Generative AI and Teachers - For Us or Against Us? A Case Study

Jenny Pettersson, Elias Hult, Tim Eriksson and Tosin Adewumi*
Machine Learning Group, EISLAB, Luleå University of Technology, Sweden
firstname.lastname@ltu.se

Abstract— **We present insightful results of a survey on the adoption of generative artificial intelligence (GenAI) by university teachers in their teaching activities. The transformation of education by GenAI, particularly large language models (LLMs), has been presenting both opportunities and challenges, including cheating by students. We prepared the online survey according to best practices and the questions were created by the authors, who have pedagogy experience. The survey contained 12 questions and a pilot study was first conducted. The survey was then sent to all teachers in multiple departments across different campuses of the university of interest in Sweden: Luleå University of Technology. The survey was available in both Swedish and English. The results show that 35 teachers (more than half) use GenAI out of 67 respondents. Preparation is the teaching activity with the most frequency that GenAI is used for and ChatGPT is the most commonly used GenAI. 59% say it has impacted their teaching, however, 55% say there should be legislation around the use of GenAI, especially as inaccuracies and cheating are the biggest concerns.**

I. INTRODUCTION

Recent advances in artificial intelligence (AI), especially generative artificial intelligence (GenAI), have caused a stir in the Education sector around the world [1], [2]. ChatGPT[1], the leading Large Language Model (LLM) by OpenAI, has both been beneficial and controversial. Some of the concerns about GenAI are the generation of deepfakes and the nature of these "unexplainable" models [13]. In spite of these concerns, many recognize the benefits inherent in these technologies [9], [13]. In this case study, we seek to understand how university teachers perceive GenAI and investigate the following research questions.

1) To what extent are university teachers open to adopting generative AI in their teaching and classrooms?
2) What is the correlation between the impact of GenAI on teachers' teaching activities and their encouragement of their students to use it?

There's increasing study of the impact of GenAI on students [9], [13]. It is equally important to study the impact on other stakeholders or teachers' teaching activities in Education [11]. United Nations Educational, Scientific and Cultural Organization (UNESCO) hopes GenAI will be a tool that benefits teachers, students, and researchers. Our main **contributions** include the following: (1) We show through

data that university teachers in this case study are open to adopting GenAI; (2) we demonstrate the correlation between the positive impact of GenAI on teachers to their willingness to encourage their students to adopt it; (3) we provide many qualitative examples of comments of teachers on the impact of GenAI, ways they encourage students, and their concerns.

The rest of this paper is organised as follows. The literature review is discussed in Section II. The method employed in this work is described in detail in Section III. The findings are discussed in Section IV. We conclude with closing remarks in Section V.

II. LITERATURE REVIEW

The subject of GenAI in teaching is gaining increasing attention. It's impact on pedagogy cannot be ignored. Recent LLMs, such as ChatGPT, Aurora-M [17], Large Language Model Meta AI (LLaMA)-2 [23] and a host of others, have compelling abilities to generate human-like content, based on their training with big data [13]. This has prompted UNESCO to publish the guidance for GenAI in education and research [13], which builds on their recommendation on the ethics of AI.

To gauge the awareness of educators and their adoption of GenAI, it is useful to conduct a survey, similarly to that done with students [8], [22]. The views of educators on how teaching should change was surveyed by [7]. They found that most believed GenAI will have a big impact on teaching and are in favour of more face-to-face relational learning, among other things. [2] discuss the unparalleled opportunities and challenges presented by GenAI in Education. They observe that academic integrity and plagiarism are some of the concerns teachers have of their use. Meanwhile, [5] mention formative assessment feedback as one of the many benefits of GenAI. They acknowledge inaccuracies and biases as some of the drawbacks of this technology.

Some best practices for designing the questionnaire for a survey were identified by [16]. As simple as it may sound, the ordering of questions is an important consideration [14], [15]. These best practices are essential to have quality data from the survey. [16] also show that it is important to design questions to avoid *acquiescence* bias, which is the endorsement of a statement, regardless of the content. Testing a survey in a pilot study usually improves the quality of the full survey, as emphasized by [4], [6], [18], [19]. Using either closed or open questions have their own benefits, as demonstrated by [10], [20], [21]. In our work, we combined both types of questions to get the best out of the survey.

