Can large language models improve student learning?

Tune H Pers

Faculty of Health and Medical Sciences
University of Copenhagen, Denmark

Motivation and justification

Supervision has been a key cornerstone of student learning throughout history. In the ancient Greece students (also referred to as disciples) followed philosophers such as Socrates for learning and during the Middle Ages apprentices would follow experienced craftsmen to learn back-then complex crafts. Nowadays supervision still is at the heart of student learning: at the universities young researchers typically continuously meet with their supervisors and more experienced researchers meet with their mentors. Across our society, human-to-human master-apprentice ('mesterlære' in Danish) systems are critical components of nearly all efforts that require a complex profession to be mastered.

Large language models (henceforth loosely referred to as artificial intelligence, or AI), such as ChatGPT (OpenAI, 2023), increasingly help to solve complex tasks, which previously only could be solved by humans. ChatGPT and related chatbots that have been trained to provide answers to so-called prompts input by the user. These chatbots have been trained on presumably all publicly available information, which renders them potentially knowledgeable partners when it comes to learning in general. Given that the upper capability limits of (current) AI models have yet to be established, a logical question is whether and how such systems can be used to enhance student learning within the context of master-apprentice setups. In this project, I would like to qualitatively assess whether and how ChatGPT can improve student learning in the context of academic supervision.

To better understand ChatGPT, I consulted resources from the internet and – with my student hat on – I partnered up with ChatGPT to

potentially increase my (and others) learnings from this project. In this project, I used ChatGPT to help me draft questions for the survey sent to students and to make me aware of sentences with grammatical errors.

Theoretical considerations related to student learning

Supervision can be defined as a "method of teaching where the control of and responsibility for the process is more equally divided between the teacher and student than is the case in a teacher-driven lecture or class teaching situation" (Ulriksen, 2014). In general, student supervision involves multiple phases that depend, among other factors, on the students *i)* academic level (bachelor (BSc)-, master-, PhD- and postdoc level). *ii)* the student's knowledge level and autonomy. *iii)* the academic nature of the supervision (e.g. project supervision, PhD-study supervision, scientific supervision). *iv)* the student's and supervisor's personalities, and *v)* their available time (see (Wichmann-Hansen, 2021) for a more complete overview). In order to identify how ChatGPT can improve student learning, I will focus on two major themes in student learning during one-to-one supervision.

A safe learning environment

A safe learning environment is key to student learning during supervision and in general. During supervision there are multiple aspects that foster a safe learning environment. These include setting clear expectations, communicating openly, providing non-judgemental feedback, showing respect and empathy, demonstrating flexibility and maintaining confidentiality (Ulriksen, 2014; Wichmann-Hansen, 2021). As described by Ulriksen (2014) the responsibility during the learning process is shared between the student and the teacher. This means that the student relative to the supervisor should have a significant role in the content and importantly the process of the learning process. According to Ulriksen, the supervisor should "support" the student rather than "persuading" them to learn things in specific ways. In practice, this balance can be affected by interpersonal aspect (2014). Given that students learn content in different ways and at different velocities, it seems reasonable to explore whether AI can enhance student learning by providing a non-

judgemental, omnipresent learning environment, and whether that effect is less pronounced for extrovert students.

Topics amenable to AI-supported student learning

The literature and my experience point to the fact that student learning during supervision is not a one-size-fits-all solution but requires highly flexible techniques that are tailored to the individual student (Wichmann-Hansen, 2021), and a number of contextual factors, many of which have been mentioned above. Given that complexity, tools like ChatGPT may not be amenable to all topics. Hence, delineation of topics for which it works well and those for which it works less well seems to be an important next step in the process towards potentially including this tool as an additional 'partner' in the master-apprentice relationship.

Problem formulation, approach and data collection

I have used a qualitative approach to explore the extent to which ChatGPT can enhance student learning during supervision. With help from ChatGPT, I constructed a questionnaire comprising questions broadly related to the following categories *i)* academic background and learning preferences, and, *ii)* specific interactions and utility of ChatGPT across different contexts.

