# Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study



Citation for the original published paper (version of record):
Stöhr, C., Ou, W., Malmström, H. (2024). Perceptions and usage of AI chatbots among students in higher education across genders,
academic levels and fields of study. Computers and Education: Artificial Intelligence, 7.
http://dx.doi.org/10.1016/j.caeai.2024.100259



(article starts on next page)

Computers and Education: Artificial Intelligence 7 (2024) 100259

Contents lists available at ScienceDirect

Computers and Education: Artificial Intelligence

journal homepage: www.sciencedirect.com/journal/computers-and-education-artificial-intelligence

# Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study

Christian Stöhr *, Amy Wanyu Ou, Hans Malmström

*Department of Communication and Learning in Science (CLS), Chalmers University of Technology, SE-41296, Gothenburg, Sweden*

ARTICLE INFO

*Keywords:*
Chatbots
Artificial intelligence
ChatGPT
Student attitudes
Higher education

ABSTRACT

AI chatbots have ignited discussions and controversies about their impact on teaching and learning practices in higher education. This study explores students' adoption and perceptions of ChatGPT and other AI chatbots in higher education. Based on survey data from a large sample (n = 5894) across Swedish universities, the study employs descriptive statistical methods to analyze usage, attitudes, and concerns, and inferential statistics to identify relations between attitudes and usage and background variables (gender, academic level, and field of study). The results reveal broad awareness and use of ChatGPT among students, but not of other AI chatbots. More than half of the students expressed positive attitudes towards the use of chatbots in education, but almost as many expressed concerns about future use. Statistically significant differences were found across all examined grouping variables, particularly between genders and fields of study. Female students and students from the humanities and medicine consistently expressed more negative attitudes and concerns about AI's role in learning and assessment, while males and technology and engineering students showed higher usage and optimism. These findings not only validate the continued relevance of student backgrounds as a determinant of technology adoption but also expose several challenges and considerations surrounding AI and chatbot usage in education. The study supports the development of local solutions to AI in education tailored to student attributes and needs, and it provides insights for developers, educators, and policymakers.

## 1. Introduction

Artificial intelligence (AI) plays an important role in higher education and is profoundly influencing the academic and everyday lives of students (Chen et al., 2020). Foremost in developed countries and China, AI applications are increasingly implemented in education with significant potential to impact teaching and learning across all levels (Tahiru, 2021). Examples include AI-based adaptive learning platforms that provide personalized learning experiences to students, automated assessment, or AI-powered writing tools that enhance students' writing quality by providing real-time feedback on issues with grammar, punctuation, and style (Holmes & Tuomi, 2022; Zawacki-Richter et al., 2019). Concomitantly, AI in education (AIED) has emerged as a vibrant academic field, extending the capability of AI not just to learners but also to educators and institutions alike (Chen et al., 2020; Hwang et al., 2020; Nemorin et al., 2023).

Among current AI innovations, chatbots – conversational agents simulating human dialogue through natural language processing and machine learning algorithms – have gained particular traction through the launch of OpenAI's ChatGPT in November 2022 (OpenAI, 2022). This represents a paradigm shift in the domain; with over 100 million users within the first two months, the chatbot has become the fastest-growing application worldwide (Dennean et al., 2023), forcing other major tech companies to push their AI development programs. ChatGPT is based on the large language model Generative Pre-trained Transformer (GPT), which is trained on massive collections of data in the form of books, articles and openly accessible webpages. In contrast to most previous chatbots, ChatGPT has not only impressed with the quality of its responses, but also showcased its unique capability to "remember" a certain number of previous interactions within the same conversation. The versatility of ChatGPT as a conversational AI extends well beyond specific, single-purpose applications, and its potentials to increase efficiency, accuracy and cost savings were quickly highlighted (Deng & Lin, 2022). This seemingly opens up a multitude of opportunities and new challenges, and the potentially disruptive impact of generative AI is subject to intensive discussion in academia and the public (Li et al., 2023; Lo, 2023).

* Corresponding author.
*E-mail addresses:* christian.stohr@chalmers.se (C. Stöhr), amywa@chalmers.se (A.W. Ou), mahans@chalmers.se (H. Malmström).

https://doi.org/10.1016/j.caeai.2024.100259
Received 28 August 2023; Received in revised form 22 April 2024; Accepted 23 June 2024
Available online 24 June 2024

### 1.1. Advantages of AI chatbots in education fields

In education, ChatGPT's user-friendly and intuitive interface potentially reduces barriers to its wide adoption across educational settings and for different groups of teachers and learners (Kasneci et al., 2023), thereby overcoming many of the obstacles reported in the AIED literature. ChatGPT and similar AI applications can serve as self-study tools (Nisar & Aslam, 2023), assisting students in acquiring information, answering questions (Chen et al., 2023), facilitating group discussions, and resolving problems instantaneously (Rahman & Watanobe, 2023), thereby enriching students' learning experiences, offering personalized support, and potentially boosting academic performance (Kasneci et al., 2023). ChatGPT has also been effectively deployed in the design of educational materials and creative assessments, providing new avenues for content creation and curriculum development (Cotton, Cotton, & Shipway, 2023; Dijkstra et al., 2022).

### 1.2. Limitations of AI chatbots in education fields

Nonetheless, as the advancement and influence of ChatGPT permeate teaching and learning, concerns are being voiced by stakeholders and scholars. Critical issues in higher education involve assessment, examination, and academic integrity (Cotton et al., 2023; Eke, 2023; Yeadon & Halliday, 2023). Advanced generic AI tools such as ChatGPT pose a significant challenge as they are able to closely mimic students' work and are therefore difficult to distinguish from the students' own contributions. This raises concerns about untraceable plagiarism and cheating, prompting a reevaluation of many established assessment methods (Farazouli et al., 2023). Critics have also pointed out ChatGPT's technical limitations, as it is known to sometimes create incorrect information and hallucinate (Weise & Metz, 2023), leading to reliability concerns. Bogost (2022, p. 1) notes that "ChatGPT and the technologies that underlie it are less about persuasive writing and more about superb bullshitting." Large language models like ChatGPT heavily depend on training data sourced from the internet, resulting in outputs that can reflect existing biases present in the data. This raises concerns about the potential reinforcement of societal biases, imbalances, or prejudices that are prevalent in online content (Deng & Lin, 2022). Consequently, a lack of understanding of AI capabilities and limitations amongst end users might lead to misuse and over-reliance on AI technologies by teachers and learners (Kasneci et al., 2023). As AI becomes more deeply integrated into educational systems, there are also significant concerns surrounding copyright issues, data privacy, and security (ibid.). On the other hand, the power of ChatGPT and other AI tools is likely to provide an advantage to some users over non-users, creating an imbalance in the educational landscape (Cotton et al., 2023). This can further reinforce inequalities, as students with access to and experience with sophisticated AI tools can outperform those without them (Adiguzel et al., 2023). Lastly, the sustainability of these technologies and their energy consumption cannot be ignored, since the substantial power needs of AI systems can contribute to a significant environmental footprint (Kasneci et al., 2023). These considerations indicate that the emergence of ChatGPT – in addition to the euphoria about its potential – has also resulted in a more critical discourse about AI and its impact on education and society; it is likely that AI deployment in education will affect diverse groups in various educational contexts differently, raising questions of power, disadvantage and marginalization (Selwyn, 2022).

## 2. Aim of the study

Comprehensive and systematic empirical research to support or reject claims made about the benefits and challenges of AI in education, and in particular large language models is, however, scarce, and research on AI chatbots in education is still very much in a state of evolution (Hwang & Chang, 2021; Rudolph et al., 2023). Data from stakeholder perspectives are sorely needed to provide a basis for more informed discussions and decision-making (e.g., Bates et al., 2020; García-Peñalvo, 2023; Rudolph et al., 2023). To this end, this paper aims to examine students' familiarity with, usage of and attitudes towards ChatGPT and other AI chatbots for different student populations based on gender, academic level and field of study. We thereby draw from a dataset of nearly 6000 questionnaire responses across many universities and diverse academic disciplines in Sweden. A preliminary summary of the survey data (i.e., general frequencies of AI chatbot usage and attitudes and respondent comments) has been reported in Malmström et al. (2023). In this study, we focus on descriptive, inferential, and correlational statistical analyses to provide a nuanced and comprehensive understanding of different groups of university students' AI usage and attitudes. Specifically, the study is guided by the following four research questions.