III. METHOD

Our chosen method for evaluating the use of GenAI by teachers at LTU is through a survey. It was designed to gather information about their habitual use of GenAI in teaching. In creating the survey, a few factors were taken into consideration to keep the inquiry objective and to avoid phrasings that could potentially skew the answers of the participants, based on best practices [16] . The factors are given below but are not limited to them.

- Keep the questions simple and concise, so as not to produce off-topic answers [12], [3].
- Include a broad, exhaustive list of viewpoints [16]
- Avoid questions where people tend to agree or disagree with statements regardless of their actual feelings or beliefs [16]
- Avoid single or double negation questions [16].

To reach a larger target audience, the survey questionnaire was created in both Swedish and English using an online tool.[2] The Swedish translation was carried out by some of the authors of this work, who are native speakers. The translations in the Results and Appendix sections were machine-translated and vetted by the native speakers. The survey consists of 12 questions to cover a relatively broad range of concepts that are necessary to answer the research questions. For the purpose of the survey, a simplified definition was provided for GenAI: *Generative AI, such as ChatGPT or DALL-E, is a tool that can answer questions and create images and other media based on prompts from the user* [13]. A copy of the questionnaire is available online.[3]

The survey was anonymous and a pilot study involving 8 teachers was initially carried out to ascertain if the forms and questions needed any adjustments. This full study involved sending the online survey via a link to all the teachers in multiple departments in different geographical campuses of LTU. The following are the 12 questions in the survey and their answer options, as created by the authors based on their pedagogy experience and the factors mentioned earlier. *Q1* and *Q8* were multiple choice questions. For teachers who answer "*None*" to the first question ("*filter question*" [16]), they were directed to continue to question 6 onwards.

1) *Have you used any generative AI in any of your teaching activities (e.g. preparation, teaching, assessment, or none)?*
   i) Preparation ii) Teaching iii) Assessment iv) Research[4]
   v) Administration vi) None vii) Other
2) *Which ones? e.g GenAI, Ex. ChatGPT, DALL-E, Bing AI, Google Bard etc. Others*
3) *How often do you use the one you use most?* i) Once a month ii) Once a week iii) Twice or more a week iv) Less than once a month
4) *Do you think the use has impacted your teaching?* i) Yes ii) No iii) Not sure
5) *Briefly describe the impact on your teaching.*

6) *Do you think AI will replace teachers in your subject if the trend of AI development continues?* i) Yes ii) No iii) Not sure
7) *Do you think there should be legislation around the use of generative AI?* i) Yes ii) No iii) Ambivalent
8) *What are some of your ethical concerns about generative AI?* i) Gender bias ii) Racial bias iii) Inaccuracies iv) Cheating v) None vi) Other concerns
9) *Will you encourage any of your students to use generative AI (in an ethical manner)?* i) Yes ii) No
10) *If you answered "Yes" in the previous question, In what way? And if "No", please say why.*
11) *Your gender* i) Woman ii) Man iii) Non-binary iv) Prefer not to say
12) *Your Division and Department*

The following LTU departments were involved in filling the survey.

1) SRT: The Department of Computer Science, Electrical and Space Engineering, among other subjects, contains Pervasive and Mobile Computing, Digital Services and Systems, Computer Science, Signals and Systems, Robotics and AI, Space Technology, Cyber-Physical Systems and Machine Learning.
2) SBN: The Civil, Environmental and Natural Resources Engineering department, among other subjects, contains Urban Water Engineering, Architecture, Structural Engineering, Building Materials, Engineering Acoustics, Soil Mechanics, Ore Geology, Applied Geophysics, Applied Geochemistry, Chemical Technology, and Process Metallurgy
3) ETKS: The Department of Social Sciences, Technology and Arts, among other subjects, contains Industrial Marketing, Political Science, Human Work Sciences, Performing Arts, Musical Performance, Economics, and Design.
4) HLT: The Department of Health, Education and Technology, among other subjects, contains Occupational Therapy, English and Education, Physiotherapy, Biomedical Engineering, Medical Science, Nursing, and Psychology.
5) TVM: The Department of Engineering Sciences and Mathematics, among other education, contains Mechanical Engineering, Automotive Engineering, Sustainable Energy, Electrical Power, Engineering Physics, and Mechanical Engineering.