The questionnaire was administered anonymously, comprised a total of 21 questions, and was designed to take approximately 15 minutes to answer. Please refer to **Appendix 1** for the questionnaire and **Appendix 2** for the email sent to the students and postdocs. I sent the questionnaire to all BSc students, master (MSc) students, PhD students, and postdocs at the Novo Nordisk Foundation Center for Basic Metabolic Research at the University of Copenhagen. The recipients had 14 days to complete the questionnaire.

Results, reflections and future perspectives

A total of 18 MSc students (n=2), PhD students (n=9) and postdocs (n=7) completed the questionnaire. The average response time was about 17 minutes, and most participants answered all questions. A summary of the

results from the questionnaire can be found in **Appendix 3**. An interactive version with the complete results can be found here. The fact that no BSc students and only two MSc students answered the questionnaire means that the results and proposed relationships reported here are predominantly applicable to understand the role of ChatGPT at the PhD-student and postdocs levels, and less to its utility at the BSc and MSc levels.

Based on the reviewing of the results, there are three introductory remarks I wanted to make. First, only 27% of the participants indicated having "reading/writing" learning styles. Because ChatGPT does not include any illustrations in its free version (version 3.5), this suggests that a substantial amount of students interact with ChatGPT in a sub-ideal way. Human supervision more easily allows for visualization, in fact, I typically use the white board in my office typical during all student supervision sessions. Second, 33% of the participants responded "Never" to the question on "How often do you study collaboratively in groups versus studying alone?". This relatively high number most likely reflects the low number of BSc and MSc students among the participants. Finally, all participants have used an online learning or AI-educational tool (presumably ChatGPT) prior to the questionnaire (although I realized that the question was ambiguous in that it is not clear whether prior is relative to the questionnaire or to the invention of ChatGPT). The large fraction of individuals who have used ChatGPT indicates that the participants could be technology-affine early-adopters, and hence might not constitute a representative sample of the general student population. This notion is backed up by the result that programing is by far the most mentioned topic that ChatGPT has been used for among the participants. These limitations aside, there are three major themes related to student learning that are worth highlighting.

Information retrieval

The first is related to the reported utility of using ChatGPT to retrieve information. Several participants used ChatGPT to get some basic information on topics they were less familiar with. On the question on

"Which topics or subjects did you find ChatGPT to be most effective in assisting with? Why?" answers included:

- "Topics where I have the least prior knowledge, and where I'm not sure how to phrase and ask the correct questions. Likely because this is where googling is least helpful - both due to the stricter requirements in question phrasing, and because answers usually require some level of prior knowledge, which can usually be tuned via chatGPT"
- "Summarising scientific concepts while explicitly asking for the sources. It provided a broad understanding of the topic, and a list of potential authors to look out for"
- "Sometimes it helps to get a synopsis on a topic for initial familiarization").

However, most participants found that ChatGPT is more beneficial for topics that they had experience with. Examples answers included:

- "Probably the ones, that are new to me, because it seems to perform better at more general tasks compared to complex and niche tasks"
- "It was easier to know whether it contained mistakes, if I knew the topic".

The answers indicated that the inaccuracy of ChatGPT makes it difficult to use to retrieve insights into topics the students are less familiar with. However, there seemed to be a group of individuals to whom ChatGTP was indeed useful to construct a basic level of knowledge on topics they were less familiar with. On the related question of whether ChatGPT was beneficial for subjects they already had some experience in or for topics that were new to them, example answers in that category of responses included:

- "for those new for me. It does not to the in detail research, but mostly gives an overview about a topic"
- "Most in those that were new to me".

These seemingly contradicting answers underline that students can have very different approaches to learning. Some feel less comfortable using imperfect systems to gain new information, while others seem to accept inaccuracies in the light of receiving new information.

Breakdown of complex problems

The second theme I wanted to highlight is related to questions concerning ChatGPT's usefulness when it comes to breaking down complex questions into more manageable parts. Fifty-three percent of the participants rated ChatGPT as being "Good" when it came to its "ability to explain theoretical concepts versus practical or applied problems". Answers to the related question "For complex topics or multi-step problems (e.g., mathematical problems, scientific processes), how well did ChatGPT aid in your understanding" included:

- "Very well, if you ask multiple (informed) questions."
- "Ok, it often takes more than one question to get there - and you need background knowledge to figure out if the answer is valid or 'good enough'"
- "In general very well!"
- "It was good, I should use it more"
- "Quite well I would say"
- "Quite ok!".

