

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Beyond Accurate AI: Counter-Bias Advice Calibrates Human Judgment

Permalink

<https://escholarship.org/uc/item/6kn9k4mq>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Kagawa, Rina

Honda, Hidehito

Sanuki, Masaru

Publication Date

2026

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Beyond Accurate AI: Counter-Bias Advice Calibrates Human Judgment

Rina Kagawa (kagawa-r@aist.go.jp)^{1,3}, Hidehito Honda², Masaru Sanuki³

¹Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology (AIST)

²Department of Artificial Intelligence and Cognitive Science, Ottemon Gakuin University

³Institute of Medicine, University of Tsukuba

Abstract

With the performance improvements and widespread social adoption of artificial intelligence (AI), including large language models, AI-assisted decision-making is rapidly expanding. On the other hand, as human decision-making is increasingly delegated to AI, concerns about degrading the quality of autonomous human decision-making when AI is unavailable has become an urgent social issue. This study focused on the insight from the wisdom of crowds that diverse opinions improve decision-making accuracy. We call advice that judges based on bias that counterbalances the human bias, “Adaptive Counter-Bias Advice.” Through theoretical analysis and behavioral experiments, we confirmed that people who repeatedly made judgments while referring to Adaptive Counter-Bias Advice showed reduced bias and improved accuracy in autonomous decision-making in situations where advice was no longer available compared to those who used accurate advice. These findings offer new insights into using and selecting AI systems to enhance autonomous human decision-making performance.