IV. RESULTS AND DISCUSSION

The survey took 3:53 minutes for each teacher to complete on average. From Table I, 32 (48%) of the teachers do not use GenAI in any of their teaching activities. The remaining 35 (52%) use GenAI for one or more teaching activities, where *Preparation* is the most frequent activity, being 27%. Besides the identified teaching activities, one teacher explained that she uses GenAI "*To see if student work is AI generated*". The wordcloud of Figure 1 shows that ChatGPT has the lion share of usage with 52% of frequency of mentions (26 out of 50), in answer to *Q2*. The nearest is DALL-E, with 6%.

Question	Option	Total	Gender			Department				
			M	W	Not say	SRT	SBN	ETKS	HLT	TVM
Q1	Preparation	27	15	10	2	3.6	3.6	14.4	3.6	1.8
	Teaching	14	10	2	2	6	2	6	0	0
	Assessment	3	3	0	0	0	0	0	0	3
	Research	20	9	10	1	7.5	5	7.5	0	0
	Administration	3	3	0	0	3	0	0	0	0
	None	32	17	13	2	9.41	7.53	9.41	5.65	0
	Other	1	0	1	0	0	1	0	0	0
Q3	Once a month	18.42	13.16	5.26	0	3.07	6.14	6.14	3.07	0
	Once a week	18.42	10.53	7.89	0	9.21	0	9.21	0	0
	> once a week	31.58	15.79	7.89	7.89	21.05	0	5.26	0	5.26
	< once a month	31.58	15.79	15.79	0	3.95	11.84	11.84	3.95	0
Q4	Yes	58.97	41.03	12.82	5.13	19.66	14.74	19.66	0	4.91
	No	17.95	10.26	7.69	0	0	8.97	4.49	4.49	0
	Not sure	23.08	5.13	15.38	2.56	11.54	0	7.69	3.85	0
Q6	Yes	4.48	4.48	0	0	4.48	0	0	0	0
	No	85.07	41.79	35.82	7.46	29.10	15.67	29.10	11.19	0
	Not sure	10.45	8.96	1.49	0	3.48	3.48	1.74	0	1.74
Q7	Yes	55.22	29.85	19.40	5.97	23.36	12.74	14.87	2.12	2.12
	No	17.91	11.94	4.48	1.49	5.12	5.12	7.67	0	0
	Ambivalent	26.87	13.43	13.43	0	6.72	2.24	6.72	11.19	0
Q8	Gender bias	15.34	7.98	5.52	1.84	5.41	1.80	6.32	0.90	0.90
	Racial bias	15.34	7.36	6.13	1.84	5.96	1.70	5.96	0.85	0.85
	Inaccuracies	34.36	18.40	13.50	2.45	11.75	7.23	10.85	3.62	0.90
	Cheating	26.99	15.34	10.43	1.23	8.70	4.35	7.84	5.22	0.87
	None	1.23	1.23	0	0	1.23	0	0	0	0
	Other	6.75	2.45	3.68	0.61	1.12	2.25	3.37	0	0
Q9	Yes	76.12	44.78	26.87	4.48	30.45	17.40	23.92	2.17	2.17
	No	23.88	10.45	10.45	2.99	4.34	2.17	6.51	10.85	0
Q11	Woman	37.31								
	Man	55.22								
	Non-binary	0								
	Prefer not say	7.46								

Gemini, Stable Diffusion Web, and Midjourney come in next at 4% while all the other GenAIs have 2% (only 1 mention). Figure 2 shows the distribution of activities across gender.