Interestingly, the answers from the other participants were mostly "not relevant, "N/A" or left empty. Together these two categories of answers suggest that those students who have spent more time on ChatGTP also have been more successful in using it to move beyond "simple" tasks and leverage it for more complex tasks that carry greater potential for enhancing student learning. Along the same lines, the utility of ChatGPT to help break down complex problems into more manageable smaller pieces was supported by the result that 53% answered "Good" to the question on "ChatGPT's ability to explain theoretical concepts versus practical or applied problems".

Human vs. machine supervisors

Of the four participants who had characterized themselves as introverted, three answered "Yes" to the question "If you tend to be introverted, did you find using ChatGPT more comfortable than asking questions in a classroom or group setting?" In contrast, no participant answered "Yes" to the question "If you identify as $f\Omega$ extroverted, did you miss any human interaction elements when learning with ChatGPT?" These results could

suggest that i) there is a potential added benefit of ChatGPT to students who are less extroverted and hence may find it more difficult to ask their supervisor about subjects they may perceive as common knowledge, and ii) that the lack of a human supervisor may be tolerable in specific phases of student learning in the context of complicated problems.

Moreover, the answers to the question “How would you describe your level of engagement when interacting with ChatGPT compared to human tutors or peers” divided participants into two groups. One group clearly preferred to human aspect of supervision:

- “Usually go to human tutors or peers first before ChatGPT”
- “It's a tool, can't be compared to any interaction with my peers”
- “Much poorer with Chatgpt than with peers”
- “It’s more like a game of generating text that matches a prompt than actual distillation of knowledge.”

In contrast, the other group found that ChatGPT had utility because of its objectiveness:

- “Very high - much easier to go back and forth a lot of times, and ask "stupid" questions, but also those that are quite complex.”
- “I would ask more questions until being fully clarified from a human. ChatGPT could not fully clarify my doubts, but it helped when breaking down the problems to smaller bits.”
- “Ask way more questions as there is no fear of wasting someones time with silly questions that I know I could easily google”
- “I can ask as many questions I want without feeling bad or wasting someone's time. Then I can ask human tutor if I need personal opinion or discussion”

These responses suggest that there is a sizable group of students who have a substantial benefit of using ChatGTP because they can ask questions they otherwise may not have asked their supervisors. That finding suggests that addition of AI-driven approaches to current supervision models carries potential to increase student learning.

Personal experiences and future perspectives

I am routinely using ChatGPT to explore concepts and topics that I am less familiar with. Personally I am OK with getting answers that I will

need to double check for the correctness. The same is basically true when it comes to answers from peers. I think that ChatGPT and similar approaches offer a very promising addition to traditional supervision approaches and will benefit groups of students who like to explore new complex concepts on their own.

Depending on the nature of the supervision and its phase, the task of the supervisor may vary from simpler tasks such as clarifying concepts to more complicated tasks such as "challenging students' choices, assumptions and prejudices" [2]. Given that the time that a supervisor can offer a given student is limited, the supervision time arguably is best spent by discussions ignited by formative feedback rather than by clarifying concepts. In that light, AI may help students to develop ideas and clarify concepts prior to supervision sessions, which in turn would make them more well-prepared at supervision meetings and may free up time for the advanced aspects of supervision rather than the simpler aspects.

Despite these promising outlooks, I think approaches like ChatGTP will give rise to new major challenges when it comes to student learning. Apart from plagiarism, in my opinion a major challenge will emerge in relation to the way we approach problems. If we stop thinking about potential solutions to complicated problems and simply prompt AI systems for answers, then we over time may lose our human-centric ability to solve complicated novel problems on our own. I do not think that we will skip the thinking process in the foreseeable future just because there is a tool that helps us to break down complex problems into more manageable subparts, but I do think it constitutes a future risk that may leave certain groups of students behind.