- *What is the overall prevalence and pattern of AI chatbot usage among students in higher education? (RQ1)*
- *What are the general attitudes towards AI chatbots among students in higher education? (RQ2)*
- *How do gender, academic level, and field of study influence the usage and attitudes toward AI chatbots in an educational context? (RQ3)*
- *What is the relationship between students' attitudes towards AI chatbots and student's reported use of ChatGPT in their learning process? (RQ4)*

RQ 1 and RQ2 highlight the present state and impact of AI chatbots in higher education, focusing on AI chatbots as student-facing tools. Prior research has identified generally positive attitudes of the public towards the use of AI in education (Latham & Goltz, 2019). Given the recency of the breakthrough of generative AI, little empirical research is available from within the higher education sector, though there are a few recent reports to build on. A recent student survey from a Belgian university found that, while a large majority of students had used some forms of AI tools for coursework, only 13% of the students had used ChatGPT (Lobet et al., 2023). A U.S. survey (Welding, 2023) among college students revealed that 43% of the respondents had experience with ChatGPT or similar applications and about one third (32%) indicated that they used or planned to use AI tools for assignment completion. About half (47%) of the American students were concerned about the impact of AI on their education and 60% reported that their instructors or schools had not (yet) specified how AI tools could be used ethically or responsibly. Nevertheless, further monitoring of students' use and attitudes toward artificial intelligence in educational settings is essential to enhance informed decision-making by stakeholders in higher education.

Technology adoption and usage have been a long-standing theme within theoretical and empirical information systems research. Several theories and models have been widely employed due to their extensive predictive and explanatory power, among them Davis' (1989) Technology Acceptance Model (TAM), the Theory of Planned Behavior (Ajzen, 1991), and Venkatesh et al.'s (2003) Unified Theory of Acceptance and Use of Technology (UTAUT) as a consolidated framework of several earlier models and theories including the two others. These theories link perceptions and beliefs with a user's intent and actual use of technology moderated by individual and contextual factors (e.g., Liu & Ma, 2023). Loosely building on this theoretical tradition, RQ3 investigates the link between demographics, attitudes, and AI chatbot usage. By examining the influence of variables such as gender, academic level, and field of study on AI chatbot usage and attitudes, we aim to uncover existing disparities or biases that may need to be addressed by educators and educational institutions. This is vital for ensuring that all students, regardless of their demographics or academic backgrounds, can benefit from AI chatbots in education. Additionally, it offers a crucial background for understanding the social dynamics that might be at play in the adoption and acceptance of AI chatbots in a learning environment. Gender has been shown to be a moderating variable in AI adoption (e.g., Nouraldeen, 2022). However, gender effects tend to be

stronger for older people as gender stereotypes are less likely to be prominent among younger generations (Morris et al., 2005). Further, technology adoption has been linked to performance expectancy for males and ease of use for females (Venkatesh & Morris, 2000). Both effects are relevant for the context of this study as students tend to be younger and ChatGPT stands out with its ease of use. Thus, while gender effects might be present among students, we expect the differences to be low.[1] The relationship between academic level and technology adoption in general, or AI in particular, has received very little scholarly attention. While Abu-Shanab (2011) found that the educational level did play a moderating role, Sandu and Gide (2019) could not identify a relationship between gender, age, or the level of education of the students and the adoption of chatbot technology in higher education. Similarly, research on the academic discipline of students as a moderating factor is scarce (Chiu et al., 2023), though existing evidence suggests that there are greater barriers to technology adoption within the fields of arts and humanities compared to technological fields (e.g., Mercader & Gairín, 2020).

RQ4 delves into the interplay between perception and practice, a crucial aspect for the effective integration of AI chatbots in education. Investigating the relationship between students' attitudes towards AI chatbots and their actual usage of these tools in their learning process helps us understand the behavioral dynamics behind the adoption of AI in education. As an analysis for all AI chatbots would exceed the scope and length constraints of this article, we focus on ChatGPT as the currently most popular AI chatbot. In line with UTAUT, positive attitudes regarding chatbots are expected to be substantially correlated with chatbot usage (Alzahrani, 2023); conversely, concerns about perceived risks and ethical issues that reflect a lack of trust in current AIED are anticipated to diminish the adoption of chatbots (e.g., Qin et al., 2020).

In answering these four questions, this paper will contribute to a more nuanced understanding of the role of AI chatbots in higher education, in the context of the potentially disruptive impact of ChatGPT. We expect our findings to guide developers, educators, and policymakers in exploring the potential of AI while remaining cognizant of students' perspectives and needs.

## 3. Methodology

### 3.1. Research approach

The methodology employed for this study follows a quantitative research paradigm, suitable for examining usage, attitudes, and correlations across a large sample of participants (Creswell, 2014). Our research design centered on survey research. We applied single-item measures to keep the survey doable within less than 5 minutes. While this method introduces a limitation, as reliability measures cannot be applied, it reduces respondent burden, survey fatigue and attrition. Additionally, past research supports the predictive validity of single-item measures for constructs that are concrete and easily and uniformly imagined (Bergkvist & Rossiter, 2007; Rossiter, 2002) as we argue is the case in this study.

The survey items, partly inspired by the groundwork laid by Welding's (2023) study on AI usage in American colleges, were developed and piloted among a small sample of the target population with the objective to make sure the questions were understandable and to verify that the time it took to complete the survey. Subsequently, some minor refinements were made to a few items based on the pilot feedback. The finalized survey was then launched through the online survey platform Questback during the spring of 2023. The survey was split into two main sections addressing chatbot usage and attitudes towards AI in education.

### 3.2. Instrument and data collection

The survey questions can be found in Appendix 1. The first section of the survey aimed to gauge the students' familiarity with and usage of ten AI chatbots that supposedly are used for educational purposes, including but not limited to ChatGPT.[2] For each chatbot, respondents were asked: *'Rate your familiarity and frequency of use with a selection of AI chatbots,'* followed by a four-item ordinal scale as answer categories: *Unfamiliar*; *Familiar but never use it; Familiar but rarely use it;* and *Familiar and regularly use it.*

The second section focused on student attitudes towards AI in education. This covered general attitudes towards chatbots in education, perceived effects of chatbot use on learning and academic performance, ethical concerns, and issues related to institutional guidelines on chatbot usage. The latter section used a response format comprising ten agree-disagree statements, along with a "don't know/prefer not to say" option. Background information pertaining to gender, field of study, and academic level was also collected from respondents (see Table 1). Although we also asked about university affiliation, this aspect was not analyzed in this study. A link to the survey was disseminated via multiple channels, encompassing networks with various Swedish universities and a promotional campaign on social media platforms, such as LinkedIn and Facebook. The survey was open from April 8, 2023, to May 5, 2023, securing a sample of 5894 students; this convenience sample broadly mirrored the national distribution of students in terms of gender, academic level, and discipline.

### 3.3. Data analysis

The analysis of the collected data encompassed both descriptive and inferential statistical methods using the Statistical Package for the Social Sciences (SPSS) software, version 28.0.1.1 (14) developed by IBM. Descriptive statistics were employed to address RQ1 and RQ2 by summarizing the distribution of responses about chatbot usage and attitudes

**Table 1**
Descriptive statistics of the demographic characteristics of participants.

| Characteristics | Frequency | Proportion (%) |
|---|---|---|
| **Gender** | | |
| Male | 2711 | 46.0 |
| Female | 2871 | 48.7 |
| Non-binary | 86 | 1.5 |
| n.d. | 226 | 3.8 |
| **Academic Level** | | |
| First cycle (undergraduate) | 3950 | 67.0 |
| Second cycle (graduate/masters) | 1714 | 29.1 |
| Third cycle (postgraduate/PhD) | 142 | 2.4 |
| n.d. | 88 | 1.5 |
| **Field of Study** | | |
| Technology (including Engineering) | 1933 | 32.8 |
| Social Sciences (including Law, Business and Pedagogy) | 1532 | 26.0 |
| Humanities (including Theology and Art) | 998 | 17.0 |
| Medicine and health care | 539 | 9.2 |
| Natural science | 494 | 8.4 |
| Other/n.d. | 398 | 6.8 |

[1] It is important to recognize that gender is not a binary concept and many individuals do not strictly identify as male or female. Unfortunately, our data on non-binary individuals was insufficient to be included in the main analysis, but the mean values for both chatbot usage and attitudes were consistently in-between those of the male and female populations.

[2] In the survey we list Bing AI and CoPilot as separate chatbots. As of November 2023 Microsoft rebranded Bing AI as CoPilot. However, at the time of the survey, those were still separate chatbots that we asked the respondents about.

towards chatbots.

