ChatGPT: The Future of Comprehensive Foundational Science Question Banks

Thekra N. Hindi, MS2 and Kori L. Brewer, Ph.D. 2

1Brody School of Medicine and 2Department of Emergency Medicine at East Carolina University, Greenville, NC

IDEA

Generating questions using ChatGPT, an AI language model, would provide students and faculty with additional question banks for practice and exam purposes.

The purpose of this project is to determine whether AI-generated questions are equivalent to faculty-written questions regarding quality and student performance.

METHODS

Part 1: Pre-work

- Goal: create a template that will consistently make relevant questions of appropriate difficulty.
- Process: Create as many questions as possible given a certain topic to identify factors needed to produce desired question types.
- Modifications required to make the questions will be noted.
- Once noted, attempts to decrease the modifications to reach the desired result will be made.
- After determining the consistency of results by verifying with the principal investigator (PI), a template will be created to unify the results produced regardless of the question creator.

Part 2: Question-Generation

- Goal: create questions using ChatGPT and to obtain questions written by faculty members about the same topic.
- Process: The select group of faculty will be divided into two groups.
- Faculty-written questions will be provided by the first group of faculty.
- AI-generated questions will be produced by the primary researchers for this project using the template created in part 1.

Part 3: Question-Testing

- Goal: determine the quality of the AI-generated questions prior to evaluating student performance with the generated questions.
- Process: The second group of faculty previously selected will evaluate the two groups of questions that are deidentified using a rubric. Some of the faculty will be experts on the subject that is discussed in the group of questions, and the other part of the faculty will be non-experts.
- Experts will determine the substance of the question and the accuracy of the answer choice that is selected as the best answer choice.
- Non-experts will determine how the question may be interpreted by students and the quality of the question.

POTENTIAL IMPACT

Increased Access to Practice Resources: Resources can be limited, and having an AI system generate practice questions can help students gain more exposure.

Efficient Learning and Revision: Students can use AI-generated questions for regular self-assessment and revision.

Improving Critical Thinking: Practice questions can challenge students to think critically, apply complex medical scenarios, and apply their knowledge to practical situations.

Augmenting Classroom Instruction: Practice questions can complement traditional teaching methods by providing additional resources for medical students to reinforce their understanding of medical concepts.

Addressing Question Scarcity: Medical school practice questions are often limited. ChatGPT can help alleviate this issue by generating a vast number of questions on various medical topics, ensuring a broader pool of resources for students.

Identifying Weak Areas: Through AI-generated practice questions, students can identify their weak areas and knowledge gaps.

EVALUATION PLAN

Part 5: Question Performance

- Goal: determine how questions that are AI-generated compared to faculty-written questions with regard to student performance.
- Process: Both groups of questions will be presented to students using Examplify, which is an assessment-taking software.
- The questions metrics will be compared to determine what kind of difference is observed based on student performance.
- The biserial and discrimination factors will both be evaluated.
- A positive biserial above 0.4 will be associated with a well-performing question.
- A positive discrimination factor above 0.35 will be associated with a well-performing question.
- If both of those values are lower, then a sum of both values being below 0.7 is associated with a well-performing question.

Part 6: Resource Utilization

- Goal: determine the likelihood that this resource will be utilized as outlined in this project.
- Process: A survey will be sent to both students and faculty.
- Students will be asked if they used ChatGPT to generate practice questions for course that lack practice-question resources.
- Faculty will be asked if they used ChatGPT to generate exam question banks to renew exam questions.

NEED/RATIONALE

ChatGPT is an excellent, interactive medical education tool that supports learning due to its natural language processing capabilities and ability to understand context.1

New questions are important in the learning process to provide additional support to students and faculty who write some of the questions currently provided to students.2 Some courses lack practice questions, making this an important need.

REFERENCES