Of the 52% who use these tools, there are as many teachers who use them twice or more in a week (i.e. more than once) as there are those who use them less than once a month (31.58%). About 59% say GenAI has impacted their teaching. Indeed, we see a strong positive correlation between the *Yes* choices of the departments to *Q4* (*Do you think the use has impacted your teaching?*) and *Q9* (*Will you encourage any of your students to use generative AI*?), based on the Spearman's correlation coefficient (ρ) +0.9474,[5] for $p = 0.01438$ (2-tailed). This implies the association between the two variables can be considered statistically significant. Several examples of the ways teachers feel GenAI has impacted their teaching are given in the appendix.

Most (85%) of the 67 teachers do not think AI will replace teachers. Many (55%) say there should be legislation around the use of GenAI, possibly because of the ethical concerns

[5]r = 1 and -1 are perfect positive and negative correlations, respectively
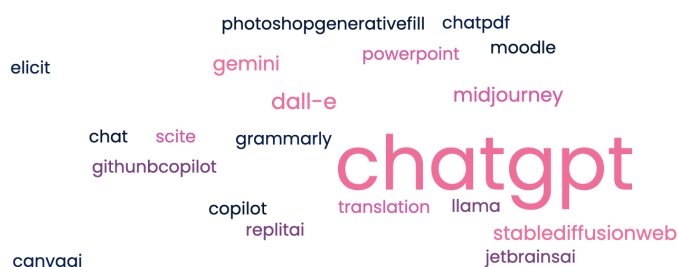

Fig. 1. WordCloud of GenAI

observed. Inaccuracies (34.36%) and cheating (26.99%) are the two most common concerns teachers have. Figure 3 shows the concerns across the departments. Given that the majority of teachers encourage their students to use GenAI, some of the ways they go about it are listed in the appendix, including these:

- "*...I do 1 or 2 sessions on how and why they should use it. I also show them when it can give wrong results and how to fact check it...*"
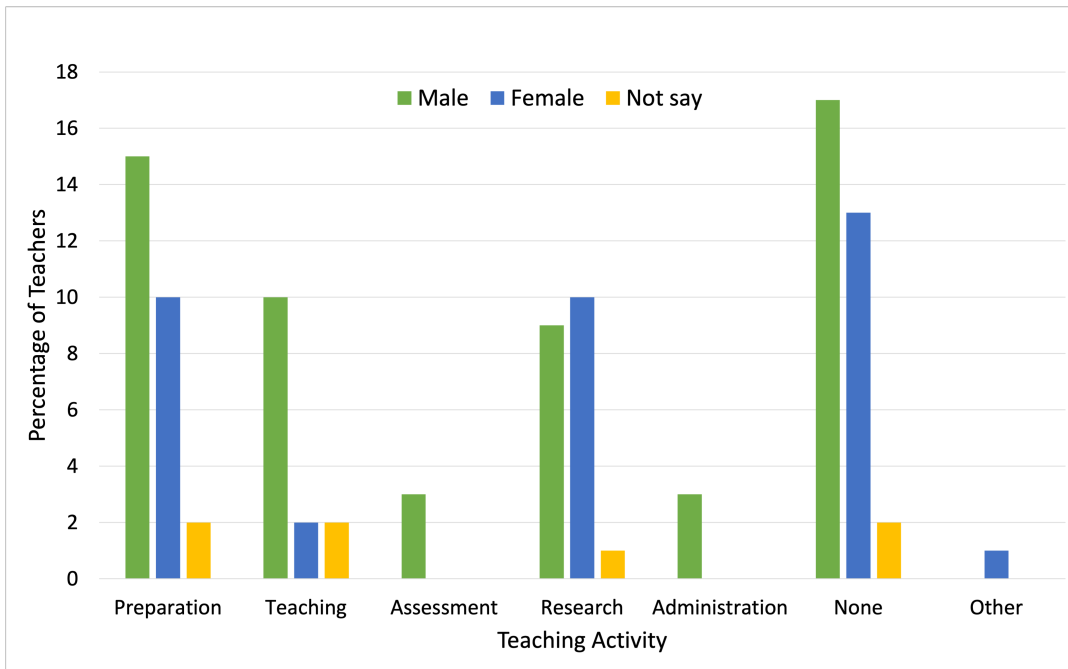
Fig. 2. GenAI usage activities across gender out of a total percentage of 100%.
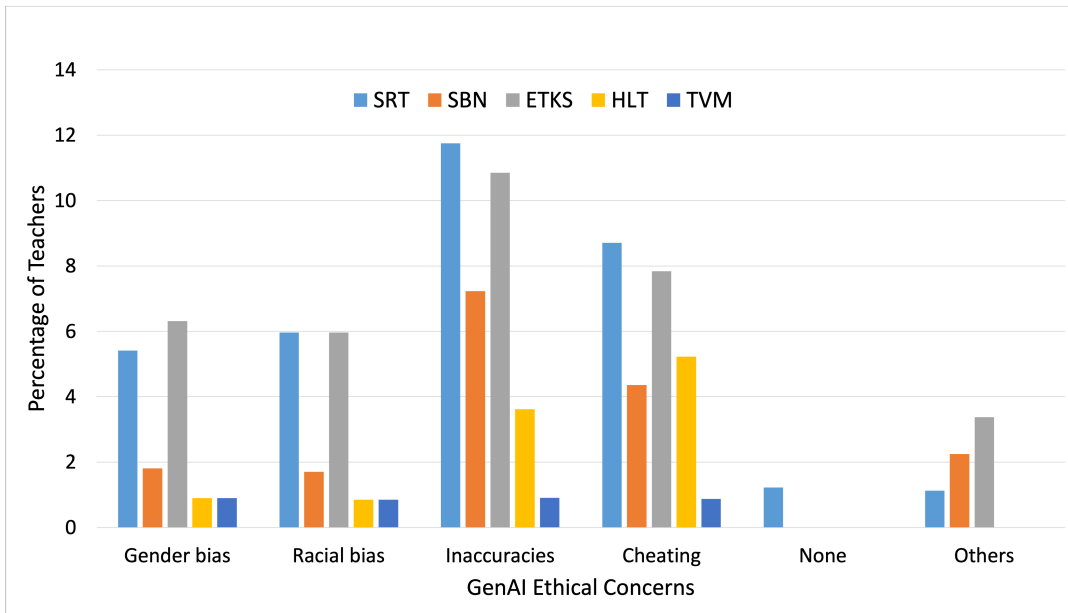

