

Labeling in Their Shoes: Improving Text Annotation with Cognitive Empathy Priming

Sung Hyun Kwon, Jessica M. Clark, Il-Horn Hann, Jui Ramaprasad
Robert H. Smith School of Business, University of Maryland, College Park, MD 20740
shkwon12@umd.edu, jmclark@umd.edu, ihann@umd.edu, jramapra@umd.edu

September 6, 2025

Abstract

Labels created by human annotators are essential for developing and assessing machine learning systems, particularly in areas involving subjective judgment. In cases where personal viewpoints influence labeling decisions, crowd-sourced annotators often struggle to achieve agreement and consistently deviate from expert consensus. Detecting sexist material represents one such perspective-sensitive context. We find consistent disagreement both among annotators themselves and between annotators and experts consensus, a pattern that continues even with many annotators. This misalignment significantly impacts the performance of large language models when fine-tuned on these inconsistent annotations. To address this, we introduce cognitive empathy priming (CEP), a scalable psychological intervention that enhances annotators’ ability to consider perspectives beyond their own. CEP markedly enhances annotation accuracy: treated annotators demonstrate 8-20 percentage point improvements in alignment with expert consensus compared to those who did not receive priming. These enhancements directly translate directly to model performance. Large language models fine-tuned on CEP-treated labels show roughly 16% better alignment with expert consensus labels compared to models trained on standard annotations. Additional robustness tests confirm these findings hold even when accounting for potential expert biases. This study provides a practical, cost-effective solution to enhance the quality of AI training datasets, with particular relevance for content filtering and bias detection.

Keywords: Natural Language Processing, Cognitive Empathy, Label Quality, Crowdsourced Labeling

Note: This paper is being presented as a non-archival 1 page abstract at AIES 2025.

Full Paper: Available [here](#)