References

OpenAI. ChatGPT, 2023.

Ulriksen, L. (2014) Teaching at further and higher education programmes. Frydenlund ISBN 978-87-7118-381-8

Wichmann-Hansen, G. (2021) DUT Guide on Supervision. Dan. Univ. Tidsskr. 16 (31). https://doi.org/10.7146/dut.v16i31.127292.

Appendices

Appendix 1

1. What is your current student/work position?
   - BSc student
   - MSc student
   - PhD student
   - Postdoc
   - Staff scientist

2. Which of the following best describes your preferred learning style?
   - Visual
   - Auditory
   - Kinesthetic
   - Reading/writing

3. How often do you study collaboratively in groups versus studying alone?
   - Daily
   - Weekly
   - Monthly
   - Never

4. Do you consider yourself more introverted or extroverted?
   - Introvert
   - Extrovert
   - In-between
   - I prefer not to answer

5. How comfortable are you asking questions or admitting lack of understanding in a classroom setting?

6. On a scale of 1-10, how persistent are you when faced with difficult academic challenges?
   (Please circle one number from 1 to 10)

7. Have you had prior exposure to online learning tools or AI-based educational tools?

8. List up to three academic topics or subjects you primarily used ChatGPT for during the past three weeks
(Enter your answer)

9. Which topics or subjects did you find ChatGPT to be most effective in assisting with? Why?
(Enter your answer)

10. Were there any topics or subjects where ChatGPT seemed less effective or less helpful? If so, please specify.
(Enter your answer)

11. For complex topics or multi-step problems (e.g., mathematical problems, scientific processes), how well did ChatGPT aid in your understanding?
(Enter your answer)

12. Did ChatGPT's assistance make you more confident in any particular subjects or topics? Which?
(Enter your answer)

13. Based on your experience, are there certain topics or areas you'd recommend other students use ChatGPT for?
(Enter your answer)

14. Were there any unexpected topics or subjects where ChatGPT proved particularly beneficial?
(Enter your answer)

15. How would you rate ChatGPT's ability to explain theoretical concepts versus practical or applied problems?
- Excellent

- Good
- Fair
- Poor
- Very good

16. Did you find ChatGPT more beneficial for subjects you already had some experience in or for those you that were new to you?
(Enter your answer)

17. If you usually study in groups, did you discuss or share ChatGPT's responses with peers?
- Yes
- No
- Maybe

18. If you tend to be introverted, did you find using ChatGPT more comfortable than asking questions in a classroom or group setting?
- Yes
- No
- Maybe

19. If you identify as extroverted, did you miss any human interaction elements when learning with ChatGPT?
- Yes
- No
- Maybe

20. Depending on your personality, did you actively seek feedback from ChatGPT, or did you mainly use it for clarifications?
(Enter your answer)

21. How would you describe your level of engagement when interacting with ChatGPT compared to human tutors or peers?
(Enter your answer)

Appendix 2

From: Tune H Pers tune.pers@sund.ku.dk
Date: Monday, 9 October 2023 at 09.54
To: XXX
Subject: Can ChatGPT enhance academic learning?

Dear CBMR students, PhD-student and Postdocs

I am conducting a survey to assess the utility and impact of ChatGPT, an AI chatbot by OpenAI, on academic learning. I am currently working on a pedagogy project focused on this topic. The survey will take approximately 10-15 minutes of your time. Your feedback will be instrumental in my qualitative analysis to determine how ChatGPT might enhance student learning. I would greatly appreciate it if you could set aside time to complete the survey by Monday, October 23rd. You can access the survey here: https://forms.microsoft.com/e/t9zdAg9b6P.

Please rest assured that your responses will remain completely anonymous. No personal identifiers will be collected. The data will be used solely for research purposes and will be removed after the project's completion on December 1st.

Thank you in advance for your time and valuable input. Should you have any questions or concerns regarding the survey, please feel free to reach out to me.

