# Chatbots can increase student learning

Danielle Collins[a], Cole English[a], Iskande V. Larkin[a], Heather T.D. Maness[b], Patrick Larkin[a]
a College of Veterinary Medicine, b Information Technology

## Background

Chatbots are gaining popularity in educational settings and can potentially have a large effect on student learning outcomes by closing the gap between learner-instructor presence in online learning and their ability to personalize student learning (Ondáš et al., 2019; Pereira et al., 2023; Wu & Yu, 2024).

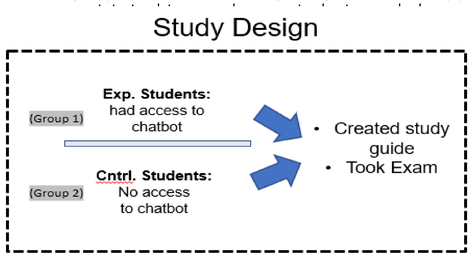
There are conflicting reports on the learning gains of using chatbots in education. Some studies have found that grades decreased with the use of ChatGPT, even though students found it useful (Forero & Herrera-Suárez (2023), others found no difference in knowledge scores or learning outcomes (Han et al., 2022; Kumar 2021), while others found that student learning and motivation increased with the use of chatbots (Chang et al., 2022; Yin et al., 2021).

The **Aim** of our study was to *understand if the use of chatbots specifically trained on lecture material could increase student learning outcomes, as measured by exam scores.* A survey was ~~also~~ given to gain a better understanding of students' perceptions and usage of chatbots.

***Our hypothesis is: chatbots that are trained on specific lecture content in a class will help increase student learning***

## Methods

- ChatBotKit™ was used to train ChatGPT4 on lecture transcripts from the "The Dog" (VME3001) and "The Cat" (VME3002) classes at the University of Florida.
- Experiments were conducted during the Summer 2024 semester (Dog Class) and in the Fall 2024 semester (Cat Class). 47 total students participated in these studies.
- Terracotta™, a LMS plugin, was used to create randomized experimental and control groups in Canvas. The study design, which was flipped during th




- Students were provided with information on how to use the chatbots, including a training video
- They were prompted to use the chatbot to complete study guides as a study tool for the exam
- Students received a survey at the end of the semester to learn more about their previous AI use and their perceptions before and after using our chatbot
- After the study, student performance and engagement data were exported from Terracotta into JSON format. The JSON file was converted to Excel CSV format using a novel Python program. Data from the excel CSV file was then used for statistical hypothesis testing.

## Results

- **Only a subset (60%) of the experimental students used the chatbots.** Therefore, students were divided into three groups: (0) Control (no access to chatbot), (1) Experimental (access to chatbot, but did not use), (2) Experimental (used the chatbots) [Table A]
- When exam performance of both Experimental groups were compared to students in the control group, **a significant difference was found in students in the experimental group using the chatbot tool compared to students in the experimental group not using the tool (P<0.04)** [Table B]

A

| Group | Least Sq Mean | Std Error |
|---|---|---|
| Exp., used chatbot (2) | 83.73 | 3.45 |
| Exp., did not use chatbot (1) | 75.95 | 3.73 |
| Control (0) | 81.8 | 3.05 |

B

| Level | Difference | p-Value |
|---|---|---|
| 2 vs. 1 | 7.78 | 0.04 |
| 0 vs 1 | 5.84 | 0.06 |
| 2 vs 0 | 1.93 | 0.45 |

- **Overall, students in the experimental group had a 10.2% improvement in their exam performance compared to students that did not use the chatbot (group 2 vs. 1)**

Figure 1: Have you ever used a Chatbot and/or AI tool for educational use before this class?

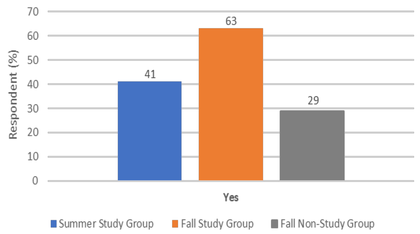



Figure 2: Which of the following best describes your preference for using the Chatbot tool for completing your study guide?

