Contribution ID: **162**   Type: **Symposium**

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

Keywords

AI, Emotion Detection

Authors

Hudson Golino

Affiliation

University of Virginia

Symposium title

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

Coordinator

Hudson Golino

Affiliation

University of Virginia

Abstract

This symposium explores innovative applications of artificial intelligence in educational assessment, intervention, and emotion analysis. Lara Russell-Lasalandra will present AI-GENIE, a novel framework for automated item generation and validation using network-based evaluation approaches. Dr. Golino will discuss P-AI-GENIE's capabilities in developing performance-based items and estimating item difficulty in-silica via Exploratory Graph Analysis. Dr. Tomasevic will demonstrate advanced techniques for emotion detection in video content, combining zero-shot image classification with dynamic exploratory graph analysis. Dr. Teles will present findings from a controlled study examining the effectiveness of AI tutoring systems in enhancing student learning outcomes. Together, these presentations showcase cutting-edge developments in AI-driven educational tools and assessment methods while highlighting practical applications for improving learning and measurement in educational settings.

Keywords

AI, LLMS, Interventions, Items, Emotions

Number of communicatios

Communication 1

Leveraging AI Tutors to Enhance Student Learning: A Controlled Educational Intervention Study

Authors

Mariana Teles

Affiliation

University of Virginia

Abstract

This study investigates the effectiveness of AI-tutored learning environments in implementing evidence-based learning techniques among undergraduate students. Drawing from cognitive science principles, particularly those outlined in Willingham's (2023) work, we developed an innovative intervention utilizing AI tutors to simulate personalized learning environments focused on three key areas: effective note-taking, complex text comprehension, and exam preparation strategies.

In this controlled experiment, 40 first-year psychology students were randomly assigned to experimental (n=20) and control (n=20) conditions. The experimental group participated in eight structured sessions with AI tutors over one semester, while the control group maintained standard learning practices. The intervention's effectiveness is being assessed through a mixed-methods approach combining quantitative academic performance metrics with qualitative analysis of student responses.

Our analytical framework employs a novel combination of traditional pre-post comparisons and advanced natural language processing techniques. Specifically, open-ended student responses are being analyzed using zero-shot classification implemented through Facebook's BART model, complemented by sentiment analysis using the transformemotion package in R. This methodological approach allows for both systematic categorization of learning outcomes and nuanced understanding of students' emotional engagement with the AI-tutored environment.

While data collection is ongoing, this study contributes to the growing body of research on AI-enhanced educational interventions and provides a methodologically rigorous framework for evaluating their effectiveness. The findings will have important implications for implementing scalable, evidence-based learning support systems in higher education.

Keywords

Communication 2

Generative Psychometrics via AI-GENIE: Automatic Item Generation and Validation via Network-Integrated Evaluation

Authors

Lara Russell-Lasalandra

Affiliation

University of Virginia

Abstract

The rapid advancement of artificial intelligence (AI), particularly large language models (LLMs), has introduced powerful tools for various research domains, including psychological scale development. This study presents a fully automated method to efficiently generate and select high-quality, non-redundant items for psychological assessments using LLMs and network psychometrics. Our approach called, Automatic Item Generation and Validation via Network-Integrated Evaluation (AI-GENIE), reduces reliance on expert intervention by integrating generative AI with the latest network psychometric techniques. The efficacy of AI-GENIE was evaluated through Monte Carlo simulations using the Mixtral, Gemma 2, Llama 3, GPT 3.5, and GPT 4o models to generate item pools that mimic Big Five personality assessment. The results demonstrated improvement in item selection efficiency, with overall average increases of 9.78-17.80 in normalized mutual information in the final item pool across all models. After, each model in AI-GENIE generated a Big Five inventory that was administered to independent, representative samples (N = 1000 each) in the U.S. The empirical results show that the items produced across all models were diverse, theoretically consistent, and structurally stable. Taken together, these findings demonstrate that AI-GENIE is a highly effective tool to automate and streamline scale development and validation processes.

Keywords

LLMS, Item Development, Structural Validity

Communication 3

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

Abstract

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.

Keywords

LLM, Item Difficulty, Structural Validity

Communication 4

Decoding Emotion Dynamics in Videos using Dynamic Exploratory Graph Analysis and Zero-Shot Image Classification.

Authors

Aleksandar Tomašević

Affiliation

Novi Sad University

Abstract

We propose a novel approach for modeling and understanding the dynamics of emotion facial expression recognition (FER) scores. Recent advancements in deep learning and transformer-based neural network architectures enable the time series analysis of FER scores extracted from images and videos. This type of data

can be important for psychological research of affective dynamics and emotion expression dynamics. However, the properties of such data are not well understood in the current literature. We propose a new method to simulate FER scores based on a modified version of the Damped Linear Oscillator with a measurement model (DLO-MM). We use this model to conduct a large-scale simulation and use dynamic Exploratory Graph Analysis to investigate the dimensionality of the data and use network scores to recover the values of the latent dimensions—positive and negative sentiment of the expressed emotions. Our results show that the DLO-MM model can be used to simulate FER scores for different patterns of emotion dynamics and that DynEGA can be used to uncover the latent structure of emotion dynamics expressed through FER scores. All methods presented in the paper are implemented in the transforEmotion R package and the tutorial section provides a step-by-step guide on how to simulate FER scores using DLO-MM and how to estimate FER scores from YouTube videos using transformer-based machine learning models.

Primary authors: Dr TOMAŠEVIĆ, Aleksandar (Novi Sad University); GOLINO, Hudson (University of Virginia); Ms RUSSELL-LASALANDRA, Lara (University of Virginia); Dr TELES, Mariana (University of Virginia)

Presenters: Dr TOMAŠEVIĆ, Aleksandar (Novi Sad University); GOLINO, Hudson (University of Virginia); Ms RUSSELL-LASALANDRA, Lara (University of Virginia); Dr TELES, Mariana (University of Virginia)

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

Track Classification: Design/Research methods: Design/Research methods