To answer RQ3, we applied chi-square tests to test for differences between the sub-groups, complemented by Cramer's V as non-parametric effect size measure and a post hoc evaluation of corrected standardized residuals. Those tests were suitable as they correspond to the categorial and ordinal nature of our variables without the need to meet the assumption of normality. Non-binary students were not included in the analysis of gender effects due to the low response frequency.

To address RQ4, we adopted the Kruskal-Wallis Test as a non-parametric equivalent of the one-way analysis of variance (ANOVA) without the need to meet assumptions of normality and homogeneity of variance. Post-hoc analyses were conducted using Mood's Median Test with a significance level of 0.05. and an additional component to identify homogeneous subsets of our data.

Finally, we conducted a correlational analysis to assess the relationship between the usage of ChatGPT as the currently most popular AI chatbot in this analysis and students' attitudes towards AI in education. Again, we employed Spearman's rho correlation as the best fit for the ordinal nature of our data.

Overall, we chose a conservative approach as a rigorous statistical foundation for our analysis. In choosing to analyze our data as ordinal, we minimize the risk of Type I errors (falsely rejecting the null hypothesis) by prioritizing the inherent order of the response options, while not imposing assumptions of continuity and interval equality of the response options. However, this methodological choice comes with the trade-off of potentially increased Type II errors, where we might fail to detect actual differences or relationships due to the tests' more stringent criteria for significance. Despite this, we believe that this approach strengthens the validity of our findings. By adhering closely to the ordinal nature of our data, we reduce the interpretative leaps and assumptions required by parametric tests, such as ANOVA, which would treat the data as quasi-continuous. At the same time, this ensures that any observed differences or relationships are not artifacts of the analytical method but are indicative of the phenomena under study. By opting for a conservative statistical approach, we aim to enhance the credibility and reliability of our findings, acknowledging the limitations of our data's scale of measurement.

## 4. Results

### 4.1. Descriptive analysis

The descriptive analysis of students' usage of different AI chatbots, as illustrated in Table 2, reveals a varied usage pattern. The most frequently used chatbot is ChatGPT, with 35.4% of students being familiar and regularly using it. It is the only chatbot with a Median higher than *unfamiliar*. By contrast, other AI chatbots such as YouChat, ChatSonic, DialoGPT, Socratic, and Jasper Chat were rarely or never used by the majority of participants (in many cases, over 90% of the respondents claimed to be unfamiliar with these other chatbots).[3]

Table 3 provides a descriptive summary of students' attitudes toward the use of AI chatbots in education. More than half of the students (55.9%) had a positive attitude towards the use of AI chatbots in education. However, almost as many (54.2%) expressed concern about the future impact of AI chatbots on students' learning, with the result that the Median is *Agree* for both statements.

Regarding the effect of AI chatbots on learning and performance, 47.7% of respondents agreed that the chatbots they used made them more effective learners, whereas only 17.3% confirmed the positive effect of chatbots in improving their study grades. Meanwhile, only 26.8% thought that chatbots improved their general language abilities. Even fewer (17.9%) believed that AI chatbots generated better results than they could produce on their own.

Concerning ethical aspects and academic integrity, a majority of students (61.9%) expressed the view that using chatbots to complete assignments and exams amounts to cheating. However, 58% disagreed with the statement that using chatbots goes against the purpose of education, and 60.3% disagreed with the prohibition of chatbots in educational settings. Lastly, only 19.1% of the students reported that their teachers or universities have rules or guidelines on the responsible use of AI chatbots.

### 4.2. Inferential analysis

Tables 4 and 5 present differences in chatbot usage and attitudes by gender. The chi-square tests indicate statistically significant gender-based differences in familiarity and usage across all chatbots with weak to medium effects (between 0.2 and 0.38). Post-hoc analysis of the corrected standardized residuals shows significant deviations from the expected distribution in all categories with males consistently expressing high familiarity and usage than females. For ChatGPT in particular, it is interesting to note that females were more likely than expected to be familiar with the chatbot but never actually use it (r = 13.7).

Concerning attitudes towards chatbots (see Table 5), chi-square tests also reveal statistically significant gender differences across all attitude statements. The gender effects are generally weak (between 0.1 and 0.3) and manifest in different directions. Female respondents were ostensibly more concerned about the impact of AI on education (r = 5.6), considered the use of chatbots as potentially contrary to the purpose of education (r = 9.3), and viewed the use of chatbots in assignments and exams as cheating (r = 6.1) that should be prohibited (r = 9.6). Males, on the other hand, had an overall more positive attitude towards chatbots (r = 14.8) and perceived them to a greater extent as tools that can improve their learning (r = 13.3) and grades (r = 12.1).

Furthermore, we examined differences in chatbot familiarity, usage and attitudes by academic level (Tables 6 and 7). While the chi-square tests indicate statistically significant differences in the familiarity and usage of chatbots as well as most attitude statements across academic levels, these effect sizes are generally very weak (<0.1), as indicated by Cramer's V.

The corrected standard residuals show that first-cycle students had a greater likelihood of being unfamiliar with chatbots (r between 3.4 and 7.4), particularly with Bing AI and Bard AI. They also reported lower rare (r between −2.5 and −3.7) and regular chatbot usage (r between −2.2 and −3.8, non-significant for Bard AI). Second-year students generally showed a reversed trend with fewer students than expected being unfamiliar with chatbots across all five chatbots (r between −3.6 and −6.7). A higher proportion of second-year students used ChatGPT (r = 4.5) and CoPilot (r = 2.0) on regular base. Notably, third-cycle students, while showing no significant differences for ChatGPT and most other categories, had significant positive residuals for the regular usage of all other chatbots (r between 3.3 and 5.1) indicating a higher reliance on AI technologies beyond the popular ChatGPT among this group.

With regard to attitudes towards chatbots (see Table 7), first-cycle students generally were more negative compared to other students. Not only did we find less agreement towards an overall positive attitude (r = −3.9), but also regarding the efficacy of chatbots in improving their learning effectiveness (r = −3.9), language ability (r = −5.9), and study grades (r = −4.6). Further, these students had stronger reservations about the role of chatbots in education. This is reflected in the tendencies towards seeing the use of chatbots for completing assignments as cheating (r = 3.4) that should be prohibited (r = 3.7) and that goes against the purpose of education (r = 4.4).

In contrast, second-cycle students displayed overall more positive attitudes towards chatbots (r = 3.9). For example, they were more likely to agree that chatbots enhance their effectiveness as learners (r = 4.0), improve their study grades (r = 4.3) as well as their language ability (r

[3] For that reason, YouChat, ChatSonic, DialoGPT, Socratic, and Jasper Chat were excluded from the inferential analysis.

**Table 2**
Descriptive analysis of students' usage of different chatbots.

| Chatbots | Familiar & regularly use it (4) | Familiar but rarely use it (3) | Familiar but never use it (2) | Unfamiliar (1) | N | Median |
|---|---|---|---|---|---|---|
| ChatGPT | 2085 (35.4%) | 1623 (27.6%) | 1877 (31.0%) | 297 (5.0%) | 5882 | 3 |
| Bing AI | 136 (2.3%) | 378 (6.5%) | 1840 (31.6%) | 3472 (59.6%) | 5826 | 1 |
| CoPilot | 133 (2.3%) | 172 (3.0%) | 833 (14.3%) | 4672 (80.4%) | 5810 | 1 |
| OpenAI playground | 113 (1.9%) | 316 (5.4%) | 1274 (21.9%) | 4108 (70.7%) | 5811 | 1 |
| Bard AI | 17 (0.3%) | 78 (1.3%) | 1248 (21.4%) | 4476 (76.9%) | 5819 | 1 |
| YouChat | 13 (0.2%) | 28 (0.5%) | 464 (8.0%) | 5306 (91.3%) | 5811 | 1 |
| ChatSonic | 12 (0.2%) | 52 (0.9%) | 581 (10.0%) | 5165 (88.9%) | 5810 | 1 |
| DialoGPT | 11 (0.2%) | 25 (0.4%) | 333 (5.7%) | 5432 (93.6%) | 5801 | 1 |
| Socratic | 10 (0.2%) | 43 (0.7%) | 439 (7.6%) | 5310 (91.5%) | 5802 | 1 |
| Jasper Chat | 9 (0.2%) | 41 (0.7%) | 452 (7.8%) | 5302 (91.4%) | 5804 | 1 |

