Initial (March 2023) uses and perceptions of ChatGPT in a sample of students and instructors at the University of Bergen (UiB), Norway

Anja Møgelvang, Anne E. Bjune, Raquel Coelho, Sehoya Cotner, David Grellscheid & Ståle Ellingsen

Centre for Excellence in Biology Education

Centre for the Science of Learning & Technology

SLATE Research Report 2023-1

Initial (March 2023) uses and perceptions of ChatGPT in a sample of students and instructors at the University of Bergen (UiB), Norway

© Anja Møgelvang[ab], Anne E. Bjune[b], Raquel Coelho[c], Sehoya Cotner[b], David Grellscheid[d] & Ståle Ellingsen[b] (2023)

[a]Western Norway University of Applied Sciences
[b]bioCEED, Department of Biological Sciences, University of Bergen
[c]SLATE, University of Bergen
[d]Didactics research group, Department of Informatics, University of Bergen

Contact Information:
Sehoya Cotner (Sehoya.Cotner@uib.no)
bioCEED, Department of Biological Sciences
University of Bergen
Postbox 7807
5020 Bergen

ISBN: 978-82-93789-09-3

To cite this report:

Møgelvang, A., Bjune, A. E., Coelho, R., Cotner, S., Grellscheid, D. & Ellingsen, S. (2023). Initial (March 2023) uses and perceptions of ChatGPT in a sample of students and instructors at the University of Bergen (UiB), Norway. *SLATE Research Report 2023-1*, Bergen, Norway: Centre for the Science of Learning & Technology (SLATE). ISBN : 978-82-93789-09-3

ChatGPT Plus with DALL-E 3 was used to generate the cover image.

Table of Contents

INTRODUCTION

The spring semester of 2023 in higher education was impacted by the release of Large Language Models (LLMs) to the general public. OpenAI's ChatGPT was at the forefront, with Google's Bard, Microsoft's Bing, and others following suit. In short, LLMs are a type of "Generative AI (GAI)," machine learning applications that train artificial neural networks on large amounts of input text. By assessing the probability of subsequent words in a sequence, these models can produce text that closely resembles human language, performing various tasks such as summarizing, writing, coding, and answering questions in response to input prompts.

In Norwegian higher education, the introduction of LLMs has spurred many discussions. In general, these discussions revolve around concerns, uncertainties, and potential opportunities. Early discussions in academia were dominated by increasing apprehension regarding students using ChatGPT to cheat on exams (Hystad & Fanghol, 2023). Furthermore, LLMs come with a wider range of challenges, including privacy matters, equity issues, biases inherent in the datasets, and the need for validating the accuracy of the text produced. LLMs can generate false claims and information, often termed "hallucinations". Such potential pitfalls need careful consideration when strategizing LLM integration in educational settings (Alkaissi & McFarlane, 2023). Nevertheless, these concerns must be weighed against the constructive possibilities this technology represents for higher education. These include for example, assisting educators and students in adapting to novel work dynamics, personalizing learning, reducing time-consuming tasks, improving accessibility and inclusivity, and providing multilingual support. As educators, we have the opportunity to leverage LLMs for various purposes, including summarizing content, conducting literature reviews, creating outlines, generating computer code for various programming languages and tasks, and providing tailored feedback that caters to individual learning needs.

Educating students about the workings of LLMs, bias in their training datasets, and potential misuses is essential to fostering an understanding of the complex challenges associated with implementing GAI in higher education. As educators, it is our role to guide students in understanding these tools and better prepare them to use LLMs constructively both during their studies and in their post-student lives. Moreover, there is increasing evidence indicating that students themselves are concerned about the ethical and appropriate use of these new tools. While students see value in using these tools to support their learning, they want to steer clear of any allegations or perceptions of cheating (Coelho et al., in review). In summary, LLMs have the potential to significantly influence higher education. It is the responsibility of educators to facilitate their use in constructive, discipline-relevant ways.

In response to the ongoing discussions in higher education, we conducted a survey to examine the use and understanding of LLMs among students and instructors at the Faculty of Mathematics and Natural Sciences at the University of Bergen, Norway. Given that the survey was distributed in February and March 2023 - shortly after ChatGPT was launched and before similar tools from other providers were in wide use - we focused primarily on ChatGPT. Our aim was to gain a deeper understanding of existing perceptions and misconceptions about ChatGPT, enabling us to design more informed implementation strategies for its integration into our courses. Further, by considering usage patterns and perceptions shortly after the launch of ChatGPT, we established a baseline for later comparisons as these tools become increasingly embedded in our routines.

METHODS

We conducted a cross-sectional survey among students and instructors at the Faculty of Mathematics and Natural Sciences at the University of Bergen (UiB) in Norway. The survey was distributed over a two-week period in late February and early March 2023. Students who consented to participate in the survey (n=178) were recruited from a large introductory biology course (n=124) and an introductory course in programming (n=54). The participating instructors (n=74) represented various disciplines within the Faculty of Mathematics and Natural Sciences (MN). Since not all participants responded to every question, the

sample sizes for each item differ. These per-item sample sizes are noted under each figure in the results section.

The data were collected digitally through SurveyXact by Rambøll (Rambøll, 2023). The collection process followed the General Data Protection Regulation (GDPR) and was registered with RETTE, UiB's data protection portal. All participants, both students and instructors, were informed of the survey's purpose, assured that their participation was voluntary, and guaranteed that all data would be anonymized.