Fig. 3. GenAI concerns across departments out of a total percentage of 100%.

- "*seed texts, help in checking texts and results; use AI as study buddy*"
- "*Det är ett mycket effektivt hjälpmedel och bör uppmuntras. Är man orolig för fusk så examinerar man studenterna på fel sätt. (It is a very effective aid and should be encouraged. If you are worried about cheating, you are examining the students in the wrong way.)*"

Overall, 37.31%, 55.22%, and 7.46% of the teachers who completed the survey were women, men and those who

preferred not to say.

V. CONCLUSION

We have shown in this case study that teachers are open to adopting GenAI, as over 50% currently use it. We also observe a strong positive correlation between the positive impact of GenAI on their teaching activities and their willingness to encourage their students to adopt GenAI. We agree with the comments of some of the teachers that students "*are guaranteed to use it already*", therefore we believe teachers should be knowledgeable about these tools

in order to provide the appropriate guidance for students. Future work can investigate some of the ways of addressing the concerns of teachers expressed in this study.

ACKNOWLEDGMENT

APPENDIX

Examples by teachers of impact on teaching

1) I am able to provide more bang for the bucks and student feedback shows the impact.
2) I made my subject knowledge deeper.
3) ChatGPT hjälper mig att sammanfatta innehållet i t ex en workshop eller en föreläsning (*ChatGPT helps me summarize the content of, for example, a workshop or a lecture*)
4) t.ex. bättre bilder till mitt undervisningsmaterial, hjälp med bra översättning till engelska etc. (*for example better pictures for my teaching material, help with a good translation into English, etc.*)
5) preparation of slides and text material is more efficient and result more impactful.
6) Bättre språk, fler exempel (*Better language, more examples*)
7) I show its use for students to use properly as a tool, as well as when or how not to use it.
8) improved clarity