**Table 3**
Descriptive Analysis of Students' Attitudes towards AI chatbots in Education.

| Statements | Agree (3) | Don't know/ Prefer not to say (2) | Disagree (1) | N | Median |
|---|---|---|---|---|---|
| **General attitudes towards chatbots in education** | | | | | |
| The use of chatbots is common among my fellow students. | 2272 (38.7%) | 2393 (40.8%) | 1206 (20.5%) | 5871 | 2 |
| Overall, I have a positive attitude towards the use of chatbots in education. | 3284 (55.9%) | 745 (12.7%) | 1844 (31.4%) | 5873 | 3 |
| I am concerned about how AI-chatbots will impact students' learning in the future. | 3186 (54.2%) | 722 (12.3%) | 1969 (33.5%) | 5877 | 3 |
| **Effects of chatbots on learning and performance** | | | | | |
| The chatbots I use make me more effective as a learner. | 2791 (47.7%) | 1552 (26.5%) | 1512 (25.8%) | 5855 | 2 |
| The chatbots I use improve my general language ability. | 1570 (26.8%) | 1711 (29.2%) | 2575 (44.0%) | 5856 | 2 |
| Chatbots generate better results than I can produce on my own. | 1049 (17.9%) | 1856 (31.7%) | 2952 (50.4%) | 5857 | 1 |
| The chatbots I use improve my study grades. | 1013 (17.3%) | 2477 (42.3%) | 2365 (40.4%) | 5855 | 2 |
| **Ethical aspects and academic integrity** | | | | | |
| Using chatbots goes against the purpose of education. | 1638 (27.9%) | 833 (14.2%) | 3406 (58.0) | 5877 | 1 |
| Using chatbots to complete assignments and exams is cheating. | 3633 (61.9%) | 813 (13.9%) | 1424 (24.3%) | 5870 | 3 |
| Using chatbots should be prohibited in educational settings. | 1377 (23.5%) | 951 (16.2%) | 3539 (60.3%) | 5867 | 1 |
| My teacher(s) or university has rules or guidelines on the responsible use of chatbots. | 1120 (19.1%) | 3229 (55.0%) | 1519 (25.9%) | 5868 | 2 |

= 4.5). They also expressed more disagreement with ethical concerns such as perceiving the use of chatbots as going against the purpose of education (r = 5.9), as cheating (r = 2.6), or that chatbots should be prohibited (r = 6.0).

Third-cycle students, besides perceiving chatbots as contributing to their language ability to a higher extent (r = 4.6), did not show any significant residuals at all.

Table 8 shows the results of a Kruskal-Wallis test examining the effect of field of study on university students' chatbot usage and attitudes. The test value column shows that there were statistically significant differences between at least two groups of students with different fields of study for all items with $p < 0.001$. Effect sizes are small to medium and tend to be particularly strong for the items related to usage of AI chatbots. The differences between the fields can be seen in the subset columns. Each subset column shows those study fields that build a homogenous subgroup, meaning there were no significant differences between the responses from students from fields listed in the subgroup. Subsets and groups are ordered from low to high, meaning that the group on the left has the lowest usage or agreement and the highest can be found on the right.

For example, for ChatGPT, medicine and healthcare students were least familiar with it, followed by students in the humanities. The differences between the students in these fields were not statistically significant, so they build subgroup 1 with the lowest familiarity of ChatGPT. In the same way, students from the natural sciences and social sciences showed significantly higher usage of ChatGPT than the former two, building subgroup 2. Students of technology and engineering expressed significantly higher familiarity with ChatGPT than all other groups and are therefore alone in subgroup 3. Students from a certain field of study can also belong to several subgroups (as for example for Bard AI). Here, students from the natural sciences showed higher usage than students from medicine, humanities and social sciences, but the difference was only statistically significant when compared to the lowest group, medicine students. Thus, students of humanities and social sciences are both in subgroup 1 (with medicine and healthcare students) and 2 (with natural sciences students). Looking at all the items in the survey, some consistent patterns emerge. Students in technology and engineering clearly stood out from all other groups. They used chatbots to a statistically significantly higher degree and had the least concerns about the ethical aspects of AI usage in education. Notably, for most items, the difference is statistically significant in relation to all other academic fields. Conversely, students from the arts and humanities as well as medicine and health care showed the opposite pattern; students from those fields were significantly less familiar with ChatGPT. They also expressed more concern, an overall less positive attitude towards the use of chatbots in education, and more students from these groups believed that the use of chatbots goes against the purpose of education. They were also more supportive of a prohibition of chatbots in education.

### 4.3. Correlational analysis

Table 9 shows the Spearman's rho correlation coefficients between the usage of ChatGPT and attitudes towards AI in education and chatbot usage. The results show a statistically significant relation for all attitude questions except one. The use of ChatGPT is strongly positively

**Table 4**
Chi-square and corrected standardized residuals for Chatbot Usage by Gender.

| Chatbot familiarity and usage | $\chi^2$ | p-value | Cramer's V | N | Corrected Standard Residuals r* | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | df = 3 | | | | | Unfamiliar | Familiar but never use it | Familiar but rarely use it | Familiar and regularly use it |
| ChatGPT | 420.079 | <0.001 | 0.275 | 5574 | Male | −11.2 | −13.7 | 2.3 | 16.2 |
| | | | | | Female | 11.2 | 13.7 | −2.3 | −16.2 |
| Bing AI | 811.955 | <0.001 | 0.383 | 5523 | Male | −28.0 | 20.0 | 13.1 | 7.8 |
| | | | | | Female | 28.0 | −20.0 | −13.1 | −7.8 |
| CoPilot | 393.351 | <0.001 | 0.267 | 5510 | Male | −19.3 | 13.5 | 8.7 | 9.7 |
| | | | | | Female | 19.3 | −13.5 | −8.7 | −9.7 |
| OpenAI playground | 228.143 | <0.001 | 0.204 | 5509 | Male | −14.5 | 9.6 | 9.2 | 3.8 |
| | | | | | Female | 14.5 | −9.6 | −9.2 | −3.8 |
| Bard AI | 696.552 | <0.001 | 0.355 | 5518 | Male | −26.4 | 25.5 | 4.5 | 2.1 |
| | | | | | Female | 26.4 | −25.5 | −4.5 | −2.1 |

*all differences are statistically significant ($p < 0.005$, corrected standardized residuals ± 1.96).

**Table 5**
Differences in chatbot attitudes by gender.

| Attitudes towards chatbots | $\chi^2$ | p-value | Cramer's V | N | Corrected Standard Residuals r* | | | |
|---|---|---|---|---|---|---|---|---|
| | df = 2 | | | | | Disagree | Don't know | Agree |
| The use of chatbots is common among my fellow students. | 88.556 | <0.001 | 0.126 | 5564 | Male | −3.2 | −6.7 | 9.4 |
| | | | | | Female | 3.2 | 6.7 | −9.4 |
| Overall, I have a positive attitude towards the use of chatbots in education. | 220.849 | <0.001 | 0.199 | 5567 | Male | −10.3 | −7.8 | 14.8 |
| | | | | | Female | 10.3 | 7.8 | −14.8 |
| I am concerned about how AI-chatbots will impact students' learning in the future. | 68.817 | <0.001 | 0.111 | 5571 | Male | 8.2 | −3.4 | −5.6 |
| | | | | | Female | −8.2 | 3.4 | 5.6 |
| The chatbots I use make me more effective as a learner. | 176.018 | <0.001 | 0.178 | 5550 | Male | −7.5 | −7.6 | 13.3 |
| | | | | | Female | 7.5 | 7.6 | −13.3 |
| The chatbots I use improve my general language ability. | 62.072 | <0.001 | 0.105 | 5552 | Male | −5.4 | 7.4 | |
| | | | | | Female | 5.4 | −7.4 | |
| Chatbots generate better results than I can produce on my own. | 86.708 | 0.003 | 0.125 | 5552 | Male | −6.4 | 8.5 | |
| | | | | | Female | 6.4 | −8.5 | |
| The chatbots I use improve my study grades. | 146.995 | <0.001 | 0.163 | 5552 | Male | −5.1 | −4.2 | 12.1 |
| | | | | | Female | 5.1 | 4.2 | −12.1 |
| Using chatbots goes against the purpose of education. | 171.801 | <0.001 | 0.176 | 5570 | Male | 13.1 | −6.7 | −9.3 |
| | | | | | Female | −13.1 | 6.7 | 9.3 |
| Using chatbots to complete assignments and exams is cheating. | 68.334 | <0.001 | 0.111 | 5564 | Male | 8.3 | −1.7 | −6.1 |
| | | | | | Female | −8.3 | 1.7 | 6.1 |
| E3. Using chatbots should be prohibited in educational settings. | 174.644 | <0.001 | 0.177 | 5561 | Male | 13.2 | −6.6 | −9.6 |
| | | | | | Female | −13.2 | 6.6 | 9.6 |
| E4. My teacher(s) or university has rules or guidelines on the responsible use of chatbots. | 10.663 | 0.005 | 0.044 | 5563 | Male | −3.2 | 2.1 | |
| | | | | | Female | 3.2 | −2.1 | |