The survey included both constrained-choice and open-ended questions developed by the authors to examine understandings, uses, and perceptions of LLMs within our sample. As large-scale examinations of this topic had just begun, employing validated scales was not feasible. However, the inclusion of open-ended questions allows us to triangulate the responses to our constrained-choice questions and obtain more nuanced responses. One example of a constrained-choice question was: *Do you think that, overall, ChatGPT will have a positive or negative impact on your education?* with the response options being: "positive/negative/unsure". An illustrative open-ended question was: *How do you think this tool should be used in higher education, in a way that is fair and supports your learning?* Additional examples of the questions posed can be found in the "Results and Discussion" of this report.

Frequency analyses of the quantitative data were performed in IBM SPSS Statistics 28 (IBM, 2023) and visualized with diagrams. The qualitative data were subjected to a range of analytical steps. First, two researchers independently created a set of categories based on 30 random responses. Second, after comparing these categories, a unified set was established. Third, to refine these collective categories even further, another 70 random responses (35 each from students and instructors) were independently examined and then compared. Fourth, the responses were coded independently and compared. Any discrepancies in coding were discussed and agreed upon. Our presentation of the results integrates these findings in a sequential mixed-methods approach (Warfa 2016), where comments from both students and instructors provide additional context to the patterns observed in the constrained-choice responses.

RESULTS AND DISCUSSION

Demographics

The student sample consisted of 63% females and 37% males with an average age of 22,5 years. A large majority were continuing-generation students: 49% reported that both parents, and 38% reported that one parent had graduated from higher education. First-generation students, i.e., students who do not have parents with a higher education degree, made up 13% of the total student sample. These distributions mirror the broader student body, which is dominated by women and continuing-generation students. The majority of instructors were aged between 41-50 years (28%) and 61-70 years (28%), followed by those aged 51-60 years (21%) and 31-40 years (17%), and 30 years or younger (7%). Regarding their teaching experience, the majority (57%) had been teaching for over 10 years, while 22% for 2-5 years, 16% had taught for 6-10 years, and 5% for less than 2 years.

Understandings

To grasp how students and instructors understand ChatGPT and its uses, we asked them to define the tool using their own words. We also encouraged elaboration, specifically related to what it does and what it can be used for. We analyzed the comments from students and instructors about ChatGPT with a comparative breakdown based on four themes: i) Descriptive; ii) General Impressions; iii) Exposure and Experience; iv)

Recommendations. Below are some examples of statements given by students and instructors. Note that the comments were given in either English or Norwegian, but we report them all in English.

Descriptive Themes

ChatGPT is an AI-based system

Student: “ChatGPT is a generative artificial intelligence built on GPT 3.5. It will respond as best it can to questions sent to it.”

Instructor: " ChatGPT is a language code that generates answers to questions in free text format, and it is trained on data from the entire internet, especially in English, until around autumn 2021."

ChatGPT is a language model

Student: " ChatGPT is an AI-based chat system. It is based on language models, and should be able to answer all possible questions, in all languages.”

Instructor: "ChatGPT is an artificial intelligence that is trained to generate text that is likely to match the question asked. It is a language model that considers and finds the most likely next word in a sentence. "

ChatGPT is trained on source data

Student: "ChatGPT is a program that has learned grammar etc by reading a large amount of texts and practicing predicting what is the next word in the sentence. It also has access to very large and widely different text databases, from social media to scientific texts."

Instructor: " It differs from previous models of this type in that it functions as a dialogue between the user and the AI. ChatGPT is trained on large amounts of text in several languages, but there are many biases in the selection of texts (both culture/nationality/politics). It can give useful answers to a number of questions, but also has a good imagination - or it hallucinates and gives answers it thinks the user wants."

ChatGPT has diverse functions

Student: “It can be used to generate text, codes, solve problems, and responds uncritically to most things.”

Instructor: "Yes, I've heard about ChatGPT and would describe it as a generative chat-bot trained on a vast database of written sources. It can be used to perform various tasks, from computations (processing and analysis of numerical and textual data) to linguistic operations (translations, descriptions/summaries, completion, including with programming languages) and covers a very large range of fields."

General Impressions

ChatGPT isn’t always accurate or reliable

Student: "ChatGPT does not know whether what it produces itself is true."

Instructor: " The text often looks reasonable, but if you know a lot about a topic, you will see that it often presents incorrect information as if it were correct, and gives no indication of uncertainty"

ChatGPT can be useful

Student: "In general, I see ChatGPT as a useful tool that can easily be used to collect and effectively present information on various topics, for example in text form, but also to generate code."

Instructor: "It is a writing tool that formulates texts based on questions or statements you make in the writing field. It collects information to answer the problem, but this should always be fact-checked. It is not a search engine, but a tool that can help you formulate and understand better."

ChatGPT can be used for cheating

Student: “I have heard that it can be used as a cheat, but also that it can be a useful tool when you need a little inspiration.”

Instructor: “But can also be used for cheating. Something that can be difficult to reveal.”

One instructor was especially concerned: “I have read quite a bit about this one. It is an AI that finds information on the internet. It is well documented that it makes up things that do not exist. It should absolutely be banned in all teaching. ChatGPT, and its further development, is going to lead to the downfall of our civilization.”

Exposure and Experience

Student: "Personally, I use ChatGPT as an aid for submissions or assignments, most often when I'm stuck or don't understand the assignment. When I ask a question related to my assignment, I get an answer. I never take the answer directly into the task but use the answer from ChatGPT as a starting point or a starter for how I can solve the task. I usually look up the information it gives me in order to obtain sources for the information or to gain a deeper understanding of what it provides."

Instructor: "I have used ChatGPT quite a bit for different tasks, including coding."

Recommendations

Student: “The text often looks reasonable, but if you know a lot about a topic, you will see that it often presents incorrect information as if it were correct and gives no indication of uncertainty. It should therefore not be used for important things.”

