UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

A Computational Account of Epistemic Vigilance: Learning from Selective Truths through Bayesian Reasoning

Permalink

https://escholarship.org/uc/item/99z8j19z

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Authors

Fang, Ke
Hawkins, Robert
Wu, Charley M
et al.

Publication Date

2025

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

eScholarship.org Powered by the California Digital Library
University of California

A Computational Account of Epistemic Vigilance: Learning from Selective Truths through Bayesian Reasoning

Ke Fang
Stanford University, Stanford, California, United States

Robert Hawkins
Stanford University, Stanford, California, United States

Charley Wu
University of Tübingen, Tübingen, Germany

Michael Franke
University of Tübingen, Tübingen, Germany

Abstract

Strategic actors often manipulate others' beliefs not by lying outright, but through selective truth-telling—also known as lying by omission or paltering—by withholding crucial details while avoiding falsehoods. For example, a pharmaceutical-funded investigator might truthfully report that some patients improved, while omitting that most did not. To guard against such selective disclosures, listeners must engage in epistemic vigilance: critically evaluating information in light of the speaker's potential agenda. In this work, we develop a Bayesian computational model of this process. We present three key findings: (1) credulous listeners who assume informative intent learn quickly in cooperative settings but are highly susceptible to persuasion; (2) vigilant listeners who account for potential bias more accurately recover the underlying true world states, even from purely persuasive speakers—albeit with slower convergence; and (3) this robustness stems from their ability to discount biased framings by reasoning about alternative utterances the speaker could have chosen otherwise.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* ©2025 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).