*only statistically significant differences are listed ($p < 0.005$, corrected standardized residuals ± 1.96).

associated with students' belief that the use of chatbots is common among fellow students (rho = 0.406), their positive attitude towards chatbots in education (rho = 0.581), and the belief that chatbots make them more effective learners (rho = 0.644). A weak positive correlation was found between the usage of ChatGPT and the belief that chatbots improve general language ability (rho = 0.197) and study grades (rho = 0.189). Medium and strong inverse correlations were found between the usage of ChatGPT and concern about AI-chatbots' impact on students' learning in the future (rho = −0.304), between the views that using chatbots contradicts the purpose of education (rho = −0.557), constitutes cheating (rho = −0.327), and that chatbots should be prohibited in educational settings (rho = −0.564). Finally, a weak negative correlation was found between ChatGPT usage and their awareness of rules or guidelines on the responsible use of chatbots (rho = −0.049). No statistically significant correlation was found between ChatGPT usage and the belief that chatbots generate better results than the respondents could produce on their own.

## 5. Discussion

This study attempted to empirically investigate university students' AI chatbot usage and attitudes towards artificial intelligence in education, following the introduction of potentially disruptive large language models, particularly ChatGPT, on a wide scale. Through this exploration, a complex landscape has been unveiled, wherein multiple factors interact to shape students' perceptions and behaviors. Notably, the analysis has provided a more nuanced understanding of the roles that gender, discipline, and academic level play in this new learning context.

First, we were interested in the general usage frequency and familiarity with different chatbots among students. Here, the hype around ChatGPT is reflected in our data since 95% of the respondents were familiar with ChatGPT and more than one third (35.4%) claimed to use it regularly. This was confirmed by a similar proportion of students stating that the use of chatbots is common among their fellow students. With regard to the true representativeness of the total student population, these numbers need to be interpreted with caution, as students already familiar with chatbots might be somewhat more likely to participate in the survey compared to unfamiliar ones. Still, our results indicate an increase in familiarity and usage compared to earlier reports (Lobet et al., 2023; Vogels, 2023; Welding, 2023) that would suffer from a similar bias, thus supporting the claim that ChatGPT, despite its novelty and limitations, has quickly attained more widespread recognition and use among students in higher education. Interestingly though, this

**Table 6**
Differences in chatbot usage by academic level.

| Chatbot familiarity and usage | $\chi^2$ | p-value | Cra mer's V | N | Corrected Standard Residuals r* | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | df = 6 | | | | | Unfamiliar | Familiar but never use it | Familiar but rarely use it | Familiar and regularly use it |
| ChatGPT | 60.152 | <0.001 | 0.072 | 5798 | First cycle | 3.4 | 5.9 | −2.8 | −4.8 |
| | | | | | Second cycle | −3.8 | −5.9 | 3.1 | 4.5 |
| | | | | | Third cycle | | | | |
| Bing AI | 73.068 | <0.001 | 0.080 | 5744 | First cycle | 7.4 | −5.7 | −2.5 | −2.4 |
| | | | | | Second cycle | −6.7 | 5.8 | | |
| | | | | | Third cycle | −2.7 | | 2.7 | 3.8 |
| CoPilot | 43.213 | <0.001 | 0.061 | 5728 | First cycle | 4.6 | −2.2 | −3.7 | −3.0 |
| | | | | | Second cycle | −3.6 | | 2.7 | 2.0 |
| | | | | | Third cycle | −3.4 | | 3.0 | 3.3 |
| OpenAI playground | 41.288 | <0.001 | 0.060 | 5729 | First cycle | 5.2 | −3.0 | −3.6 | −2.2 |
| | | | | | Second cycle | −5.4 | 3.6 | 3.3 | |
| | | | | | Third cycle | | | | 5.1 |
| Bard AI | 84.369 | <0.001 | 0.086 | 5737 | First cycle | 7.4 | −6.8 | −2.7 | |
| | | | | | Second cycle | −6.4 | 6.3 | | |
| | | | | | Third cycle | −3.4 | | 3.8 | 4.1 |

*only statistically significant differences are listed ($p < 0.005$, corrected standardized residuals ± 1.96).

popularity appeared to be restricted to ChatGPT in particular and did not expand to other chatbots at the time of the survey. Some chatbots from larger companies, such as Bing AI and Bard, were also familiar to a substantial proportion of university students in Sweden, but most had never used any of the other applications. While this pattern does not apply to AI applications in general (see e.g., Malmström et al., 2023) and might change in the future, it underscores ChatGPT's prominent role as the key driving force behind the ongoing popularization of generative AI in education. It also stresses the urgent need for educators and educational institutions to adapt education to this new situation and find ways to address the potentials and challenges connected to this technology. At the time of the survey, more than four out of five students were not aware of any rules or guidelines from teachers or their universities in this regard.

Examining students' attitudes towards chatbots more specifically, we saw that while over half of the respondents expressed an overall positive attitude towards the use of chatbots in education, almost as many expressed concerns about their impact in the future. This disparity of optimism and concern highlights the complex relationship that students have with this emerging technology. With regard to the effect of chatbots on learning and performance, almost half of the students indicated that chatbots make them more effective as learners, pointing at the potential but also the ease with which ChatGPT seems to be utilized as self-study tool, facilitator, or assistant for learning (Chen et al., 2023; Nisar & Aslam, 2023; Rahman & Watanobe, 2023).

Nevertheless, fewer than one in five of the students felt that chatbots produced better results or improved their grades. This could indicate that most students use chatbots as a complement in the learning process rather than to complete assignments and exams, though the high number of students choosing "don't know/prefer not to say" requires caution when interpreting these results. However, these findings receive some confirmation when questions about academic integrity are considered: more than sixty percent believed that the use of chatbots during examination is cheating, Nonetheless, a majority of students were against the prohibition of AI in education settings and neither thought that chatbots go against the purpose of education. Thus, it appears that many students were aware of the potential of chatbots to support the actual learning process, and more insight is needed on precisely how students use chatbots in practice. Qualitative research could potentially address this aspect.

Regarding the relationship between the familiarity and usage of AI chatbots and student attitudes, our findings provide empirical support for the predicted relationship between students' attitudes towards AI chatbots and their level of familiarity and usage. We found a strong positive correlation between ChatGPT familiarity and usage and positive attitudes towards chatbots as well as perceived benefits from their use. Conversely, experience with ChatGPT was strongly negatively correlated with concerns about the impact of AI on future learning and ethical concerns surrounding chatbot usage in education. These findings are consistent with the broader predictions of UTAUT (Venkatesh et al., 2003) and other empirical research (e.g., Alzahrani, 2023; Qin et al., 2020). They underscore the importance of the relationship between exposure and hands-on experience and beliefs about technological innovations in educational contexts.

To what extent do chatbot usage and attitudes differ for different groups of students? Our results found statistically significant and consistent differences in the responses for all three examined grouping variables (gender, academic level and field of study). Regarding gender, we found particularly strong differences in familiarity with and usage of chatbots. Female students were overall more negative and concerned about the impact of AI on learning and assessment. Thus, as predicted by UTAUT (Venkatesh et al., 2003) and confirmed by recent AIED empirical research (Nouraldeen, 2022), our results indicate that gender is still a relevant factor for AI adoption for the current student generation and in the context of ChatGPT; this may be indicative of underlying societal factors or personal experiences that necessitate further investigation. Moreover, these results prompt an essential reflection on the design and implementation of educational technology. If AIED tools like ChatGPT are to be effectively utilized and integrated across diverse student populations, the identified differences must be acknowledged and addressed. Gender-sensitive approaches, tailored interventions, and inclusive design principles may be required to ensure that AI-powered educational solutions cater to the unique needs and preferences of various student demographics.