Instructor: "ChatGPT is a valuable tool that can enhance the learning experience. It can be employed to engage students, provide quick answers to questions, and foster critical thinking by encouraging students to explore topics in-depth. Incorporating ChatGPT into educational settings can be a powerful way to support learning and research."

Overall, both groups recognize ChatGPT's reliance on vast amounts of training data and its capabilities as a text generator (and consequently, the potential for diverse uses). However, they also express concerns about its reliability and the potential for misuse, especially in academic contexts.

Uses

Our examination of the frequency in use of ChatGPT, based on constrained-choice survey responses, shows differences between students and instructors. Whereas 116 (74%) of the students at the time of data collection had started using ChatGPT, only 31 (51%) of the instructors had. We also found that students use ChatGPT more frequently than instructors do (Figure 1). These findings corroborate those from the open-ended responses (discussed above).

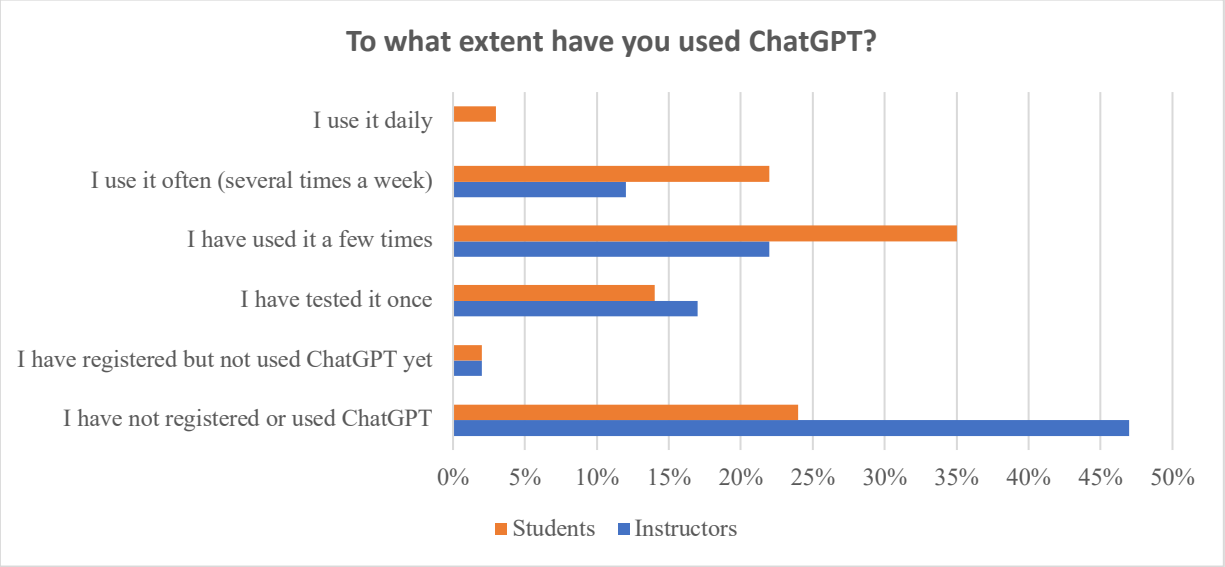

Figure 1. A comparison of the frequency in use among students (n=157) and instructors (n= 60)

Having established differences in the frequency in use among students and instructors, we moved on to examine different ways of using ChatGPT. Specifically, we asked the students and instructors to indicate tasks for which they use ChatGPT (Figure 2). Here we also identified differences between students and instructors, especially when it came to generating text and finding answers to questions related to schoolwork or teaching.

We find that three months after its release, the predominant use of ChatGPT was for finding answers related to schoolwork/studies, while for instructors, it was mostly employed for generating text.

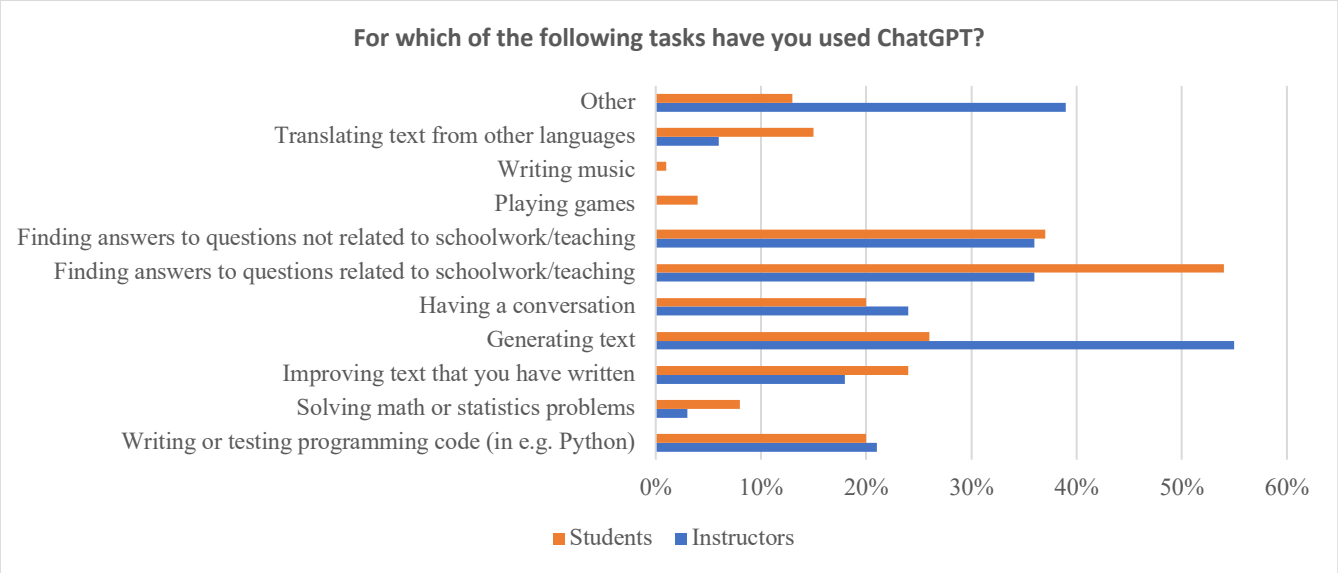

Figure 2. A comparison of the tasks for which students (n=125) and instructors (n=33) had used ChatGPT

