

3 - 5 October 2025

Milan , Italy

Using Ai Tools in Educational Assessments and Assignments: a Comparative Study

Karen Kar Lin Hor

Singapore University of Social Sciences

Abstract

Generative AI is commonly used to complete assessments and assignments by students in educational settings. This makes assessment integrity a challenge and thus teachers' accuracy in assessing students' learning. This paper compares written responses by 17 ten- to twelve-year-old students and AI's responses in open-ended mathematics questions. Open-ended questions allow for various response options, do not conform to a set answer and allow the individual to respond however one chooses. Here, the study uses three different types of open-ended mathematics questions converted from commonly found closed questions in Singapore primary schools' mathematics textbooks. The three different types of open-ended questions are: (1) problems with missing data or hidden assumptions, (2) problem requiring an explanation of a concept, procedure, or error and (3) problem-posing. The AI tools used here which are compared with human responses are ChapGPT, Barb (now called Gemini) and Deepseek.

Findings suggest that AI struggles with logical thinking grounded in human experiences. The implication here suggests redesigning assessments to emphasize human experiences which are difficult for AI to replicate, that is, design tasks that require personal connections to lived contexts. The findings here do support several educational theories. One instance is John Dewey's (1938) educational theory on experience and education, and the other is Vygotsky's (1978) sociocultural theory of cognitive development.

Keywords: Keywords: Artificial Intelligence (Ai); Assessments; Education; Lived Experience; Openended-ended Questions