9) Jag får ett bollplank som kan hjälpa mig som lärare. (*I get a sounding board that can help me as a teacher.*)
10) Jag har använt Canva:s generativ AI för att snabbt ta fram illustrationer till mina powerpoints. Jag tror att det kan ha en viss positiv påverkan för inlärning att få "bildstöd" till anteckningarna. (*I have used Canva's generative AI to quickly produce illustrations for my powerpoints. I think having "visual support" for the notes can have some positive impact on learning.*)
11) Tidsbesparande (*Timesaving*)
12) Ibland dyker det upp aspekter som jag tidigare inte tänkt på, men som är relevanta. (*Sometimes aspects appear that I previously did not think about, but which are relevant.*)
13) Är ett fantastiskt verktyg att skapa bilder istället för att leta clip-art. Att få hjälp att förklara saker samt som kreativt verktyg i idegenereringsprocessen (*Is a great tool to create images instead of looking for clip-art. To get help explaining things and as a creative tool in the idea generation process*)

14) Jag har fått en bättre förståelse för hur studenter kan använda det som stöd samt jag har lärt mig att känna igen resultaten i studenters arbete (*I have gained a better understanding of how students can use it as support and I have learned to recognize the results of students' work*)

Examples of other concerns

1) I have had trouble with students copying AI generated information. I am afraid they are not using it as a learning tool, but rather to avoid learning.
2) ...students' usage affects the type of tasks I can give them and how I test their knowledge. That's mainly why I use it myself.
3) Nya former för examination kräva (*New forms of examination require*)
4) Efter att tidigare ha använt hemtentor i delar av kursen har jag gått över till salskrivningar. (*Having previously used take-home exams in parts of the course, I have switched to classroom writing.*)
5) Jag har tydligt kunnat visa för mina studenter varför det är viktigt att kunna ha grundläggande kunskap inom ett område, för det AI säger behöver inte nödvändigtvis vara korrekt, vilket de fick erfara i en kurs. (*I have been able to clearly show my students why it is important to be able to have basic knowledge in a field, because what the AI says does not necessarily have to be correct, as they experienced in a course.*)
6) I sin nuvarande form är generativ AI bra på att generera text som ser rimlig ut men mycket väl kan vara ful av felaktigheter. Jag ser inte detta som särskilt anändbart för mina studenter. (*In its current form, generative AI is good at generating text that looks reasonable but may well be ugly with inaccuracies. I don't see this as particularly relevant to my students.*)

Examples of ways teachers encourage their students to use GenAI

1) "Use it to learn, not to cheat". Använd för att förbättra eget material, inte för att generera från grunden. Viktigt att man inte presenterar andras material som sitt eget. Däremot är det liten skillnad att få en språkfgransking av en människa eller från AI när man väl skrivit texten. Viktigt att kunna materialet så att man kan faktagranska AI-lösningarna. (*"Use it to learn, not to cheat". Use to improve your own material, not to generate from scratch. It is important not to present other people's material as your own. However, there is little difference in getting a language check by a human or from AI once you have written the text. It is important to know the material so that you can fact-check the AI solutions.*)
2) Jag uppmuntrar dem att använda AI så mycket som möjligt om det hjälper deras lärande. (*I encourage them to use AI as much as possible if it helps their learning.*)
3) T.ex. för att bolla idéer, få hjälp med struktur i en text, hitta och sortera källor. Jag tycker det är ett väldig kraftfylld verktyg men precis som vilken verktyg som helst kan den vara farlig om den används av människor

utan rätt kunskap. Så att skaffa sig just den kunskapen för att kunna använda generativ AI på ett säkert sätt bor ingår undervisningen. (*For example. to brainstorm ideas, get help with structure in a text, find and sort sources. I think it's a very powerful tool but like any tool it can be dangerous if used by people without the right knowledge. So acquiring that particular knowledge to be able to use generative AI in a safe way is part of the teaching.*)

4) De använder det garanterat redan, så bättre att ha riktlinjer kring hur användandet bör ske. (*They are guaranteed to use it already, so better to have guidelines about how the use should take place.*)

5) Inte uppmuntra, men inte heller hindra (*Not encouraging, but not hindering either*)

6) Framförallt för att lära sig skriva vetenskaplig text på engelska och för att diskutera kursinnehåll, mjukvarukunskaper (t.ex. Excel, Matlab, python etc.) (*Mainly to learn how to write scientific text in English and to discuss course content, software skills (e.g. Excel, Matlab, python etc.)*)