The effect of academic level, even though existent, was very weak (the effect sizes were consistently small). Consistent with Abu-Shanab (2011), the post-hoc analysis of the residuals showed that advanced-level students exhibited slightly greater familiarity and usage. Second-cycle students were more likely to be users of ChatGPT, while third-cycle students showed a higher tendency to regularly use other AI

**Table 7**
Differences in chatbot attitudes by academic level.

| Attitudes towards chatbots | $\chi^2$ df = 4 | p-value | Cramer's V | N | Corrected Standard Residuals r* | Dis-agree | Don't know | Agree |
|---|---|---|---|---|---|---|---|---|
| The use of chatbots is common among my fellow students. | 23.386 | <0.001 | 0.045 | 5788 | First cycle | 3.2 | | −3.4 |
| | | | | | Second cycle | −3.5 | | 4.1 |
| | | | | | Third cycle | | | |
| Overall, I have a positive attitude towards the use of chatbots in education. | 15.881 | 0.003 | 0.037 | 5791 | First cycle | 2.8 | | −3.9 |
| | | | | | Second cycle | −3.0 | | 3.9 |
| | | | | | Third cycle | | | |
| I am concerned about how AI-chatbots will impact students' learning in the future. | 7.120 | 0.130 | n.s. | 5795 | First cycle | | | |
| | | | | | Second cycle | | | |
| | | | | | Third cycle | | | |
| The chatbots I use make me more effective as a learner. | 18.662 | <0.001 | 0.040 | 5772 | First cycle | | 3.5 | −3.9 |
| | | | | | Second cycle | | −3.6 | 4.0 |
| | | | | | Third cycle | | | |
| The chatbots I use improve my general language ability. | 48.496 | <0.001 | 0.065 | 5774 | First cycle | 2.2 | 3.3 | −5.9 |
| | | | | | Second cycle | | −3.0 | 4.5 |
| | | | | | Third cycle | −3.1 | | 4.6 |
| Chatbots generate better results than I can produce on my own. | 22.444 | <0.001 | 0.044 | 5775 | First cycle | | 4.4 | −3.1 |
| | | | | | Second cycle | | −4.1 | 2.8 |
| | | | | | Third cycle | | | |
| The chatbots I use improve my study grades. | 26.534 | <0.001 | 0.048 | 5773 | First cycle | 2.1 | | −4.6 |
| | | | | | Second cycle | | −2.0 | 4.3 |
| | | | | | Third cycle | −2.5 | | |
| Using chatbots goes against the purpose of education. | 38.992 | <0.001 | 0.058 | 5795 | First cycle | −6.2 | 3.1 | 4.4 |
| | | | | | Second cycle | 5.9 | −2.9 | −4.2 |
| | | | | | Third cycle | | | |
| Using chatbots to complete assignments and exams is cheating. | 12.051 | 0.017 | 0.032 | 5788 | First cycle | −2.8 | | 3.4 |
| | | | | | Second cycle | 2.6 | | −3.1 |
| | | | | | Third cycle | | | |
| Using chatbots should be prohibited in educational settings. | 39.510 | <0.001 | 0.058 | 5785 | First cycle | −6.1 | 3.9 | 3.7 |
| | | | | | Second cycle | 6.0 | −3.5 | −3.9 |
| | | | | | Third cycle | | | |
| My teacher(s) or university has rules or guidelines on the responsible use of chatbots. | 5.597 | 0.231 | n.s. | 5787 | First cycle | | | |
| | | | | | Second cycle | | −2.1 | |
| | | | | | Third cycle | | | |

*only statistically significant differences are listed ($p < 0.005$, corrected standardized residuals ± 1.96).

chatbots. These effects could be linked to more specialized AI utilization opportunities in advanced, research-based projects, or generally higher self-regulated learning skills among graduate and PhD students. Further, we found more favorable attitudes and more limited concerns about using chatbots in education for second-year students, along with the opposite tendency for first-year students. Interestingly, third-cycle students did not show the same attitude patterns as second-year students, indicating a potential non-linear relation between academic level and attitudes towards chatbots. However, given the paucity of empirical studies on this question, more research is required to draw definitive conclusions.

As for disciplinary differences, our results indicate clear and consistent differences, with students from engineering apparently using chatbots to a much higher degree and expressing stronger optimism towards AI. Conversely, students from the arts and humanities and from medicine and health care used chatbots less and were more skeptical. While the results for humanities and arts confirm prior research (Mercader & Gairín, 2020), the results for students from the medicine and healthcare sectors are surprising, given the fact that it is one of the fields most actively discussing AI adoption (e.g., Zhang et al., 2023). Our findings also contradict the results of prior research reporting that students generally have positive attitudes towards AI in medicine (Santomartino & Yi, 2022). Potentially, our students' skepticism is connected to chatbots in particular or reflects more on the health care sector as a whole than medicine alone. However, this should be considered in the sometimes euphoric discussion about the potential of ChatGPT in medical education (e.g., Lee, 2023). The consistency and decent effect sizes of our results also raise the question of the importance of academic disciplines and differences in disciplinary traditions and practices as explanatory factors in AIED in general. Our findings motivate further examination of this perspective, which is not well theorized and so far largely lacks empirical verification (see Orji, 2010 for an exception).

Generally, our findings also have pedagogical implications for teachers and academic institutions. Students' AI literacy development enabling them to critically evaluate, communicate with, and use AI technologies (Long & Magerko, 2020) are likely to be significantly shaped by the ways in which teachers insert and discuss AI technologies in their teaching practices. Our results suggest that AI-related policies and guidelines in that regard should not be one-size-fits-all; rather, support efforts need adaptation to the student characteristics and teaching methods in specific disciplinary contexts. Thus, solutions to

**Table 8**
Differences in chatbot usage and attitudes by field of study (kruskal-wallis test with homogeneous subgroups ($p < 0.05$), grouping from low to high usage and disagree to agree).

| | Kruskal-Wallis H (df = 4) | Effect size $\varepsilon^2$ | Subset 1 | Subset 2 | Subset 3 | Subset 4 | N |
|---|---|---|---|---|---|---|---|
| **Chatbot familiarity and usage** | | | | | | | |
| ChatGPT | 589.489* | 0,107 | MH | NS | T | | 5489 |
| Bing AI | 454.283* | 0,083 | M | SH | N | T | 5440 |
| CoPilot | 511.305* | 0,094 | MSH | N | T | | 5424 |
| OpenAI playground | 154.880* | 0,028 | MHSN | T | | | 5426 |
| Bard AI | 360.236* | 0,066 | MHS | HSN | T | | 5435 |
| **Attitudes towards chatbots** | | | | | | | |
| The use of chatbots is common among my fellow students. | 434.643* | 0,079 | HM | NS | T | | 5479 |
| Overall, I have a positive attitude towards the use of chatbots in education. | 306.434* | 0,055 | H | M | SN | T | 5481 |
| I am concerned about how AI-chatbots will impact students' learning in the future. | 94.910* | 0,017 | T | NS | MH | | 5487 |
| The chatbots I use make me more effective as a learner. | 389.312* | 0,071 | H | M | NS | T | 5466 |
| The chatbots I use improve my general language ability. | 52.291* | 0,009 | HS | SMN | T | | 5466 |
| Chatbots generate better results than I can produce on my own. | 13.887* | 0,002 | HSNM | NMT | | | 5468 |
| The chatbots I use improve my study grades. | 101.288* | 0,018 | HMSN | T | | | 5466 |
| Using chatbots goes against the purpose of education. | 280.126* | 0,050 | T | NS | HM | | 5486 |
| Using chatbots to complete assignments and exams is cheating. | 91.378* | 0,016 | T | SN | NMH | | 5480 |
| Using chatbots should be prohibited in educational settings. | 330.042* | 0,060 | T | NS | HM | | 5478 |
| My teacher(s) or university has rules or guidelines on the responsible use of chatbots. | 21.084* | 0,003 | TSM | HN | | | 5478 |

* … $p < 0.001$, M … Medicine and healthcare, H … Humanities and art, N … Natural sciences, S … Social sciences, T … Technology and Engineering (homogeneous subgroups are based on asymptotic significances with $\alpha = 0.05$).

handling AI in education should be developed and implemented "locally" to address the specific needs of the local student population. For example, the widespread enthusiasm and experience with AI chatbots among technology and engineering students can provide an opportunity to integrate AI chatbot tasks more deeply into curricula, while at the same time requiring particular efforts to create awareness of the technology's limitations and ethical concerns. In other fields, different instruction strategies might be required to address a lack of comfort and/or proficiency with AI chatbots, for example through feedback mechanisms, allowing students to share their concerns, experiences, and suggestions about AI chatbot usage. While there will always be variation between individual students that needs to be considered, our results can help teachers adapt their handling of AI tools to their particular teaching contexts.