Warm regards,

Tune

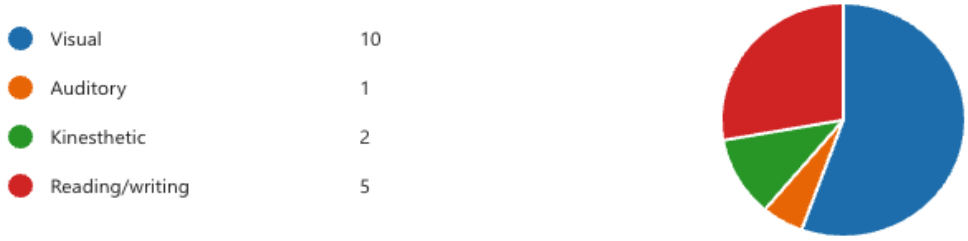
Tune H Pers
Associate Professor, Group leader
University of Copenhagen
Novo Nordisk Foundation Center for Basic Metabolic Research
Faculty of Health and Medical Sciences

Blegdamsvej 3B (Maersk Tower, room 07-8-78)
DK - 2200 Copenhagen N
DIR +45 30 29 49 03
@tunepers
tune.pers@sund.ku.dk
http://cbmr.ku.dk

Appendix 3

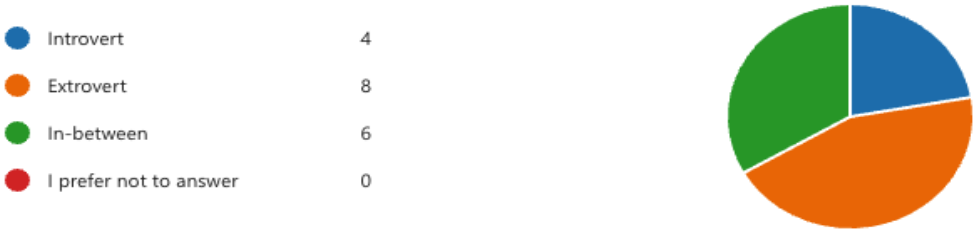
1. What is your current student/work position? (18 responses)

2. Which of the following best describes your preferred learning style? (18 responses)

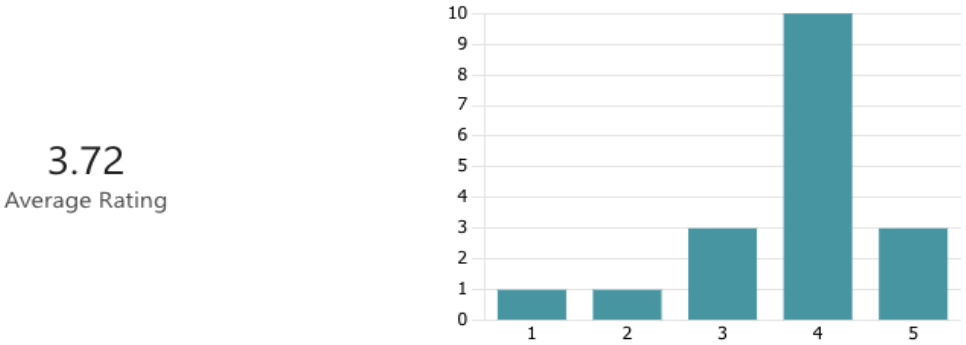


3. How often do you study collaboratively in groups versus studying alone? (18 responses)

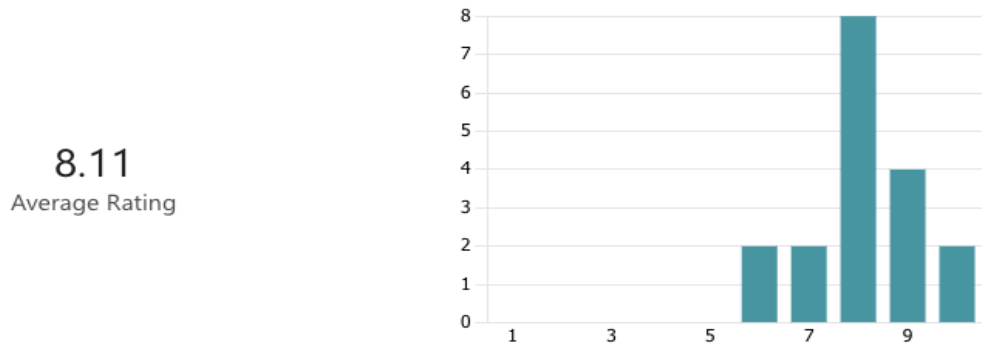
4. Do you consider yourself more introverted or extroverted? (18 responses)



