

# ChatGPT: The Future of Comprehensive Foundational Science Question Banks

Thekra N. Hindi, MS2<sup>1</sup> and Kori L. Brewer, Ph.D.<sup>2</sup>

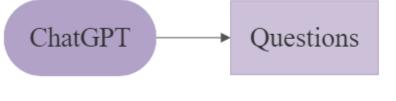
<sup>1</sup>Brody School of Medicine and <sup>2</sup>Department of Emergency Medicine at East Carolina University, Greenville, NC

Thekra Hindi, MS2 **Brody School of Medicine** East Carolina University Greenville, North Carolina 27858 hindit19@students.ecu.edu

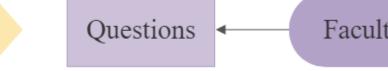
## IDEA

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 AIgenerated questions are equivalent to faculty-written questions regarding quality and student performance.







## 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.<sup>2</sup> Some courses lack practice questions, making this an important need.



Automated question generation: AI model can produce questions in a manner similar to how a human educator would create them.



**Diverse question types:** This diversity allows for a well-rounded set of practice questions to test different aspects of a student's understanding.



Adaptability to Difficulty Levels: ChatGPT can adjust the difficulty level of the questions it generates.



**Instant Feedback and Solutions:** ChatGPT can be programmed to provide instant feedback and solutions for the practice questions it generates.



Continuous Learning: Over time, ChatGPT can improve its question generation abilities and provide even better practice questions.



Topic Exploration: ChatGPT can help students explore different topics by generating questions that encourage critical thinking and deeper understanding.



Supplementing Educational Resources: ChatGPT's question generation capabilities can complement existing educational resources by providing additional practice materials.<sup>3</sup>

## METHODS



- 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.
- o Once noted, attempts to decrease the modifications to reach the desired result will be made.
- o 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.



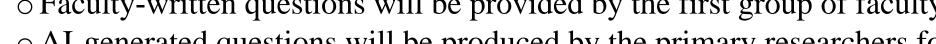
Part 1: Pre-work





- 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.

o Faculty-written questions will be provided by the first group of faculty. Part 2: Question-





Generation

o AI-generated questions will be produced by the primary researchers for this project using the template created in part 1.





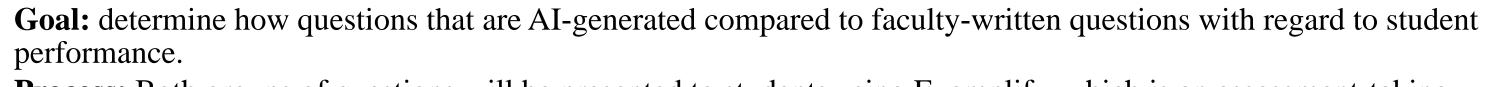
- Goal: determine the quality of the AI-generated questions prior to evaluating student performance with the generated
- **Process:** The second group of faculty previously selected will evaluate the two groups of questions that are deidentified using a rubric. Some of the faculty will be experts on the subject that is discussed in the group of questions, and the other part of the faculty will be non-experts.



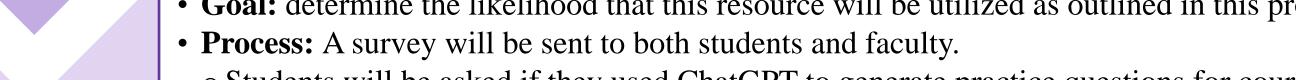


- Experts will determine the substance of the question and the accuracy of the answer choice that is selected as the best answer choice.
- Non-experts will determine how the question may be interpreted by students and the quality of the question.

## EVALUATION PLAN



- Process: Both groups of questions will be presented to students using Examplify, which is an assessment-taking
- o 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<sup>4</sup>: Part 5: Question
  - A positive biserial above 0.4 will be associated to a well-performing question.
  - A positive discrimination factor above 0.35 will be associated to a well-performing question.
  - If both of those values are lower, then a sum of both values being a positive value equal to or higher than 0.7 is associated to a well-performing question.



### Part 6: Resource Utilization

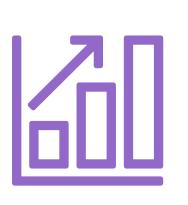
**Performance** 

稟



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

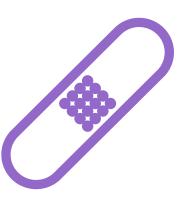
## POTENTIAL IMPACT



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

Personalized Learning: Each student can receive practice questions suitable for their current level of understanding.

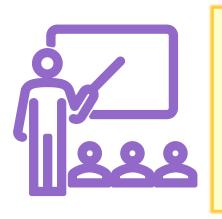




Efficient Learning and Revision: Students can use AI-generated questions for regular self-assessment and revision.

Improving Critical Thinking: Practice questions can challenge students to think critically, analyze 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.

## REFERENCES

- Gilson, A., Safranek, C., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2022). How well does ChatGPT do when taking the medical licensing exams? The implications of large language models for medical education and knowledge assessment. medRxiv, 2022-12.
- Khogali, S. E. O., Davies, D. A., Donnan, P. T., Gray, A., Harden, R. M., McDonald, J., ... & Yu, N. (2011). Integration of e-learning resources into a medical school curriculum. *Medical Teacher*, 33(4), 311-318.
- Ruksakulpiwat, S., Kumar, A., & Ajibade, A. (2023). Using ChatGPT in Medical Research: Current Status and Future Directions. Journal of Multidisciplinary Healthcare, 16, 1513-1520. https://doi.org/10.2147/JMDH.S413470
- Taib, F., & Yusoff, M. S. B. (2014). Difficulty index, discrimination index, sensitivity and specificity of long case and multiple-choice questions to predict medical students' examination performance. Journal of Taibah University Medical Sciences, 9(2), 110-114.