Society for Research on Educational Effectiveness

Focus on AI: Research Methods Webinar Series

Webinar 3

Title: Introducing pairadigm: A Coding Package for Measurement of Scalar Constructs Using Large Language Models

Date/Time: July 14, 2026 | 11:00am – 12:30pm Eastern Time

Presenter: Michael L. Chrzan, Senior Data Scientist at the Center for Educational Data Science and Innovation (EDSI) at the University of Maryland

Abstract:
As Large Language Models (LLMs) are increasingly adopted for qualitative coding and scoring in education research, the field faces a critical challenge: moving beyond opaque automation toward rigorous, verifiable measurement.

This webinar introduces pairadigm, an open-source package designed to implement Concept-Guided Chain-of-Thought (CGCoT) workflows for systematic text evaluation. Unlike standard prompting techniques that output raw scores with little justification, pairadigm grounds LLM evaluations in comparative judgment. The framework utilizes pairwise comparisons - a method well-established in psychometrics (e.g., Thurstone, Bradley-Terry) - to reduce calibration drift and enhance inter-rater reliability. By forcing models to explicitly reason about the presence of specific concepts using researcher-crafted guidance before adjudicating between two text artifacts, the tool produces interpretable, granular decision trails.

In this session, we will see how education researchers can use pairadigm to scale their measurement efforts without sacrificing validity. We will cover the end-to-end workflow: defining and generating construct-specific breakdowns, generating comparisons and rationales using LLMs (supporting OpenAI, Anthropic, Gemini, and local Hugging Face or Ollama model backends), and, crucially, validating machine outputs against expert judgments. Lastly, attendees will learn how to use pairadigm to compute Bradley-Terry scores from the final pairwise decisions to derive interval-scale measures from pairwise data, providing a robust methodology for integrating LLM-generated annotation data into their studies.