

Performance-Based Item Development and Validation in Silica: LLMs and Generative Psychometrics for Structural Validity and Item Difficulty

Wednesday 23 July 2025 15:40 (15 minutes)

Recent advances in large language models (LLMs) present opportunities for developing performance-based items in educational and psychological assessment. We introduce P-AI-GENIE (Performance-based Automatic Item Generation and Network-Integrated Evaluation), an extension of AI-GENIE that focuses on generating and validating performance items. The talk will cover how items can be developed and automatically validated in silica, including the estimation of item difficulty via Exploratory Graph Analysis without collecting data in humans. We seek to demonstrate P-AI-GENIE's potential for streamlining performance assessment development while maintaining measurement quality.

Primary author: GOLINO, Hudson (University of Virginia)

Presenter: GOLINO, Hudson (University of Virginia)

Session Classification: Symposium : "Artificial Intelligence and Large Language Models: Item Development and Validation, Educational Interventions, and Emotion Analysis of Videos"