To elaborate on the use of tasks related to schoolwork and studies, we posed two more questions for both students and instructors. For students, we asked *To what extent have you used ChatGPT for generating text for an assignment that was graded and NOT citing ChatGPT* and *To what extent have you used ChatGPT for generating text for an assignment that was graded and citing ChatGPT?*. For instructors, we asked *To what extent have you used ChatGPT for generating teaching material/text/presentations?* and *To what extent have you used ChatGPT for grading a student assignment?* The main reason for including these items was the ongoing discussion around student cheating. However, most (>80%) of our students report that they have not used ChatGPT for a graded assignment - at least not when asked in February/March 2023. Moreover, we wanted to examine if instructors use ChatGPT for grading and/or generating teaching material. Our results show that they do not use it for grading at all and that only three instructors (6% of those who had used the tool) had used it for generating teaching material.

Furthermore, we found that 80 (72%) of the student users have occasionally or often experienced an improvement in their ability to complete a task using ChatGPT (Figure 3). Identifying positive uses of LLMs such as ChatGPT and similar tools in higher education could help instructors appreciate the affordances—and not just the challenges—associated with the use of these tools (Cotton et al., 2023). That students experience an improvement in their own ability to complete a task using ChatGPT is something educators should keep in mind as they consider whether, and how, to incorporate LLMs into their teaching.

Perceptions

The perceptions examined in our survey largely include two types: the students' and the instructors' predictions and concerns about the use of LLMs like ChatGPT.

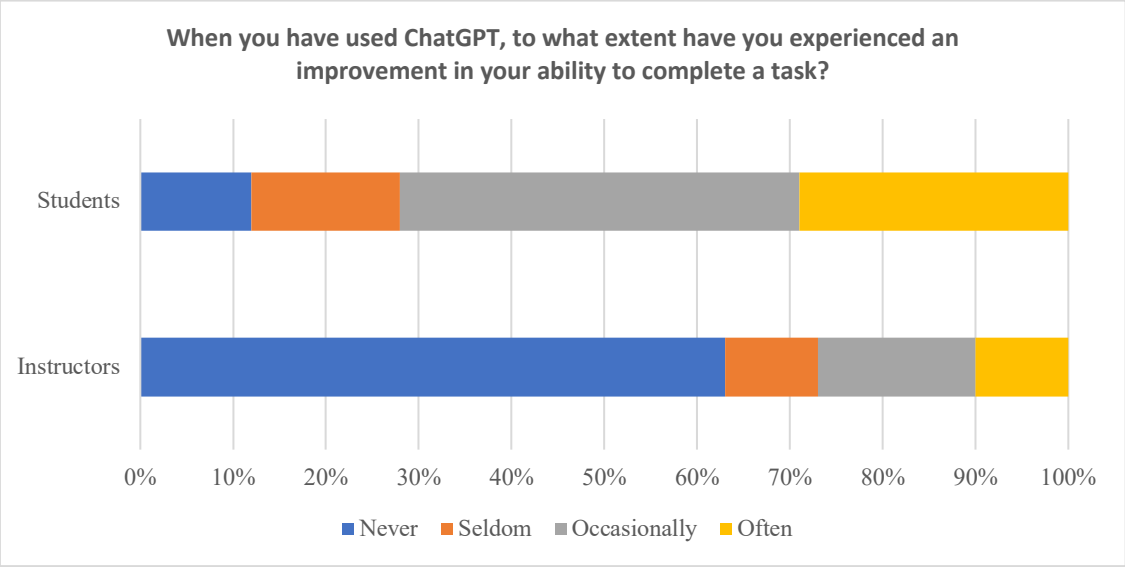

Figure 3. Students' (n=111) and instructors' (n=30) experienced improved ability to complete a task using ChatGPT

Predictions

When we asked the students and instructors about their perceptions on what impact LLMs will have on higher education in the future, we saw differences (Figure 4). The students were both more positive and

less unsure about the impact of LLMs on education than the instructors were. These differences may be associated with their different experiences with ChatGPT. Specifically, the students had used ChatGPT more often than the instructors, and almost half of the instructors had not yet used ChatGPT (Figure 1). Also, the students had used ChatGPT more than the instructors to find answers to questions related to schoolwork or teaching (Figure 2). Further, almost three quarters of students reported that they occasionally or often experienced an improvement in their ability to complete a task whereas a mere quarter of the instructors reported the same (Figure 3).