7) Studenterna (Och industrin) använder redan generativ AI i långt högre utsträckning än vad vi lärare gör. Bättre att lära från dem och uppmana dem till att använda systemen på ett klokt sätt som uppmuntrar deras lärande. (*Students (And industry) already use generative AI to a far greater extent than we teachers do. Better to learn from them and encourage them to use the systems wisely which encourages their learning.*)

8) Precis som jag skrev innan så har jag uppmuntrat studenterna att använda ChatGPT för att få förståelse/fördjupning av vissa ämnen (*Just as I wrote before, I have encouraged the students to use ChatGPT to gain understanding/deepening of certain topics*)

9) Leta material, sortera i material (*Find materials, sort in materials*)

10) ja, att använda det i den kreativa processen för att utforska en mängd idéer, för att förbättra engelskan i texter etc. (*yes, to use it in the creative process to explore a variety of ideas, to improve English in texts, etc.*)

11) Som ett skrivstöd (*As a writing aid*)

12) För språkgranskning och programmering (*For language review and programming*)

13) AI kan vara mycket hjälpfullt om man har ett koncept/begrepp som man inte förstår men vill ha förklarat så att man sen, självständigt, kan använda konceptet/begreppet. (*AI can be very helpful if you have a concept that you don't understand but want explained so that you can then, independently, use the concept.*)

14) Utmärkt för sammanfattningar och scanning av stora litteraturmängder (*Excellent for summaries and scanning large volumes of literature*)

15) utkast till texter, utkast till musik, tex. (*drafts of texts, drafts of music, e.g.*)

16) Det är ett ypperligt bollplank, framförallt då man inte har någon fysisk person att diskutera med, men det kan även ge förstärkning om man använder den då man är studerar i grupp. (*It is an excellent sounding board, especially when you have no physical person to discuss with, but it can also provide reinforcement if you use it when you are studying in a group.*)

17) I produktutveckling, som ett stöd och verktyg. (*In product development, as a support and tool.*)

18) skriv hjälp, hjälp att komma igång med ett arbete, skapa bilder för presentationer, hjälp mot skrivkramp (*write help, help to get started with a work, create images for presentations, help against writing cramp*)

19) 1. Använda det för att stava rätt i inlämningar. 2. För att ha någon att bolla idéer kring studentarbete. 3. Fråga om grundläggande koncept som AI kan behärska och förklara. (*1. Use it to spell correctly in submissions. 2. To have someone to bounce ideas off of student work. 3. Ask about basic concepts that AI can master and explain.*)

20) Vi lärare behöver skapa och formulera premisser som är rimliga för detta, nu när man inte kan backa bandet med AI. Hur - vet jag inte än. (*We teachers need to create and formulate premises that are reasonable for this, now that you cannot reverse the trend with AI. How - I don't know yet.*)

REFERENCES

[1] ADEWUMI, T., ALKHALED, L., BUCK, C., HERNANDEZ, S., BRILIOTH, S., KEKUNG, M., RAGIMOV, Y., AND BARNEY, E. Procot: Stimulating critical thinking and writing of students through engagement with large language models (llms). *arXiv preprint arXiv:2312.09801* (2023).

[2] ALASADI, E. A., AND BAIZ, C. R. Generative ai in education and research: Opportunities, concerns, and solutions. *Journal of Chemical Education 100*, 8 (2023), 2965–2971.

[3] ARTINO JR, A. R., AND GEHLBACH, H. Am last page: Avoiding four visual-design pitfalls in survey development. *Academic Medicine 87*, 10 (2012), 1452.

[4] BABURAJAN, V., E SILVA, J. D. A., AND PEREIRA, F. C. Openended versus closed-ended responses: A comparison study using topic modeling and factor analysis. *IEEE transactions on intelligent transportation systems 22*, 4 (2020), 2123–2132.

[5] BAIDOO-ANU, D., AND ANSAH, L. O. Education in the era of generative artificial intelligence (ai): Understanding the potential benefits of chatgpt in promoting teaching and learning. *Journal of AI 7*, 1 (2023), 52–62.

