

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Fallible feedback: Judgments of "novice" automated and human writing tutors

Permalink

<https://escholarship.org/uc/item/6xd837s9>

Authors

Duran, Nicholas

Thomas, Anuja Mariyam

Publication Date

2023

Peer reviewed

Fallible feedback: Judgments of "novice" automated and human writing tutors

Anuja Mariyam Thomas

Arizona State University, Glendale, Arizona, United States

Nicholas Duran

Arizona State University, Glendale, Arizona, United States

Abstract

Automated Writing Evaluation (AWE) tools are steadily gaining prominence in educational settings given the ease and scale in which they can be deployed. Nevertheless, despite producing a similar performance as human raters, the accuracy of automated systems is often met with skepticism. In the current study we explore whether such skepticism extends to writing feedback believed to be generated by human tutors in training or an AI tutor under development. When both sources are fallible, are critical judgments mitigated? Participants (N=477) judged the accuracy of feedback on writing samples given by human or AI tutors where the feedback was normed to be accurate, inaccurate, and ambiguous. Results showed participants were still more likely to deem identical feedback provided by an AI, as compared to human tutors, as less accurate, and they were more confident in these evaluations. It appears that "novice" AWE systems are not completely immune to negative bias.