence of a co-created technological solution can deter social media users from endorsing misinformation in an authentic social media environment. Resistance to sharing misinformation was correlated with participants' trust and acceptance of the technological intervention, which was related to their use of social

media. These findings have broader implications about the conditions under which technological interventions can support the fight against online misinformation. Additional studies are needed to understand all the implications of this study, with larger samples and different topical contexts.

Acknowledgement

This study was funded by the European Commission, grant agreement 770302, project Co-Inform (www.coinform.eu).

References

- Anderson, J. & Rainie, L. (2017). The future of truth and misinformation online. *Pew Research Center*, 19.
- Kozyreva, A., Lewandowsky, S., & Hertwig, R. (2020). Citizens versus the internet: Confronting digital challenges with cognitive tools. *Psychological Science in the Public Interest*, 21(3), 103-156.
- Loader, B. D. & Mercea, D. (2011). Networking Democracy? *Information, Communication & Society*, 14:6, 757-769, DOI: 10.1080/1369118X.2011.592648
- Tandoc Jr, E. C. (2019). The facts of fake news: A research review. *Sociology Compass*, 13(9), e12724.
- Whipple, M. (2005). The Dewey-Lippmann Debate today: Communication distortions, reflective agency, and participatory democracy. *Sociological Theory*, 23(2), 156-178.