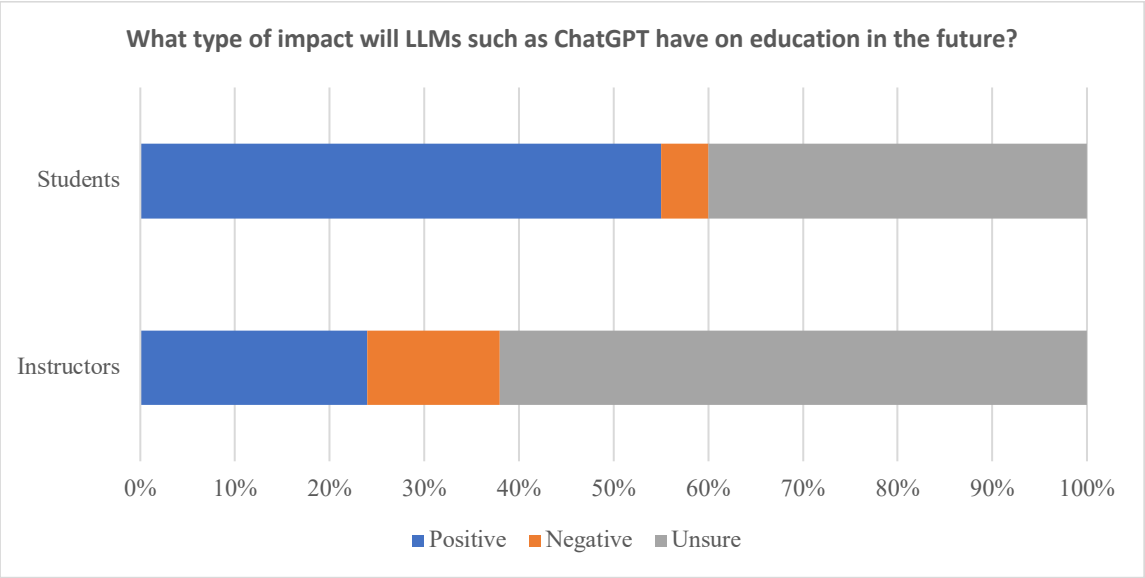

Figure 4. Students' (n=154) and instructors' (n=59) prediction of the impact of LLMs on future education

Concerns

We found that neither students nor instructors expressed a great deal of ethical concerns about the use of ChatGPT. Specifically, 39 (35%) of the students and 10 (33%) of the instructors reported that they occasionally or often had experienced ethical concerns when using ChatGPT (Figure 5).

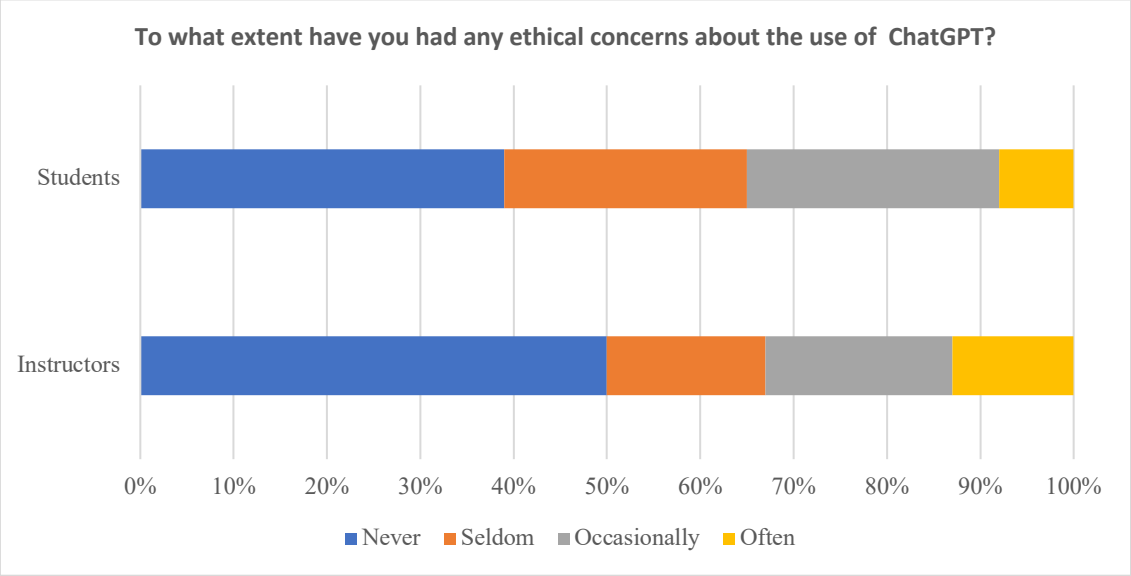

Figure 5. Students' (n=112) and instructors' (n=30) ethical concerns about the use of ChatGPT

This tendency was also observed when examining specific ethical concerns. For example, students and instructors did not differ much in their concerns about cheating, privacy, and bias (Figure 6). Further, a minority of students and instructors were occasionally or often concerned about cheating and privacy matters. In fact, around 40% of the student (n=40) and instructor users (n=13) had occasionally or often questioned whether ChatGPT is biased.

