

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Representational Similarity and Context Inference as a Shared Computational Account for False Memories in Humans

Permalink

<https://escholarship.org/uc/item/2q39k2b3>

Journal

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

Authors

Gupta, Mohan
Kha-Uong, Caleb
Taylor, Jordan
[et al.](#)

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

Representational Similarity and Context Inference as a Shared Computational Account for False Memories in Humans

Mohan W. Gupta

Princeton University, New Jersey, United States of America

Caleb Kha-Uong

Princeton University, New Jersey, United States of America

Jordan A. Taylor

Princeton University, New Jersey, United States of America

Jonathan D. Cohen

Princeton University, New Jersey, United States of America

Abstract

Human memory is constructive: representations that support generalization can also produce systematic errors. Two robust examples—the Deese–Roediger–McDermott paradigm and the misinformation effect—are typically explained by separate theories, from fuzzy-trace representations to source-monitoring failures. We propose a shared computational account in which both phenomena arise from similarity-weighted retrieval over compressed semantic representations. We leverage the Integrated Semantics with Context Inference model, a simple feedforward network that learns independent semantic relationships between items that can then be modulated by context. The model reproduces hallmark human patterns: (i) increased false recognition of critical lures as a function of semantic similarity and list length, and (ii) recognition judgments shift systematically after misleading post-event information. In the misinformation paradigm, inferred event representations shift toward the distractor, quantitatively capturing effect strength. Together, these results suggest that diverse false memory phenomena can be understood as consequences of similarity-based inference.