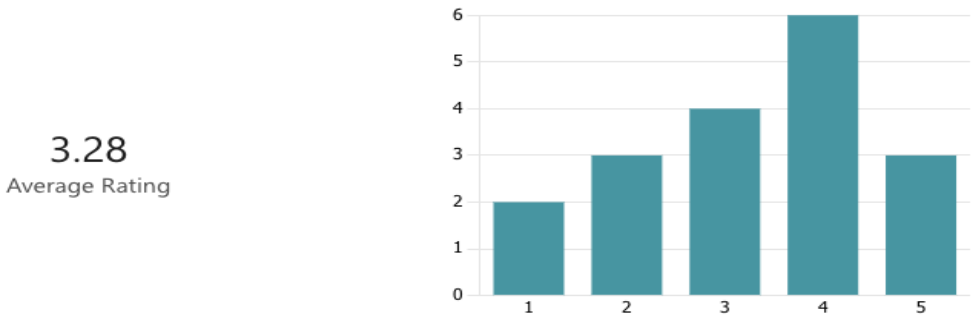
5. How comfortable are you asking questions or admitting lack of understanding in a classroom setting? (15 responses)



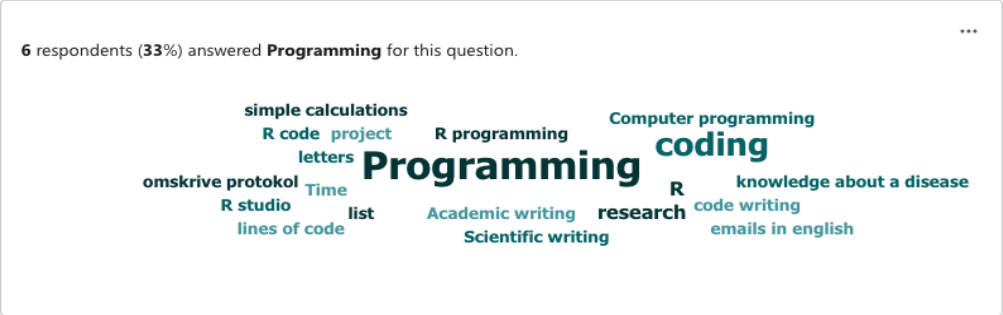
6. On a scale of 1-10, how persistent are you when faced with difficult academic challenges? (18 responses)



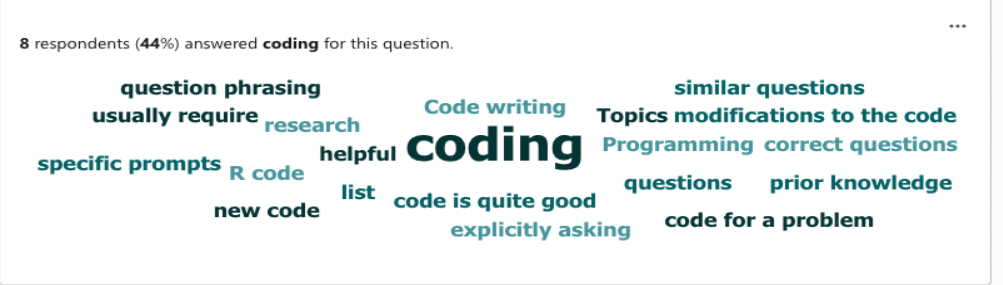
7. Have you had prior exposure to online learning tools or AI-based educational tools? (18 responses)