## 6. Conclusions and limitations

Almost 6000 university students in Sweden answered survey questions about their usage and attitudes towards ChatGPT and other AI chatbots. The insights gleaned from this research underscore the importance of understanding student perceptions and experiences of AI chatbots in educational settings. The study has revealed the widespread usage of ChatGPT among university students. Given that we are likely only seeing the beginning of large language model applications, we agree with other educators and AIED researchers to conclude that the use of ChatGPT and other chatbots in education among students is already mainstream and likely to stay (Hajkowicz et al., 2023). While concerns about academic integrity and cheating are valid and justified, many students acknowledge the usefulness of AI for their actual learning, and our efforts should be directed towards supporting these developments. Students still need substantive training and learning, and ChatGPT should be treated as a tool rather than a replacement (Berdanier & Alley, 2023), but both students and teachers need new competencies in integrating AI chatbots in the learning process.

Case-based inspiration and examples of how AI chatbots can productively support learning are currently published on a daily basis (e.g., Santos, 2023). Nevertheless, some students benefit more from these developments than others. Our results show multifaceted and sometimes conflicting views on the role of AI in education, and these views are influenced by gender and academic discipline. Addressing the needs of different student populations will require locally adapted solutions. Given the apparent lack of guidelines and rules, many teachers and decision-makers would appear to be unprepared to this end.

This study also highlights the importance of addressing students' concerns about the potential impact of AI on their future learning. Ongoing lively discussions about the potential and dangers of AI in education need to be complemented with empirical studies of the kind presented here. In addition, qualitative research is needed to better understand how students use AI tools in practice. Ultimately, the findings from this study contribute to the growing body of literature on the role of AI in education and provide a valuable resource for developers, educators and policymakers as they navigate the emerging landscape of AI chatbots within the educational sector. It is also important to acknowledge that the data underpinning our study was collected in May of 2023, marking a specific snapshot in the rapidly evolving landscape of AI. As such, the reported AI chatbot usage and attitudes are likely to be subject to change as new applications emerge and awareness grows. Thus, the temporal context serves both as a limitation and a springboard for future research highlighting the need for continuous investigation to update and adjust the observed findings. Certain further limitations of the present study must be acknowledged. First, the sample used in this study was not random. The respondents self-selected to participate, and the topic of the survey might make students who already had some degree of exposure to chatbots in their academic settings more likely choose to engage. This potential selection bias could somewhat overstate the familiarity and usage of AI chatbots. Despite these limitations, the large sample size used in this study, encompassing thousands of responses, enhances the statistical power of the analysis and allows for the detection of even small effect sizes. Furthermore, the broad mix of respondents from different academic levels and genders provides a heterogeneous sample that has offered a rich view into the range of student experiences and attitudes towards AI chatbots. This was enabled through the use of single-response items instead of larger attitudinal constructs, which may raise concerns about the reliability of our approach. Thus, while we feel that the use of single items is justified in this particular study to provide a particular, readily interpretable snapshot of certain attitudes and behaviors, we encourage future research to build on this work by employing established or newly developed multidimensional scales for a more comprehensive understanding of the factors that drive students' perceptions and interactions

**Table 9**
Correlation between usage of ChatGPT and attitudes towards AI in education and chatbot usage.

| | Spearman's rho | p-value | 95% lower bound | 95% upper bound | N |
|---|---|---|---|---|---|
| The use of chatbots is common among my fellow students. | 0.406 | <0.001 | 0.384 | 0.428 | 5869 |
| Overall, I have a positive attitude towards the use of chatbots in education. | 0.581 | 0.000 | 0.563 | 0.598 | 5871 |
| I am concerned about how AI-chatbots will impact students' learning in the future. | −0.304 | <0.001 | −0.328 | −0.280 | 5875 |
| The chatbots I use make me more effective as a learner. | 0.644 | 0.000 | 0.628 | 0.659 | 5853 |
| The chatbots I use improve my general language ability. | 0.197 | <0.001 | 0.172 | 0.223 | 5855 |
| Chatbots generate better results than I can produce on my own. | −0.009 | 0.492 | −0.035 | 0.017 | 5856 |
| The chatbots I use improve my study grades. | 0.189 | <0.001 | 0.164 | 0.215 | 5854 |
| Using chatbots goes against the purpose of education. | −0.557 | 0.000 | −0.575 | −0.539 | 5875 |
| Using chatbots to complete assignments and exams is cheating. | −0.327 | <0.001 | −0.350 | −0.303 | 5868 |
| Using chatbots should be prohibited in educational settings. | −0.564 | 0.000 | −0.582 | −0.546 | 5865 |
| My teacher(s) or university has rules or guidelines on the responsible use of chatbots. | −0.049 | <0.001 | −0.075 | −0.022 | 5867 |

with AI chatbots in educational settings. Our study serves as a steppingstone, indicating that these larger attitudinal patterns exist and warrant deeper exploration through more robust measurement tools and the investigation of explanatory factors such as teaching practices, educator attitudes, and the overall educational context and their influence on technology adoption and educational outcomes. Future research might also consider the use of other data collection methods such as direct observation, logs of actual chatbot use, or focus group discussions to reduce bias due to social desirability or inaccurate recall in self-reported data that is common in studies of this kind. Related to this, the cross-sectional design of the study does not allow for the exploration of causal relationships or changes over time. Future research could benefit from a longitudinal design, which would provide insights into the evolution of students' perceptions and usage of AI chatbots over time. Despite these limitations, this study has provided a robust starting point for further exploration into this rapidly evolving field of study. The results provide meaningful cues for stakeholders involved in integrating and regulating AI chatbots in higher education, underscoring the crucial role of student perspectives in the discourse about educational technologies.

## Statements on ethics and open data

The data can be obtained by sending request e-mails to the corresponding author. All procedures in the study were conducted in accordance with applicable laws and institutional guidelines for research ethics ensuring strict adherence to ethical principles throughout this study. No personally identifiable or sensitive data were collected, ensuring total participant anonymity. Detailed information about the study was provided to participants before beginning the survey, reassuring them of their right to voluntary participation and withdrawal. All participants provided their informed consent for participation when submitting the survey.

## CRediT authorship contribution statement

**Christian Stöhr:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Amy Wanyu Ou:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. **Hans Malmström:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Data curation, Conceptualization.

## Declaration of generative AI and AI-assisted technologies in the writing process

Statement: During the preparation of this work, the authors used ChatGPT in order to improve readability and language in certain places of the text. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors of this report express their sincere gratitude for the assistance and support provided by the following individuals: Jenny Palm (Chalmers University of Technology); Jörg Pareigis (Karlstad University); Malin Kjellberg and Andreas Eriksson (Chalmers University of Technology); all people involved in the promotion and distribution of the survey at the different universities and the anonymous reviewers for their constructive feedback. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.caeai.2024.100259.

## References

Abu-Shanab, E. A. (2011). Education level as a technology adoption moderator. *2011 3rd International Conference on Computer Research and Development, 1*, 324–328. https://doi.org/10.1109/ICCRD.2011.5764029. IEEE.

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology, 15*(3), ep429. https://doi.org/10.30935/cedtech/13152

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Alzahrani, L. (2023). Analyzing students' attitudes and behavior toward artificial intelligence technologies in higher education. *International Journal of Recent*

