

RESEARCH INSIGHT

Unpacking the Dynamics of Bias, Decision Revision, and Performance in Human-AI Collaboration

Yan Chen¹, Xueping Liang¹, Kane Smith², Arijit Sengupta¹

¹ Department of Information Systems and Business Analytics, College of Business, Florida International University

² Department of Information Technology & Decision Sciences, College of Business, University of North Texas

Human interaction with artificial intelligence (AI) is often influenced by perceptual biases, including cognitive biases related to AI and self-capabilities, as well as motivational biases concerning personal control. Prior research has largely examined such biases using single-stage or outcome-based approaches, overlooking the sequential and interdependent nature of human-AI decision processes and the role of AI explanations in shaping these processes. This study develops a stage-based decision funnel model in Figure 1 to examine how decisions are progressively filtered across stages, providing a theoretical framework for understanding how perceptual biases and AI explanations influence stage-specific decision behavior and performance. Drawing on repeated-measures data from two online experiments in the context of AI-assisted phishing email detection and using generalized estimating equations with nonlinear specifications, we show that cognitive biases exert stage-specific, detrimental, and, in some cases, nonlinear effects. These effects are most pronounced during decision transitions, when users reconsider prior decisions in response to AI aid. In contrast, illusion of control (i.e., a motivational bias) is associated with beneficial decision revisions and improved performance, suggesting that a sense of control may help users benefit from collaboration with AI. We further find that concrete explanations are associated with fewer overall revisions and higher revision accuracy, although they do not significantly improve overall performance.

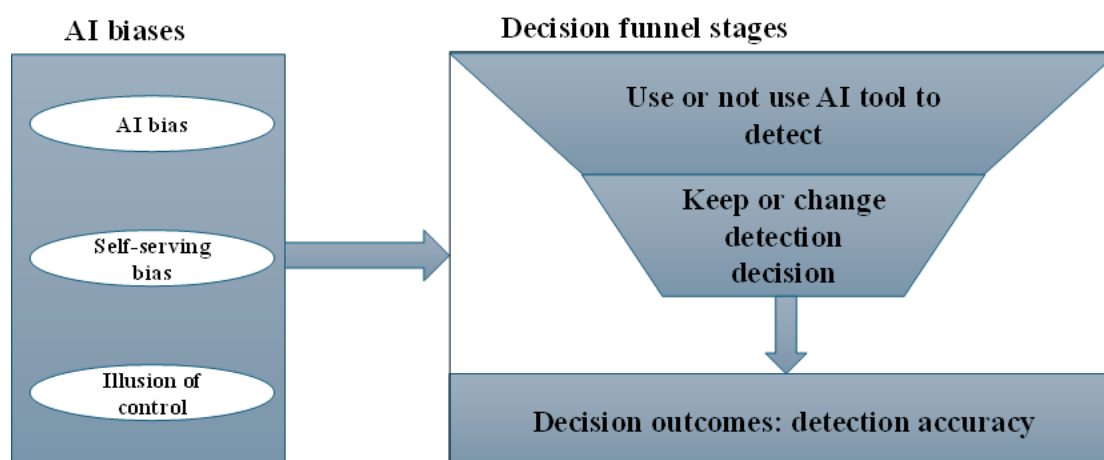


Figure 1. AI decision funnel model