



# Hijacking Epistemic Agency

## How Emerging Technologies Threaten our Wellbeing as Knowers

John Dorsch  
University of Edinburgh  
john.dorsch@ed.ac.uk

### ABSTRACT

The aim of this project to expose the reasons behind the pandemic of misinformation (henceforth, PofM) by examining the enabling conditions of epistemic agency and the emerging technologies that threaten it. I plan to research the emotional origin of epistemic agency, i.e. on the origin of our capacity to acquire justification for belief, as well as on the significance this emotional origin has for our lives as epistemic agents in our so-called *Misinformation Age* [1]. This project has three objectives. First, I plan to expose the degree to which epistemic agency is made possible by an under-researched species of emotion called *epistemic feelings* [2] [3]. Perhaps, the most epistemically significant is the feeling of confidence [4]. In particular, epistemic feelings make epistemic agency possible by making errors in reasoning salient or the potential lack thereof [5].

Second, in order to diagnose the reasons for PofM, I will analyze the emotional basis of epistemic agency in the context of emerging technologies [6]. Epistemic feelings ought to be construed as motivators of epistemic acts, specifically acts following exposure to misinformation spread by social media [7]. For example, a recent study found that subjects on YouTube have a 6.3% probability in five clicks to go from watching innocuous videos to misogynistic videos [8]. Accordingly, one prominent view holds that social media algorithms lead subjects down the proverbial rabbit hole, producing content that elicits strong emotional reactions [9] [10].

But under-researched is the degree to which these emotions are intertwined with epistemic feelings, so that subjects are likely to misjudge the information as correct due to being primed to do so by processing the AI's suggested media. Thus, when this technology is used in epistemic practices, the result is invariably epistemically bad behavior because of how this manner of producing content motivates epistemic acts through the production of erroneous feelings of confidence, which result from the perceived ease of cognitive processing [11] [12]: the content is easy to process and so, the subject confidently misjudges, it is true. As of yet, an account of this cognitive and

emotional decision-making process has not been used to analyze how emerging technologies are threatening epistemic agency, and thus our wellbeing as knowers.

Finally, I plan to develop therapeutics for PofM by outlining how to reform our collective epistemic practices. Thus, this project complements existing research in virtue epistemology [13], specifically by articulating the challenges to cultivating epistemic virtues and warding off the vices of the mind [14]. In summation, the project aims to account for how emerging technologies are committing a form of epistemic injustice [15] through the exploitation of cognitive biases and the production of erroneous epistemic feelings, with the overarching goal being a framework for convalescing from the pandemic of misinformation plaguing our age.

### REFERENCES

- [1] Cailin O'Connor and James Weatherall. 2019. *The Misinformation Age*. Yale University Press, Cambridge, MA.
- [2] Joëlle Proust. 2013. *The Philosophy of Metacognition: Mental Agency and Self-Awareness*. Oxford University Press, Oxford.
- [3] John Dunlosky and Janet Metcalfe. 2008. *Metacognition*. Sage, London.
- [4] Florent Meyniel, Mariano Sigman, and Zachary F. Mainen. 2015. Confidence as Bayesian Probability: From Neural Origins to Behavior. *Neuron* 88, 1, 78-92. DOI: <https://doi.org/10.1016/j.neuron.2015.09.039>
- [5] Harry G. Frankfurt. 1978. The problem of action. *American philosophical quarterly* 15, 2, 157-162.
- [6] Markus D. Dubber, Frank Pasquale, and Sunit Das. (Eds.). 2020. *The Oxford Handbook of Ethics of AI*. Oxford University Press, Oxford.
- [7] Chris Fuchs. 2021. *Social media: A critical introduction* (3rd ed.). Sage, London.
- [8] Kostantinos Papadamou, Savvas Zannettou, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, and Michael Sirivianos. 2021. "How over is it?" Understanding the Incel Community on YouTube. In *Proceedings of the ACM on Human-Computer Interaction* 5(CSCW2), 1-25. DOI: <https://doi.org/10.1145/3479556>
- [9] Kevin Roose. 2019. The Making of a YouTube Radical. *The New York Times*. Retrieved Dec 19, 2021 from <https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html>
- [10] Manoel Horta Ribeiro, Raphael Ottoni, Robert West, Virgilio A. F. Almeida, and Wagner Meira. 2020. Auditing radicalization pathways on YouTube. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, 131-141. DOI: <https://doi.org/10.48550/arXiv.1908.08313>
- [11] Adam L. Alter and Daniel M. Oppenheimer. 2009. Uniting the Tribes of Fluency to Form a Metacognitive Nation. *Personality and Social Psychology Review* 13, 3, 219-235. DOI: <https://doi.org/10.1177/1088868309341564>
- [12] Norbet Schwarz, Madeline Jalbert, Tom Noah, Lynn Zhang. 2021. Metacognitive experiences as information: Processing fluency in consumer judgment and decision making. *Consumer Psychology Review* 4, 1, 4-25. DOI: <https://doi.org/10.1002/arc.1067>
- [13] Duncan Pritchard. 2005. Virtue epistemology and the acquisition of knowledge. *Philosophical Explorations* 8, 3, 229-243. DOI: <https://doi.org/10.1080/13869790500219513>
- [14] Quassim Cassam. 2018. *Vices of the Mind: From the Intellectual to the Political*. Oxford University Press, Oxford.
- [15] Miranda Fricker. 2007. *Epistemic injustice: Power and the Ethics of Knowing*. Oxford University Press, Oxford.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

AIES'22, August 1–3, 2022, Oxford, United Kingdom

© 2022 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-9247-1/22/08.

DOI: <https://doi.org/10.1145/3514094.3539537>