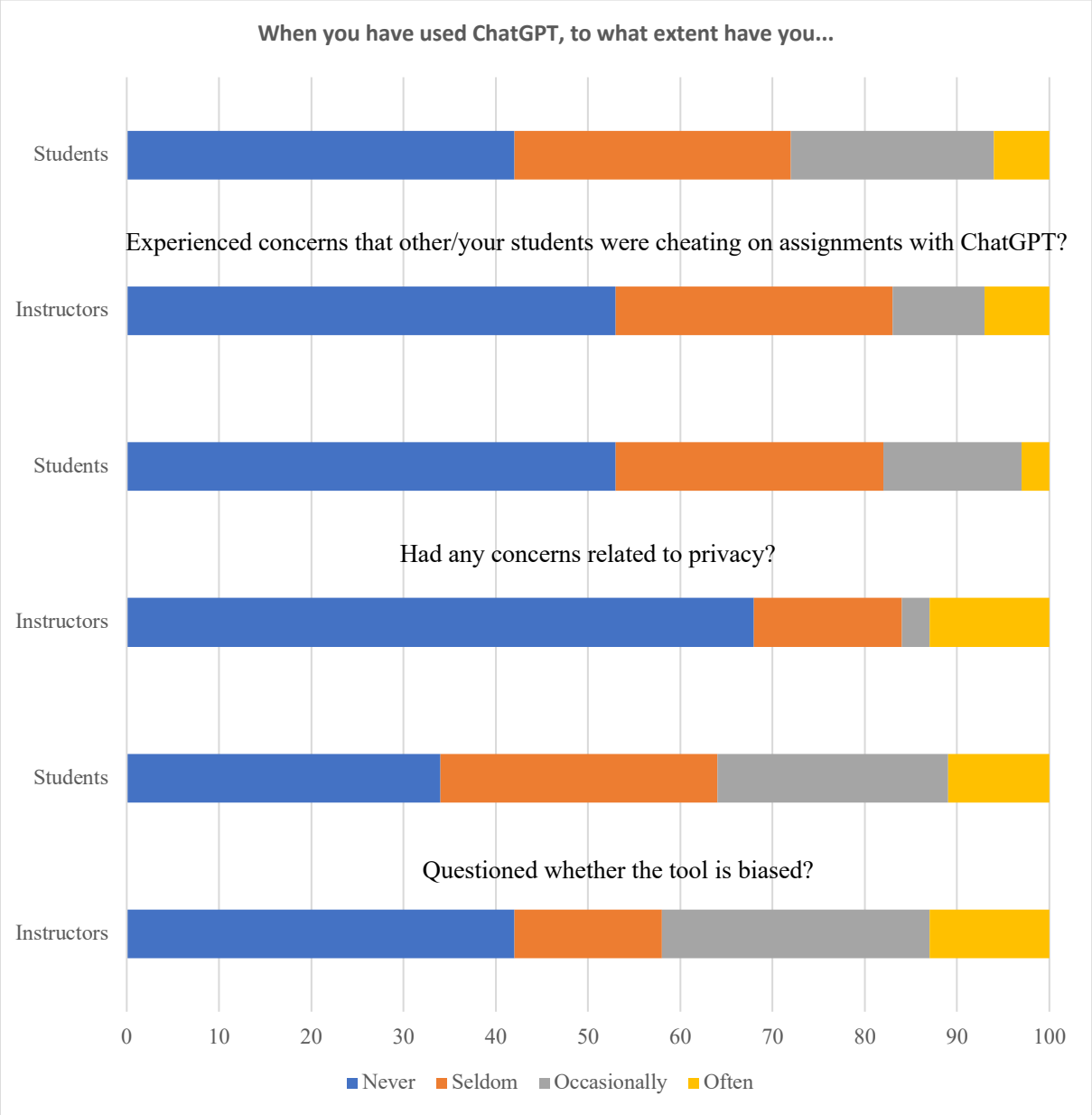

Figure 6. Students' (n=112-113) and instructors' (n=30-31) concerns about cheating, privacy, and bias when using ChatGPT

The students and instructors did differ in one concern, however, and that was in the question of accuracy (Figure 7). Here we found that 17 (55%) of the instructors but only 40 (35%) of the students using ChatGPT had often questioned the accuracy of the results. This finding is consistent with the indication that some students clearly see ChatGPT as akin to a search engine.

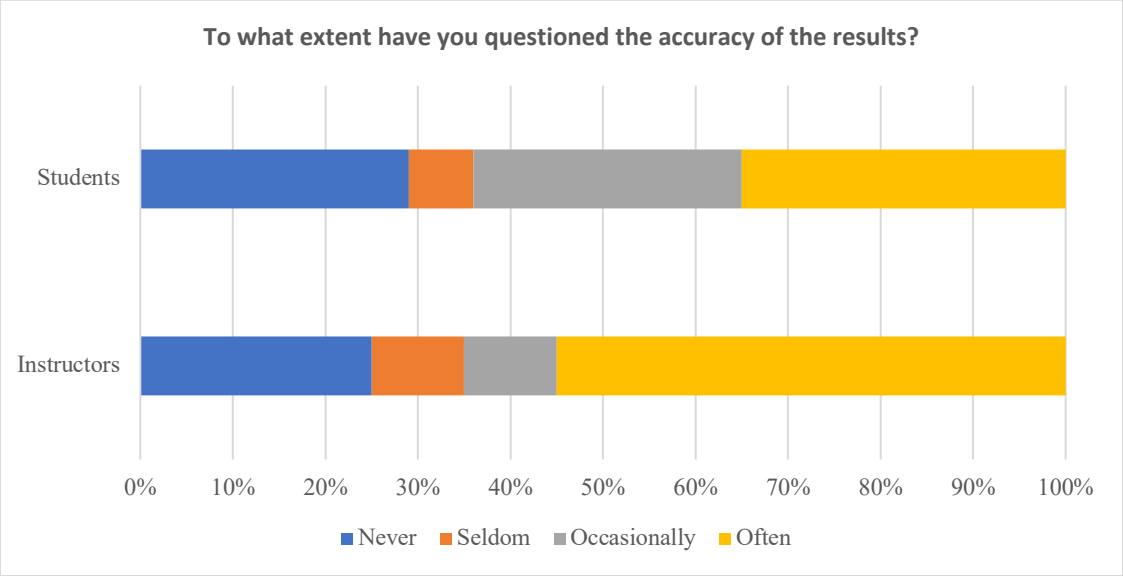

Figure 7. Students' (n=114) and instructors' (n=31) concern about accuracy when using ChatGPT

Suggestions for use in higher education

Many students and instructors responded to the question: *How do you think this tool should be used in higher education, in a way that is fair and supports your learning?* In *general* terms, several responses could be categorized as one of the following: ChatGPT...should be used in learning; is difficult to use fairly; should not be used during an exam; should be acknowledged; should not be used at all. Examples for each of these categories are provided in Table 1.

Table 1. General categories for respondents' views on how ChatGPT can be used in higher education (with examples from students).