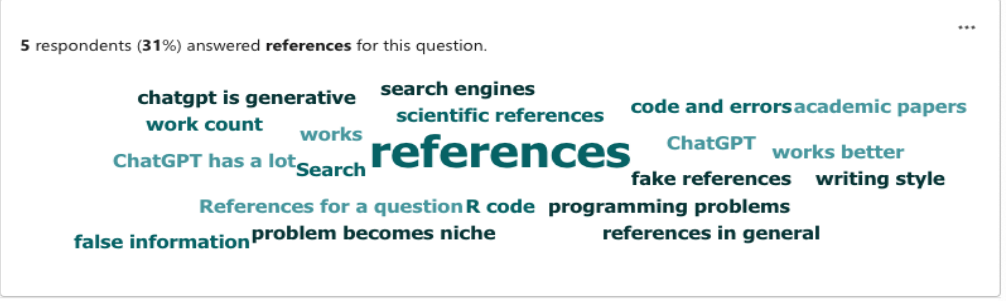
8. List up to three academic topics or subjects you primarily used ChatGPT for during the past three weeks. (18 responses)



9. Which topics or subjects did you find ChatGPT to be most effective in assisting with? Why? (18 responses)

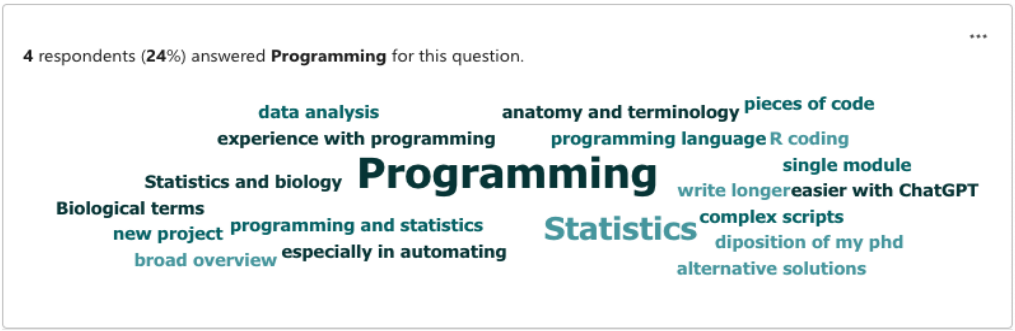


10. Were there any topics or subjects where ChatGPT seemed less effective or less helpful? If so, please specify. (16 responses)



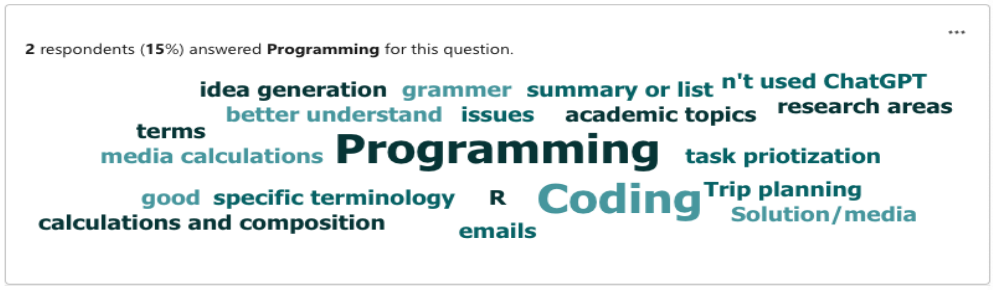
11. For complex topics or multi-step problems (e.g., mathematical problems, scientific processes), how well did ChatGPT aid in your understanding? (16 responses)

12. Did ChatGPT's assistance make you more confident in any particular subjects or topics? Which? (17 responses)



13. Based on your experience, are there certain topics or areas you'd recommend other students use ChatGPT for? (16 responses)

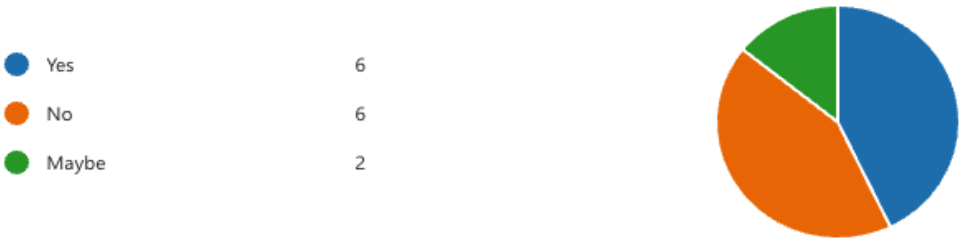
14. Were there any unexpected topics or subjects where ChatGPT proved particularly beneficial? (13 responses)