[6] BOWDEN, A., FOX-RUSHBY, J., NYANDIEKA, L., AND WANJAU, J. Methods for pre-testing and piloting survey questions: illustrations from the kenqol survey of health-related quality of life. *Health policy and planning 17*, 3 (2002), 322–330.

[7] BOWER, M., TORRINGTON, J., LAI, J. W., PETOCZ, P., AND ALFANO, M. How should we change teaching and assessment in response to increasingly powerful generative artificial intelligence? outcomes of the chatgpt teacher survey. *Education and Information Technologies* (2024), 1–37.

[8] CHAN, C. K. Y., AND HU, W. Students' voices on generative ai: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education 20*, 1 (2023), 43.

[9] CHAN, C. K. Y., AND LEE, K. K. The ai generation gap: Are gen z students more interested in adopting generative ai such as chatgpt in teaching and learning than their gen x and millennial generation teachers? *Smart Learning Environments 10*, 1 (2023), 60.

[10] CONNOR DESAI, S., AND REIMERS, S. Comparing the use of open and closed questions for web-based measures of the continued-influence effect. *Behavior research methods 51* (2019), 1426–1440.

[11] DAI, Y., LIU, A., AND LIM, C. P. Reconceptualizing chatgpt and generative ai as a student-driven innovation in higher education. *Procedia CIRP 119* (2023), 84–90.

[12] GEHLBACH, H., AND ARTINO JR, A. R. The survey checklist (manifesto). *Academic Medicine 93*, 3 (2018), 360–366.

[13] HOLMES, W., MIAO, F., ET AL. *Guidance for generative AI in education and research*. UNESCO Publishing, 2023.

[14] JENSEN, P. S., WATANABE, H. K., AND RICHTERS, J. E. Who's up first? testing for order effects in structured interviews using a counterbalanced experimental design. *Journal of Abnormal Child Psychology 27* (1999), 439–445.

[15] KREUTER, F., MCCULLOCH, S., PRESSER, S., AND TOURANGEAU, R. The effects of asking filter questions in interleafed versus grouped format. *Sociological Methods & Research 40*, 1 (2011), 88–104.

[16] KROSNICK, J. A. Questionnaire design. *The Palgrave handbook of survey research* (2018), 439–455.

[17] NAKAMURA, T., MISHRA, M., TEDESCHI, S., CHAI, Y., STILLER-MAN, J. T., FRIEDRICH, F., YADAV, P., LAUD, T., CHIEN, V. M., ZHUO, T. Y., ET AL. Aurora-m: The first open source multilingual language model red-teamed according to the us executive order. *arXiv preprint arXiv:2404.00399* (2024).

[18] OKSENBERG, L., AND KALTON, G. New strategies for pretesting survey questions. *Journal of official statistics 7*, 3 (1991), 349.

[19] PRESSER, S., COUPER, M. P., LESSLER, J. T., MARTIN, E., MARTIN, J., ROTHGEB, J. M., AND SINGER, E. Methods for testing and evaluating survey questions. *Methods for testing and evaluating survey questionnaires* (2004), 1–22.

[20] REJA, U., MANFREDA, K. L., HLEBEC, V., AND VEHOVAR, V. Open-ended vs. close-ended questions in web questionnaires. *Developments in applied statistics 19*, 1 (2003), 159–177.

[21] SCHUMAN, H., AND SCOTT, J. Problems in the use of survey questions to measure public opinion. *Science 236*, 4804 (1987), 957–959.

[22] SMOLANSKY, A., CRAM, A., RADUESCU, C., ZEIVOTS, S., HUBER, E., AND KIZILCEC, R. F. Educator and student perspectives on the impact of generative ai on assessments in higher education. In *Proceedings of the tenth ACM conference on Learning@ Scale* (2023), pp. 378–382.

[23] TOUVRON, H., MARTIN, L., STONE, K., ALBERT, P., ALMAHAIRI, A., BABAEI, Y., BASHLYKOV, N., BATRA, S., BHARGAVA, P., BHOSALE, S., ET AL. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288* (2023).