Category	Example
Should be used in learning	I think it should be used to develop learning, and to understand things better. If it is used just to finish something quickly, and for cheating, it does not give a better understanding of the subject. But as I said, if it's being used for learning, lg understanding, I think it's really good
It is difficult to use fairly	If good guidelines are developed for how to use it, so that all students benefit equally from ChatGPT, then it can certainly be an okay substitute for teaching, but I think it can be difficult to find a way to use it that is fair.
Should not be used during an exam	It should, of course, be illegal on exams. Otherwise, it may be allowed. The responsibility lies with the user in not becoming too dependent on it.
Should acknowledge its use	Should always be cited as a source, and don't use it too much as you can't learn anything from it
Should not be used/doubt that it should be used	I think chatGPT can be an obstacle to deeper learning in some cases and can be used for "cheating", so I think we should have more school exams to make sure that all students have an equal opportunity to showcase knowledge. At the same time, I think it's a good tool to use for teaching activities.

More specific responses about *how* ChatGPT can be useful in education, included "should be used as an aid/tool/supplement" to "improve code" to "can be used to prepare for an exam" (Table 2).

Table 2. Specific categories for respondents' views on how ChatGPT can be used in higher education (with numbers, n, in each category and examples from either students, S, or instructors, I).

Category (n)	S/I	Example
Focus on proper use of the tool/ethics issues surrounding its use (49)	S	I think there should be talk about both the consequences and the benefits of it in each subject. Everyone is responsible for their own learning, so it's ultimately everyone's responsibility that you learn what you have to, and using chatGTP as a cheater out I don't see as an advantage. But I think everyone should know that it can be a good tool with proper use!
Should be used as an aid/tool/supplement (31)	S	It should be used as an aid, and a tool for new input and new ideas. Not as a tool that does the work for you.
Can help finding information/encyclopedia/search engine (30)	I	It will certainly be a very good "search engine" that can provide quick and good answers to a lot. This probably requires that the user has a good insight into what the answers are based on.
Make summaries/explain difficult topics (23)	I	Learn to use it to collect, summarize and analyze information, but no one should think that it replaces knowledge.
Spelling/improve language/correct text (18)	S	Works well as spelling aid.
Inspiration to start an assignment or text (11)	S	Compare afterwards with work you have done to improve and get inspiration, but do not use it to replace your own work.
Improve code (6)	S	I honestly don't know. ChatGPT can help learn in many ways, especially in things like programming, but the problem arises when people use it blindly. If you don't understand what you're doing, you don't learn anything and it often makes mistakes, so you can't trust it. I can imagine that in the future you will separate characters so that e.g. 50% of the character comes from what you can manage without chatGPT and the rest comes from what you manage with chatGPT. or other AI.

IMPLICATIONS

Our findings indicate that, even quite soon after the release of ChatGPT, students and instructors had a range of experiences with and perceptions of the tool and related emerging tools. Students and instructors had different views on ChatGPT. Students were mostly neutral or positive, seeing ChatGPT as a useful tool, and their suggestions for its utility often focus on specific tasks like programming, gathering information, and writing drafts of text. Instructors were more divided, with some acknowledging its benefits, but others worrying about its accuracy, ethics and cheating, and its effects on education and society. Overall, while students appeared to view ChatGPT more as a tool with potential benefits, instructors seem more concerned about its implications, especially in an academic setting.

These findings, combined with the ongoing discussion of LLMs in higher education in Norway suggest that it will be useful to consider, *together with our students*, the following elements in our future teaching: i) understanding what LLMs are and what they can do, ii) possibilities of LLMs to improve students' ability to complete tasks, iii) critical exploration of the accuracy of LLMs, and iv) alternative forms of assessment in an era of LLMs.

What are LLMs and what can they do?

Our findings point to several misconceptions about what LLMs are and what they can do. When engaging in discussions on LLMs and their influence on teaching and learning in higher education, it is important that we share a common and correct understanding of what these tools are and what they can do. Thus, a natural first step would be to seek reliable information on the nature of LLMs within the context of educational development in higher education. Next, we could invite our students to explore the topic through engaging them in discussions and solving, together, purpose specific LLM tasks. Setting aside time in our teaching activities to explore topics not directly linked to the curriculum is not always easy due to ambitious learning outcomes and little time. However, by putting LLMs on the agenda, ideally within the context of our disciplines, we may help develop a much-needed common understanding of and framework for the uses of LLMs .

Can students and instructors critically explore the accuracy of LLMs?

One of our findings indicates that only 35% of the participating students occasionally or often had questioned the accuracy of the results when using ChatGPT. Knowing what we know about factual inaccuracies with LLMs such as ChatGPT, it could be a concern that few students question their accuracy. It may be the instructor's role to make students aware of this problem. This could be done by having students work on generated text around a subject with which they have some familiarity. A critical review of generated output may reveal to students that some of the text is false or irrelevant to the initial query.

Can LLMs improve students' ability to complete tasks?

A majority of the students participating in our survey report that they have experienced an improvement in their ability to complete a task when using ChatGPT. We argue that this finding is both interesting and promising. If students and instructors are able to find ways to use LLMs in constructive ways to improve the students' ability to complete tasks, much may be gained. If simple tasks can be completed quickly, students can use their time for more creative tasks or tasks that lead to deeper learning.

In one approach, 4th semester students were asked to perform basic bioinformatic analyses based on assistance from ChatGPT. The instructor was there only to announce the different steps that should be taken: 1) download coding sequence for gene "YFG" from human; 2) download sequence from same gene in additional 4 species; 3) perform multiple sequence alignment analysis of the sequences; etc. Using ChatGPT as their "teaching assistant", most students could perform accurate sequence retrieval and analysis, protein translation and functional analysis, without prior knowledge of such analyses or the databases and online tools used.