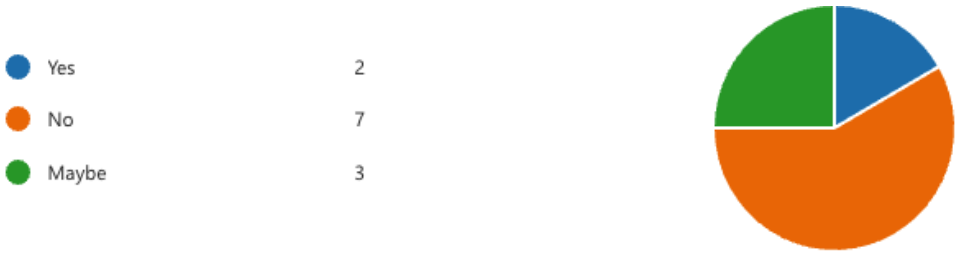
15. How would you rate ChatGPT's ability to explain theoretical concepts versus practical or applied problems? (15 responses)

16. Did you find ChatGPT more beneficial for subjects you already had some experience in or for those you that were new to you? (17 responses)

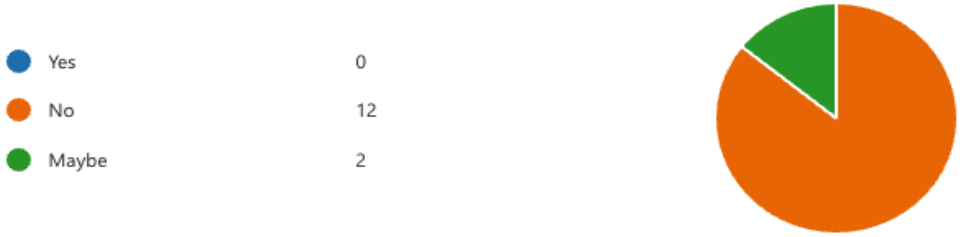
17. If you usually study in groups, did you discuss or share ChatGPT's responses with peers? (14 responses)



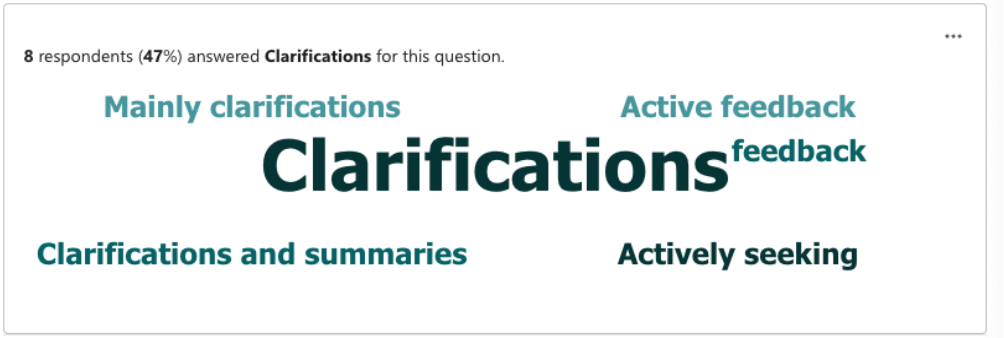
18. If you tend to be introverted, did you find using ChatGPT more comfortable than asking questions in a classroom or group setting? (12 responses)



19. If you identify as extroverted, did you miss any human interaction elements when learning with ChatGPT? (14 responses)



20. Depending on your personality, did you actively seek feedback from ChatGPT, or did you mainly use it for clarifications? (17 responses)



21. How would you describe your level of engagement when interacting with ChatGPT compared to human tutors or peers? (16 responses)