*Technology and Engineering, 11*(6), 65–73. https://doi.org/10.35940/ijrte.F7475.0311623
Bates, T., Cobo, C., Mariño, O., & Wheeler, S. (2020). Can artificial intelligence transform higher education? *International Journal of Educational Technology in Higher Education, 17*(1), 1–12. https://doi.org/10.1186/s41239-020-00218-x
Berdanier, C. G. P., & Alley, M. (2023). We still need to teach engineers to write in the era of ChatGPT. *Journal of Engineering Education, 2023*(112), 583–586. https://doi.org/10.1002/jee.20541
Bergkvist, L., & Rossiter, J. R. (2007). The predictive validity of multiple-item versus single-item measures of the same constructs. *Journal of Marketing Research, 44*(2), 175–184. https://doi.org/10.1509/jmkr.44.2.175
Bogost, I. (2022). *ChatGPT is Dumber than you think. The atlantic*. https://www.theatlantic.com/technology/archive/2022/12/chatgpt-openai-artificial-intelligence-writing-ethics/672386/.
Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access, 8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510
Chen, Y., Jensen, S., Albert, L. J., Gupta, S., & Lee, T. (2023). Artificial intelligence (AI) student assistants in the classroom: Designing chatbots to support student success. *Information Systems Frontiers, 25*(1), 161–182. https://doi.org/10.1007/s10796-022-10291-4
Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers & Education: Artificial Intelligence, 4*, 100–118. https://doi.org/10.1016/j.caeai.2022.100118
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 & Teaching International*, 1–12. https://doi.org/10.1080/14703297.2023.2190148
Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.
Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008
Deng, J., & Lin, Y. (2022). The benefits and challenges of ChatGPT: An overview. *Frontiers in Computing and Intelligent Systems, 2*(2), 81–83. https://doi.org/10.54097/fcis.v2i2.4465
Dennean, K., Gantori, S., Limas, D. K., Pu, A., & Gilligan, R. (2023). Let's chat about ChatGPT. UBS https://secure.ubs.com/public/api/v2/investment-content/documents/XILxY9V9P5RazGpDA1Cr_Q?apikey=Y8VdAx8vhk1P9YXDlEOo2Eoco1fqKwDk&Accept-Language=de-CH.
Dijkstra, R., Genc, Z., Kayal, S., & Kamps, J. (2022). Reading comprehension quiz generation using generative pre-trained transformers. In S. Sosnovsky, P. Brusilovsky, & A. Lan (Eds.), *Proceedings of the fourth international workshop on intelligent textbooks 2022: Co-located with 23d international conference on artificial intelligence in education (AIED 2022)* (pp. 4–17). http://ceur-ws.org/Vol-3192/itb22_p1_full5439.pdf.
Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? *Journal of Responsible Technology, 13*, Article 100060. https://doi.org/10.1016/j.jrt.2023.100060
Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2023). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education, 13*, 1–13. https://doi.org/10.1080/02602938.2023.2241676, 2023.
García-Peñalvo, F. J. (2023). The perception of artificial intelligence in educational contexts after the launch of ChatGPT: Disruption or panic? *Education in the Knowledge Society (EKS), 24*, Article e31279. https://doi.org/10.14201/eks.31279
Hajkowicz, S., Sanderson, C., Karimi, S., Bratanova, A., & Naughtin, C. (2023). Artificial intelligence adoption in the physical sciences, natural sciences, life sciences, social sciences and the arts and humanities: A bibliometric analysis of research publications from 1960-2021. *Technology in Society, 74*, Article 102260. https://doi.org/10.1016/j.techsoc.2023.102260
Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education, 57*(4), 542–570. https://doi.org/10.1111/ejed.12533
Hwang, G.-J., & Chang, C.-Y. (2021). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments*, 1–14. https://doi.org/10.1080/10494820.2021.1952615
Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers & Education: Artificial Intelligence, 1*, Article 100001. https://doi.org/10.1016/j.caeai.2020.100001
Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274
Latham, A., & Goltz, S. (2019). A survey of the general public's views on the ethics of using AI in education. In S. Isotani, E. Millán, A. Ogan, P. Hastings, B. McLaren, & R. Luckin (Eds.), *Lecture notes in computer science: 11625. Artificial intelligence in education. AIED 2019*. Cham: Springer. https://doi.org/10.1007/978-3-030-23204-7_17.
Lee, H. (2023). The rise of ChatGPT: Exploring its potential in medical education. *Anatomical Sciences Education, 0*(0), 1–6. https://doi.org/10.1002/ase.2270
Li, L., Ma, Z., Fan, L., Lee, S., Yu, H., & Hemphill, L. (2023). ChatGPT in education: A discourse analysis of worries and concerns on social media. *arXiv:2305.02201*. https://doi.org/10.48550/arXiv.2305.02201
Liu, G., & Ma, C. (2023). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching, 0*(0), 1–14. https://doi.org/10.1080/17501229.2023.2240316
Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), 410. https://doi.org/10.3390/educsci13040410
Lobet, M., Honet, A., & Wathelet, V. (2023). *Use of ChatGPT by students: Artificial intelligence is all the rage!*. https://newsroom.unamur.be/en/news/chatgpt-by-students-ai-all-rage.
Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3313831.3376727
Mercader, C., & Gairín, J. (2020). University teachers' perception of barriers to the use of digital technologies: The importance of the academic discipline. *International Journal of Educational Technology in Higher Education, 17*(1), 4. https://doi.org/10.1186/s41239-020-0182-x
Morris, M. G., Venkatesh, V., & Ackerman, P. L. (2005). Gender and age differences in employee decisions about new technology: An extension to the theory of planned behavior. *IEEE Transactions on Engineering Management, 52*(1), 69–84. https://doi.org/10.1109/TEM.2004.839967.
Malmström, H., Stöhr, C., & Ou, A.W. (2023). *Chatbots and other AI for learning: A survey of use and views among university students in Sweden.* (Chalmers Studies in Communication and Learning in Higher Education 2023:1) https://doi.org/10.17196/cls.csclhe/2023/01.
Nemorin, S., Vlachidis, A., Ayerakwa, H. M., & Andriotis, P. (2023). AI hyped? A horizon scan of discourse on artificial intelligence in education (AIED) and development. *Learning, Media and Technology, 48*(1), 38–51. https://doi.org/10.1080/17439884.2022.2095568
Nisar, S., & Aslam, M. S. (2023). Is ChatGPT a good tool for T&CM students in studying pharmacology? *SSRN Scholarly Paper (4324310)*. https://doi.org/10.2139/ssrn.4324310
Nouraldeen, R. M. (2022). The impact of technology readiness and use perceptions on students' adoption of artificial intelligence: The moderating role of gender. *Development and Learning in Organizations: An International Journal, 37*(3), 7–10. https://doi.org/10.1108/DLO-07-2022-0133
OpenAI. (2022). Introducing ChatGPT. Retrieved July 11, 2023, from https://openai.com/blog/chatgpt.
Orji, R. O. (2010). Effect of academic discipline on technology acceptance. In *2010 international conference on education and management technology* (pp. 617–621). IEEE. https://doi.org/10.1109/ICEMT.2010.5657581.
Qin, F., Li, K., & Yan, J. (2020). Understanding user trust in artificial intelligence-based educational systems: Evidence from China. *British Journal of Educational Technology, 51*(5), 1693–1710. https://doi.org/10.1111/bjet.12994
Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences, 13*(9), 5783. https://doi.org/10.3390/app13095783
Rossiter, J. R. (2002). The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing, 19*(4), 305–335. https://doi.org/10.1016/S0167-8116(02)00097-6
Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching, 6*(1), 342–363. https://doi.org/10.37074/jalt.2023.6.1.9
Sandu, N., & Gide, E. (2019). Adoption of AI-chatbots to enhance student learning experience in higher education in India. In *18th international conference on information technology based higher education and training (ITHET)* (pp. 1–5). https://doi.org/10.1109/ITHET46829.2019.8937382
Santomartino, S. M., & Yi, P. H. (2022). Systematic review of radiologist and medical student attitudes on the role and impact of AI in radiology. *Academic Radiology, 29*(11), 1748–1756. https://doi.org/10.1016/j.acra.2021.12.032
Santos, R. P. dos (2023). Enhancing chemistry learning with ChatGPT and bing chat as agents to think with: A comparative case study. *arXiv:2305.11890*. https://doi.org/10.48550/arXiv.2305.11890
Selwyn, N. (2022). The future of AI and education: Some cautionary notes. *European Journal of Education, 57*(4), 620–631. https://doi.org/10.1111/ejed.12532
Tahiru, F. (2021). AI in education: A systematic literature review. *Journal of Cases on Information Technology, 23*(1), 1–20. https://doi.org/10.4018/JCIT.2021010101
Venkatesh, V., & Morris, M. G. (2000). Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. *MIS Quarterly, 24*(1), 115–139. https://doi.org/10.2307/3250981
Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540
Weise, K., & Metz, C. (2023). *When A.I. Chatbots hallucinate*. The New York Times https://www.nytimes.com/2023/05/01/business/ai-chatbots-hallucination.html.
Welding, L. (2023). Half of college students say using AI on schoolwork is cheating or plagiarism. *BestColleges*. https://www.bestcolleges.com/research/college-students-ai-tools-survey/.
Yeadon, W., & Halliday, D. P. (2023). Exploring durham university physics exams with Large Language models. arXiv preprint arXiv:2306.15609 https://doi.org/10.48550/arXiv.2306.15609.
Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the

educators? *International Journal of Educational Technology in Higher Education, 16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0

Zhang, W., Cai, M., Lee, H. J., Evans, R., Zhu, C., & Ming, C. (2023). AI in Medical Education: Global situation, effects and challenges. *Education and Information Technologies*. https://doi.org/10.1007/s10639-023-12009-8

Vogels, E.A. (2023). A majority of Americans have heard of ChatGPT, but few have tried it themselves. *Pew Research Center*. https://pewrsr.ch/3ICYoIx.