Can LLMs lead to alternative forms of assessment?

One of the main concerns in the initial education-related discussions of LLMs is their potential use in student cheating, but this concern is not reflected in our local findings. In fact, when students (and instructors?) comment on the issue of cheating they focus more on the need to think differently about assessment: "Can we cheat using ChatGPT? Then maybe the assessment form needs to be adjusted?"

Repeatedly, research has shown that such high-stakes final-exams, also known as summative assessment, aren't good for overall motivation or learning, and they may increase existing inequities in education (Högberg and Horn 2022, Salehi et al., 2019). Despite this evidence the majority of higher education courses in Norway are exclusively assessed using summative assessment (Norwegian Ministry of Education and Research, 2017). Perhaps the introduction of LLMs is what we need to *finally* introduce alternative assessment forms – e.g., assessments that are authentic (i.e., related to a practical application of the course material) and ideally, formative (i.e., implemented throughout the semester and thus, assessment *for* learning rather than *of* learning; Harlap et al., 2022).

When deciding on alternative assessment forms we need to primarily consider two aspects: 1) Do the assessment forms assess the learning outcomes and how can we, through relevant learning activities, prepare our students for these assessments?; and 2) Do the assessment forms stimulate and lead to deep learning? Deep learning approaches focus on the meaning of the learning content and take place when our students apply, analyze, and reflect on the learning content in ways that enable them to generate new knowledge (Biggs, 1996). An example of an assessment form which may stimulate deep learning is to pose authentic problems or tasks which require the students to transfer the knowledge and skills they have acquired to new settings – and preferably settings that are authentic and prepare the students for life outside the classrooms (Harlap et al., 2022).

CONCLUSIONS AND NEXT STEPS

These findings are limited in scope and reflect a time, soon after the introduction of ChatGPT 3.5, we have moved beyond. As of this writing (June 2023), several new LLM tools are in use, and students are also becoming familiar with many other types of Generative AI. However, the take-away messages are relevant: students want LLMs to be used in education, they are worried about the potential negative outcome of LLM use, and they are probably looking to their instructors to provide guidance—specifically, how *should* we be using these tools in our disciplines?

It is beyond the scope of this work to say how these tools should and should not be used. However, we can reflect on these findings and urge our colleagues to use these tools, learn about them, talk about them with their students, and model ethical, conscientious use of LLMs.

Future work will explore the practicalities of LLM use in our courses, and we look forward to studying the specific implementation of LLM-anchored teaching strategies. We will also replicate some aspects of this study after students and instructors have had more time to learn about LLMs, giving us a better understanding of growing competences with these tools.

REFERENCES

Alkaissi, H., & McFarlane, S. I. (2023). Artificial Hallucinations in ChatGPT: Implications in Scientific Writing. Curēus (Palo Alto, CA), 15(2), e35179–e35179. https://doi.org/10.7759/cureus.35179

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, *32*(3), 347-364. https://doi.org/10.1007/BF00138871

Coelho, R., Cotnerb, S, Bjuneb, A., Ellingsen, S. Students' Perspectives on Large Language Models in Biology Higher Education: Insights for Collaborative Curriculum Design (under review).

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1–12. https://doi.org/10.1080/14703297.2023.2190148

Ellingsen, S., Bjune, A., & Cotner, S. (2023, April 28). Let´s not vilify our students over ChatGPT. *ScienceNorway.* https://sciencenorway.no/artificial-intelligence-ChatGPT-education/lets-not-vilify-our-students-over-ChatGPT/2190432

Harlap, Y., Jørgensen, C., & Cotner, S. (2022). Maintaining quality assessment practices in Norwegian higher education after the two-evaluator law. Nordic Journal of STEM Education, 6(1), 41–60. https://doi.org/10.5324/njsteme.v6i1.4873

Hystad, J., & Fanghol, T. A. (2023). NTNU forbyr bruk av ChatGPT på eksamen. — Nødvendig å nevne det eksplisitt. Khrono. Retrieved June 16 from https://khrono.no/ntnu-forbyr-bruk-av-ChatGPT-pa-eksamen-nodvendig-a-nevne-det-eksplisitt/779389

Högberg, B., & Horn, D. (2022). National High-Stakes Testing, Gender, and School Stress in Europe: A Difference-in-Differences Analysis. *European Sociological Review, 38*(6), 975–987. https://doi.org/10.1093/esr/jcac009

IBM. (2023). *IBM SPSS Statistics for Macintosh, Version 28.0.* IBM Corp.

Norwegian Ministry of Education and Research. (2017). Meld. St. 16 (2016–2017) - Quality Culture in Higher Education. https://www.regjeringen.no/no/dokumenter/meld.-st.-16-20162017/id2536007/

Møgelvang, A. (2023). *Cooperative Learning in Undergraduate STEM Education: Applications and Outcomes.* [Doctoral dissertation]. University of Bergen, Norway. https://hdl.handle.net/11250/3054858

Rambøll. (2021). *SurveyXact*. Retrieved 22.02.2023 from https://www.surveyxact.no

Salehi, S., Cotner, S., Azarin, S. M., Carlson, E. E., Driessen, M., Ferry, V. E., & Ballen, C. J. (2019). Gender performance gaps across different assessment methods and the underlying mechanisms: The case of incoming preparation and test anxiety. *Frontiers in Education (Lausanne), 4*. https://doi.org/10.3389/feduc.2019.00107

Warfa, A. R. M. (2016). Mixed-methods design in biology education research: Approach and uses. CBE—Life Sciences Education, 15(4), rm5.


UNIVERSITAS
BERGENSIS