The Benefits of Bias? A Rational Choice Informed Agent-Based Model on Algorithmic Technology in the Public Sector

Keywords: Agent-Based Modeling, Artificial Intelligence, Machine Learning, Predictive Policing, Algorithmic Bias

Extended Abstract

"Our bill [the proposed *Algorithmic Accountability Act*] requires companies to study the algorithms they use, identify bias in these systems, and fix any discrimination or bias they find." – U.S. Senator Ron Wyden

Both researchers and media pundits assessing the dangers of (the implementation of) algorithmic technology, such as artificial intelligence or machine learning, in the public sector often point to a singular culprit for any and all negative unintended consequences, bias, specifically algorithmic bias. Algorithmic bias is said to occur when algorithmic systems produce consistent systematic errors without a warranted rationale. Human beings also have a litany of cognitive biases that impact how they make decisions. Extant scholarship has not yet examined how receiving decision-making advice from a (theoretically) unbiased algorithmic technology systems can impact agent behavior in the context of public sector enforcement activities, such as predictive policing or fraud detection and prevention.

We address this gap by first building a rational choice model to understand the micro-level dynamics of human agents interacting with an unbiased advisor in the form of algorithmic technology and a biased advisor in the form of human political leadership/oversight. We find that unbiased advice from algorithmic technology cannot possibly influence an enforcement officer to modify their decision. We also find that human political leadership, acting with a blame avoidance bias, can modify the behavior of enforcement agents to be more cautious. This is a significant finding, as it demonstrates that unbiased algorithmic technology may not be the panacea for reducing negative unintended consequences, as many have purported.

We then tested these findings in a macro-level environment using agent-based modeling. We find that unbiased algorithmic technology can be a condition that leads to increasingly aggressive enforcement agent behavior. The findings of this paper point towards a need for further empirical research, as well as a more nuanced understanding of how the interplay between algorithmic bias and human cognitive biases impacts enforcement agent behavior and potentially contributes to negative unintended consequences.

References

Acciarini, C., Brunetta, F., & Boccardelli, P. (2020). Cognitive biases and decision-making strategies in times of change: A systematic literature review. *Management Decision*, *59*(3), 638–652. https://doi.org/10.1108/MD-07-2019-1006

Turner Lee, N., Resnick, P., Barton, G. (2019, May 22). Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms. *Brookings*. https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/

Weaver, R. K. (1986). The Politics of Blame Avoidance. *Journal of Public Policy*, *6*(4), 371–398. https://doi.org/10.1017/S0143814X00004219

Wyden, Booker, Clarke Introduce Bill Requiring Companies To Target Bias In Corporate Algorithms | U.S. Senator Ron Wyden of Oregon. (2019). Retrieved February 18, 2023, from https://www.wyden.senate.gov/news/press-releases/wyden-booker-clarke-introduce-bill-requiring-companies-to-target-bias-in-corporate-algorithms-