Assessing ChatGPT's Response Accuracy in Pharmacy Education: Insights into Cognitive Complexity

Tuan Tran, Ph.D., Victor Phan, PharmD, Uyen Le, Ph.D.
California Northstate University College of Pharmacy, Elk Grove, CA 95758

PURPOSE

- To evaluate the accuracy of ChatGPT's responses in pharmacy education using standardized exam questions.

OBJECTIVE

- Assess the accuracy of ChatGPT responses to standardized questions across various levels of Bloom's Taxonomy within the pharmacy program.
- Help educators and practitioners to enhance their understanding of ChatGPT's capabilities and develop guidelines to help students effectively integrate ChatGPT into their study routines and clinical practice.

METHOD

- **Exam Question Creation**: We administered 120 questions to ChatGPT.
  - The questions covered 3 main topics: biostatistics, calculations, and therapeutics.
  - Questions were created based on the complexity of Bloom's Taxonomy: Recall, Understand, Apply, and Analyze.
  - For conceptual questions, the format is "multiple-choice" or "multi-selection" with four possible answer choices. For calculation questions, the format is "fill-in-the-blank".
- **Data Analysis:** Responses from ChatGPT were evaluated by faculty members with relevant expertise.
  - **Question Administration**
    - 120 multiple-choice questions were fed into the ChatGPT3.5 prompt asking to select the correct options listed in each question or provide correct answers for calculations.
    - Faculty members evaluated the responses from ChatGPT to determine their correctness.
  - **Response Evaluation**: The responses by ChatGPT were evaluated by faculty members and classified into the following categories:
    - **Correct**: ChatGPT selects or provides correct answers.
    - **Incorrect**: ChatGPT selects or provides incorrect answers.
    - **Partially Correct**: ChatGPT selects one or more but not all correct options (for Multi-selection questions).

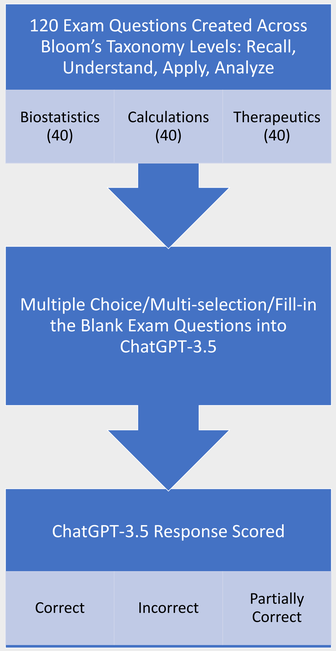


Figure 1. The Process of Question Creation, Administration, and Response Analysis.

RESULT

- The result showed that ChatGPT's responses to 120 questions were 65% (n = 78) correct, 29.17% (n = 35) incorrect, and 5.83% (n = 7) partially correct

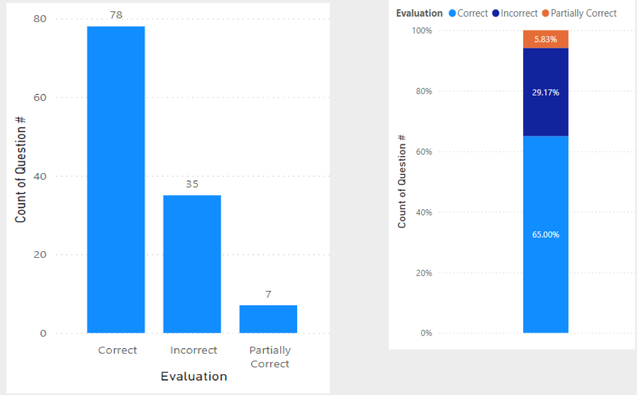


Figure 2. Evaluation of ChatGPT's Response Correctness: 65% - Correct, 29.17% - Incorrect, and 5.83% - Partially Correct.

- Biostatistics has 70% (n = 28) accuracy, calculation 65% (n = 26), and therapeutics the lowest at 60% (n = 24)

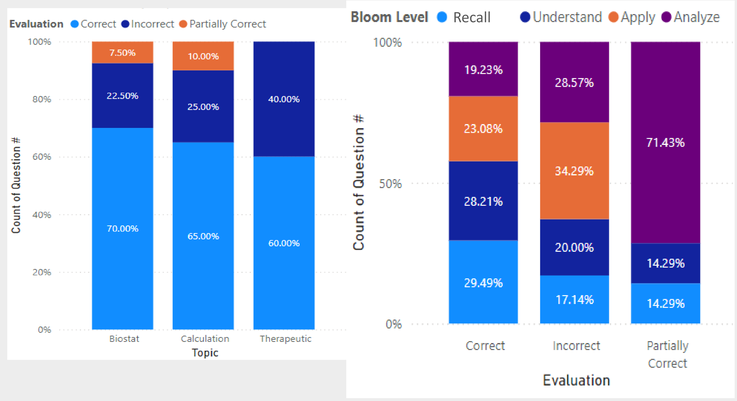
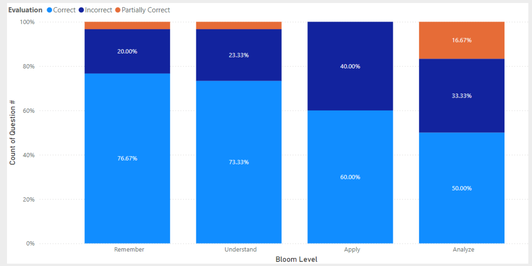


Figure 3. Evaluation of ChatGPT's Response by Topics: Correctness goes down from biostatistics to calculation to therapeutics .

RESULT (CONT.)

- Accuracy varied across Bloom's Taxonomy, from 76.67% (n = 23) in Recall and 73.33% (n = 22) in Understand to 60% (n = 18) in Apply and 50% (n = 15) in Analyze.



Evaluation	Recall	Understand	Apply	Analyze	Total
Correct	23	22	18	15	**78**
Incorrect	6	7	12	10	**35**
Partially Correct	1	1		5	**7**
Total	**30**	**30**	**30**	**30**	**120**

Evaluation	Recall	Understand	Apply	Analyze	Total
Correct	76.67%	73.33%	60.00%	50.00%	**65.00%**
Incorrect	20.00%	23.33%	40.00%	33.33%	**29.17%**
Partially Correct	3.33%	3.33%		16.67%	**5.83%**
Total	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**

Figure 4. Evaluation of ChatGPT's Response by Levels of Bloom's Taxonomy.

CONCLUSION

- Our study reveals ChatGPT as a valuable tool in pharmacy education, demonstrating a 65% accuracy rate across a variety of exam questions.
- The performance of ChatGPT, however, declines as the complexity of questions increases, particularly with higher-order cognitive cases as categorized by Bloom's Taxonomy.
- Our future work will expand the framework into the entire pharmacy curriculum.

ACKNOWLEDGMENTS

- Dr. Tran's work was supported in part by the 2024 CNUCOP Faculty Development Fund.
- Dr. Phan's travel and participation was supported by the CNUCOP Dean Office

REFERENCES

[1] ChatGPT Platform, https://openai.com/research/

[2] Bloom's Taxonomy Alternate Version, https://www.acpe-accredit.org/pdf/Blooms_Taxonomy.pdf

[3] Don Roosan, Pauline Padua, Raiyan Khan, Hasiba Khan, Claudia Verzosa, Yanting Wu, "*Effectiveness of ChatGPT in clinical pharmacy and the role of artificial intelligence in medication therapy management*," Journal of the American Pharmacists Association, 2024.