

Editor's Comment:

The paper addresses an important and relevant issue in education today: the comparison between AI-powered large language models (LLMs) and human graders in higher education settings. Artificial intelligence (AI) is being increasingly used in education, including tasks like grading.

The paper attempts to evaluate the potential of Large Language Models (LLMs) to perform grading tasks efficiently and reliably compared to traditional human graders, considering factors such as accuracy, fairness, and bias.

The methodology was sound and followed the steps of good scientific research. All the reviewers were very positive in their evaluation and a few very minor changes were recommended. These were addressed in the revised manuscript.

I therefore support publication in the Journal.

Editor's Details:

Dr. Nasser Mustapha
Senior Lecturer, University of the West Indies, Trinidad and Tobago.