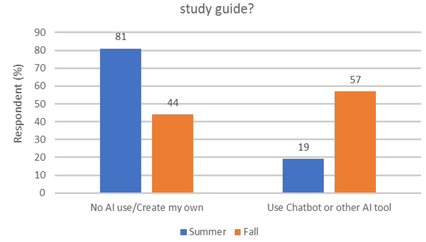
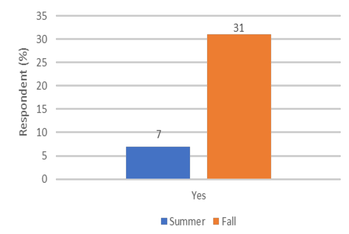



Figure 3: Did you use the Chatbot beyond the study guide assignment to prepare for the exam?

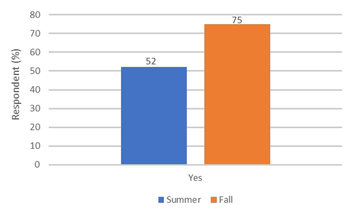



- **Not all students had used a Chatbot or AI tool before our class**. This was lower in the Summer study group (41%) compared to the Fall study group (63%) [Figure 1]. Interestingly, students with previous AI tool experience appeared to be more inclined to participate in the Chatbot study. In the Fall study group 63% had used an AI tool before, whereas in the non-study group only 29% had used an AI tool previously [Figure 1]

- When we asked students their preference in creating their study guides, **there was a shift from less than 20% of students in the Summer to over half (57%) of the students in the Fall preferring to use AI tools** [Figure 2, right side of graph]

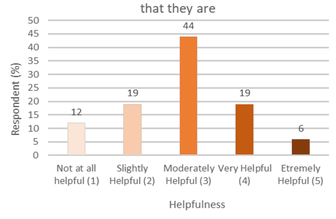
- In the Fall, **more students used the Chatbot to prepare for the exam beyond the study guide** (Summer 7% vs Fall 31%) [Figure 3]. Students also found the chatbot more instrumental in doing well on the exam in the Fall group (3.4 ± 1.09) than the Summer group (2.48 ± 1.09). This may correlate with additional emphasis by instructors using AI training videos in the fall class (9.5 minutes long) compared to the summer class (only 3.5 minutes long) to showcase the advantages to students for using chatbots.

Figure 4: Would you like to see more Chatbot integrations in future courses?



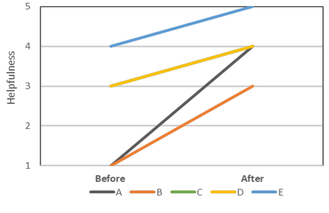

A

Figure 5: Before using the chatbot in this course, my general opinion of chatbots was that they are




B

Figure 5: Helpfulness Perspective Change




- **Students overwhelmingly (>50%) wish to see chatbots used in future courses** [Figure 4]
- Students found the chatbot moderately helpful before using chatbots in our class [Figure 5A].
- When we asked students if they had a perception change after using the chatbot in our class, 5 students said yes. **All 5 students that addressed that survey question had a positive increase in their perception of the helpfulness of chatbots** [Figure 5B]
- Finally, we asked students what they liked about our chatbots, and what they would like to see changed. These were our top responses:

"What did you find most beneficial about the Chatbot?"

- Helped clarify complex concepts… (69%)
- Receive answers anytime 24/7 … (56%)
- Made course content feel more interactive…(25%)

"What would you change for future Chatbot uses in this course?"

