

Exploring the Use of ChatGPT to Analyze Student Course Evaluation Comments



Kathryn Fuller, PharmD*; Kathryn A. Morbitzer, PharmD, MS*; Jacqueline M. Zeeman, PharmD;
Adam Persky, PhD; Amanda Savage, PharmD; Jacqueline E. McLaughlin, PhD, MS
UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC

KEY FINDINGS

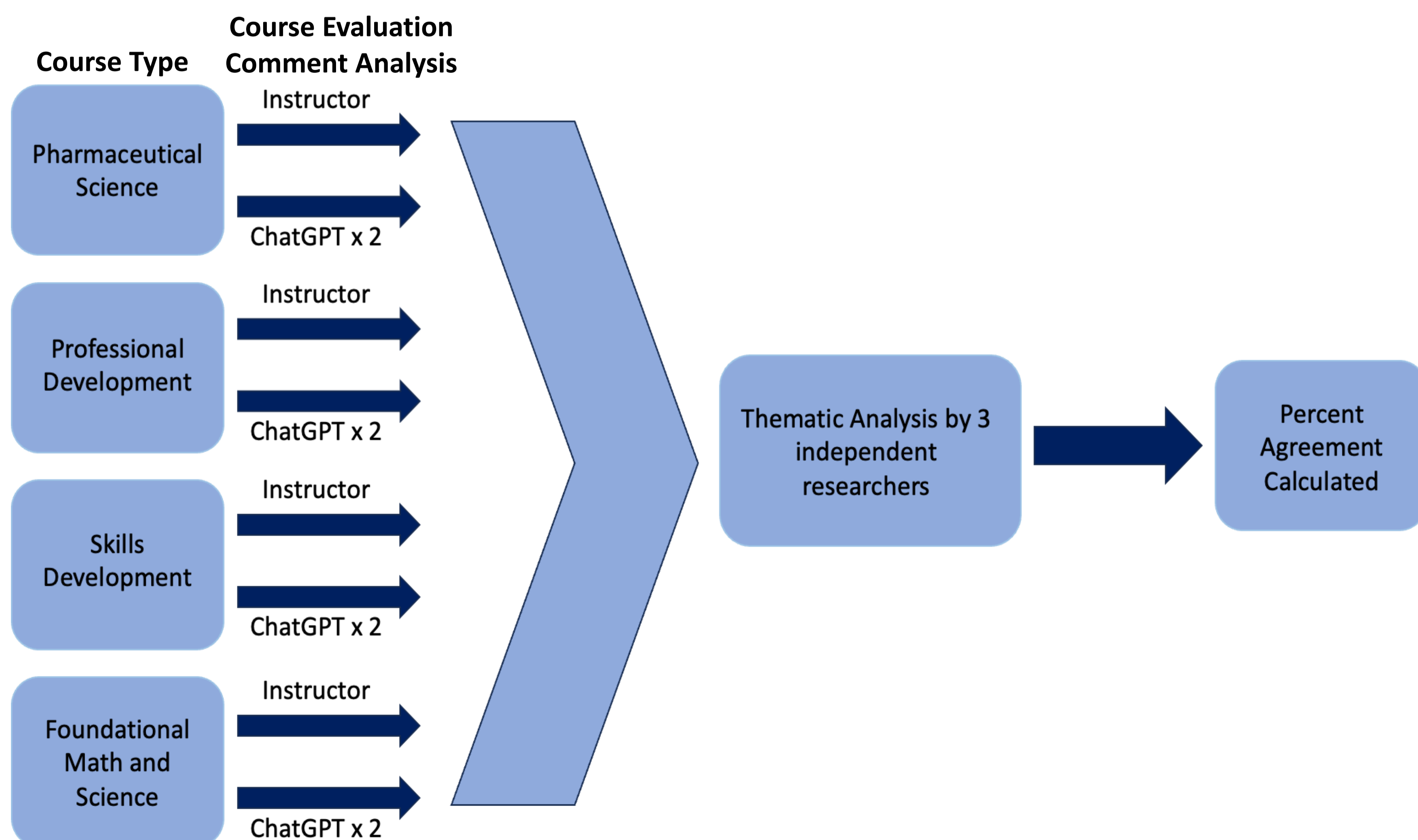
- There was **high agreement** between the instructor and ChatGPT results - with the highest agreement for course-related topics and lowest agreement for weaknesses of the course.
- ChatGPT identified **more themes**, on average, than the faculty themselves, as well as provided **more depth** and detail of the themes.
- Instructors took **27.5 minutes on average** per course to analyze their data while the **ChatGPT users took on average 10.5 minutes and 12.5 minutes** for the two accounts.
- ChatGPT **lacks contextual understanding** and may incorporate biases from the data they are trained on

Why did we do this study?

