USE OF ARTIFICIAL INTELIGENCE IN CONTENT VALIDITY

Wednesday 23 July 2025 17:15 (15 minutes)

Poster

USE OF ARTIFICIAL INTELIGENCE IN CONTENT VALIDITY

Author

Xavier G. Ordóñez & Sonia J. Romero

Affiliation

Universidad Complutense de Madrid & Universidad Nacional de Educación a Distancia

Abstract

Artificial intelligence (AI) has significantly evolved across various fields, including psychometrics, where it can enhance the development and validation of measurement instruments. As AI becomes more relevant, literacy in this technology has become essential, driving the creation of educational programs and assessment tools. However, there are still limitations in how AI literacy is measured, prompting the search for new methodologies.

Different authors have proposed models to assess AI literacy. Ng et al. (2021) identified four dimensions: cognitive, metacognitive, affective, and social. Kong et al. (2024) also suggested an approach based on dimensions, including conceptual understanding, real-world application, self-efficacy, and social and ethical awareness. Other studies have developed scales to evaluate AI literacy, such as the AI Literacy Scale by Laupichler et al. (2023) or the MAILS scale by Carolus et al. (2023), which incorporate psychological and educational aspects.

This study proposes an innovative methodology where generative AI is not only used to create items but also to validate their content, replacing the traditional process based on human judges. A total of 720 items were generated using ChatGPT, based on Bloom's Taxonomy and specific AI literacy dimensions. These items were then evaluated by eight generative AI models to analyze their congruence with theoretical dimensions, ensuring content validity. Subsequently, a second group of AI models assessed the clarity, theoretical connection, discrimination, writing quality, and usefulness of the items.

The results showed that the AI-based methodology effectively identifies items with high levels of content validity in various AI application areas, such as social media, virtual assistants, and entertainment. However, challenges were detected in the discrimination of certain items and the content validity of some factors within Bloom's Taxonomy, particularly in the levels of remembering and applying.

Compared to previous studies, this research demonstrates that generative AI can streamline the content validation process, allowing for the evaluation of a large number of items in less time than human expert panels. The methodology of the present study ensures a rigorous statistical analysis, including the Content Validity Ratio (CVR), the Content Validity Coefficient (CVC), and Aiken's V.

Among the study's limitations is the dependency on AI model training and the variability of their responses. However, this methodology opens new possibilities for the creation and validation of assessments in education and psychology. In the future, it is recommended to compare AI-generated judgments with those of human experts to assess reliability and explore AI applications in other fields of knowledge.

The poster will include examples of interactions with AI and the results of the content validity of the proposed items.

Keywords

Validity, Artificial Intelligence, Literacy, Measurement

Primary author: Dr ORDÓÑEZ, Xavier G. (Complutense University of Madrid)

Co-author: ROMERO MARTÍNEZ, Sonia Janeth (Universidad Nacional de Educación a Distancia)

Presenter: Dr ORDÓÑEZ, Xavier G. (Complutense University of Madrid)

Session Classification: Poster Session 2

Track Classification: Measurement: Measurement