- Improve ability to handle complex questions… (60%)
- Improve ability to summarize/review concepts… (60%)
- More detailed responses/explanations…(47%)

## Discussion/Conclusions

- **Chatbot usage significantly improved exam performance:** Students who used the chatbot scored 10.2% higher on exam questions compared to experimental students who did not use it ($p < 0.04$). When compared to all students not using the chatbot (control + non-users), the performance increase was 6.1% ($p = 0.08$)
- **Usage rates highlight engagement challenges:** Despite being in the experimental group, only 60% of students actually used the chatbot to complete their study guides. This usage gap emphasizes the need for better onboarding and motivation strategies.
- **Training makes a difference:** Between Summer and Fall, students using the chatbot beyond the study guide jumped from 7% to 31%, and those preferring AI for study guides increased from <20% to over 50%. This perhaps correlated with extending the Chatbot training videos used in the class from 3.5 minutes (summer) to 9.5 minutes (fall) and including usage examples like flashcards and practice exams.
- **Prior AI exposure influences participation:** In Fall 2024, 63% of the study group had prior AI experience versus 29% in the non-study group ($p = 0.0631$), suggesting prior exposure to AI tools encourages participation in educational AI studies.
- **Students want more chatbot integration:** By the end of the Fall semester, 75% of the study group said they would like to see chatbots integrated into future courses, compared to 52% in the Summer. This again demonstrates how improved training, via the training videos, may shift student attitude.

## Tool Recommendations

- **Training students how to use chatbots effectively is important**
  - How to access and use chatbot, types of questions to ask, examples, prompt engineering, re-prompting, etc.
- **Students may benefit from a trial run of the chatbot before first official use & so may instructors** (see what works well and what doesn't)
- **Use incentives and rewards for using the chatbot to increase student motivation and engagement** (i.e. extra credit points)
- **Survey your students.** Find out what they like and what they don't

## References

1. Chang, C. Y., Hwang, G. J., & Gau, M. L. (2022). Promoting students' learning achievement and self-efficacy: A mobile chatbot approach for nursing training. *British Journal of Educational Technology*, *53*(1), 171–188. https://doi.org/10.1111/BJET.13158
2. Forero, M. G., & Herrera-Suárez, H. J. (2023). ChatGPT in the Classroom: Boon or Bane for Physics Students' Academic Performance? *ArXiv Preprint ArXiv:2312.02422*. http://arxiv.org/abs/2312.02422
3. Han, J. W., Park, J., & Lee, H. (2022). Analysis of the effect of an artificial intelligence chatbot educational program on non-face-to-face classes: a quasi-experimental study. *BMC Medical Education*, *22*(1), 830. https://doi.org/10.1186/s12909-022-03898-3
4. Kumar, J. A. (2021). Educational chatbots for project-based learning: investigating learning outcomes for a team-based design course. *International Journal of Educational Technology in Higher Education*, *18*(1), 65. https://doi.org/10.1186/S41239-021-00302-W
5. Ondáš, S., Pleva, M., & Hládek, D. (2019). How chatbots can be involved in the education process. *2019 17th International Conference on Emerging Elearning Technologies and Applications (ICETA)*, 575–580. https://doi.org/10.1109/ICETA48886.2019.9040095
6. Pereira, D. S. M., Falcão, F., Costa, L., Lunn, B. S., Pêgo, J. M., & Costa, P. (2023). Here's to the future: Conversational agents in higher education- a scoping review. *International Journal of Educational Research*, *122*, 102233. https://doi.org/10.1016/J.IJER.2023.102233
7. Wu, R., & Yu, Z. (2024). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, *55*(1), 10–33. DOI: 10.1111/bjet.13334
8. Yin, J., Goh, T. T., Yang, B., & Xiaobin, Y. (2021). Conversation Technology With Micro-Learning: The Impact of Chatbot-Based Learning on Students' Learning Motivation and Performance. *Journal of Educational Computing Research*, *59*(1), 154–177. https://doi.org/10.1177/0735633120952067

## Acknowledgements

The authors of this paper would like to thank Chris Sharp (Educational Technologist, CITT, University of Florida), for his assistance with the chatbot program selection guidance and creation of the chatbot links, as well as James Colee (IFAS Stats) for all statistical analyses.