- While faculty have acknowledged the value of course evaluations, the process of identifying useful information is often time consuming and difficult
- ChatGPT's natural language processing capabilities offer an efficient alternative to text mining, automating the extraction of meaningful insights from student course evaluations

What did we do?

- Four instructors from the UNC Eshelman School of Pharmacy independently analyzed student course evaluation comments for one of their own courses
- Two ChatGPT user accounts were utilized with slightly different prompts to explore variations



What did we find?

| Examples of Course Changes Identified by Instructors and Related ChatGPT Findings | | |
|---|--|---|
| Example Change (Instructor): Please list 3-4 changes you might make to your course or teaching based on the course evaluation comments: | Example Change (Chat1): Please list 3-4 changes I might need to make to my course or teaching based on the course evaluation comments. | Example Change (Chat2): I am a professor at a pharmacy school teaching a [Course Title] course. Please list 3-4 changes you might make to your course or teaching based on the course evaluation comments: |
| End class on time | Address the concerns regarding the early morning class time and strive to end classes on time. Consider rearranging the schedule or providing more breaks to ensure that class activities fit within the allotted time frame. | Address the concerns raised by students regarding the course going past the scheduled time or feeling dragged on. Consider implementing strategies to manage time more effectively during class sessions, ensuring that topics are covered within the allocated time frames and maintaining an engaging pace throughout the course. |
| Provide more practice problems | Students expressed a desire for more interactive and hands-on learning opportunities. Incorporating more group work, class discussions, and hands-on activities can foster student engagement and deeper understanding of the course material. Consider integrating activities that encourage active participation and collaboration among students. | Identify [topics] where students have expressed a desire for more practice questions. Provide additional practice resources to reinforce their understanding of these challenging topics and support their learning process. |

| Prompt Focus | Agreement between Instructors and ChatGPT | | | |
|--------------------------------------|---|-----------|-------------|-----------|
| | Agreement – M±SD | | | |
| | I/Chat1 | I/Chat2 | Chat1/Chat2 | All |
| Course-related Topics | .80 ± .17 | .71 ± .14 | .82 ± .14 | .89 ± .04 |
| Themes related to Student Experience | .82 ± .14 | .60 ± .19 | .73 ± .13 | .85 ± .08 |
| Strengths of the Course | .68 ± .11 | .71 ± .12 | .92 ± .17 | .88 ± .06 |
| Weaknesses of the Course | .53 ± .06 | .57 ± .08 | .81 ± .24 | .80 ± .04 |
| Changes You Might Make | .71 ± .28 | .58 ± .29 | .89 ± .13 | .86 ± .10 |

I = instructor; Chat1 = ChatGPT account 1; Chat2 = ChatGPT account 2; M = Mean; SD = Standard Deviation

Highest agreement between instructors and the ChatGPTs was for course-related topics and lowest agreement was for weaknesses of the course

ChatGPT identified more themes, on average, than the faculty themselves, as well as provided more depth and detail of the themes. While the instructor may have identified one suggested change, both ChatGPT accounts identified the suggested change along with potential solutions

This study found that anonymizing the data, formatting for submission, and analyzing via ChatGPT **required less than half the time** instructors required to analyze the data.

What does it mean?

- This suggests that ChatGPT can effectively assist in the thematic analysis of student comments and streamline the process all while potentially reducing the burden on faculty members
- ChatGPT is comparable to human analysis at capturing feedback related to the course content, teaching methods, and overall course experience²²
- ChatGPT does not have the ability to “feel” the impact of a negative comment and is not influenced by emotional factors, which may therefore explain the low agreement for weaknesses of the course
- Cleaning the data and creating effective prompts are essential steps to enhance the quality and relevance of ChatGPT-generated content for qualitative analysis

Next Steps

Future research should explore how prompt language may impact the themes yielded and the integration of ChatGPT into health profession education program workflows to further assess its impact on course quality improvement and faculty workload.

REFERENCES

A. Fuller, K., Morbitzer, K.A., Zeeman, J.M. et al. Exploring the use of ChatGPT to analyze student course evaluation comments. *BMC Med Educ* 24, 423 (2024). <https://doi.org/10.1186/s12909-024-05316-2>