The Future of Engineering Education
2024 Annual Conference & Exposition
Oregon Convention Center
Portland, OR . June 23 - 26, 2024
Paper ID #44397

Characterizing Computing Students' Use of Generative AI

Maura Lyons, Codio

Maura is a Marketing Associate at Codio with a BA in Psychology and English.

Elise Deitrick, Codio

Elise has a BS in Computer Science and PhD in STEM Education. Her thesis was on interdisciplinary, collaborative computing using mixed methodologies. Elise combines her over a decade of teaching experience with her research background to create evidence-based computing education tools in her current role at Codio.

Mr. Joshua Richard Coughlin Stowell Ball, Codio

Joshua Ball is Codio's Vice President of Marketing and a Senior Fellow at the National Institute for Deterrence Studies.

Characterizing Computing Students' Use of Generative AI

Abstract

While the discussion of Generative AI in education has been centered on academic integrity and uses in learning contexts from a teacher and administrator perspective, there is less work understanding students' adoption, use, and perspectives on this new technology.

This paper reports on a survey of 371 US college students taking computing courses. We first asked what services are being used, how much they are paying for them, what they are using them for, and how long they have been using AI. We dig further into their use of AI tools in their schoolwork by asking about what subjects they use AI for, what they use AI for, and what causes them to *not* use it. Turning to their computing courses, we determine their use of AI, how useful they find AI tools, how they ensure academic integrity, and how they characterize their computing courses' framing of the use of AI tools.

We found that the majority of students pay for GenAI tools despite readily available free versions. Students use GenAI tools primarily to understand jargon, such as understanding teacher-written programming assignment prompts and developer-written compiler messages as opposed to potentially problematic uses such as generating code. In fact, students' main motivation to *not* use GenAI tools on graded assignments was they like to do their own work. Notably, students who were taught how AI works had significantly different views on AI tools' impact on academic integrity concerns.

Computing students' use of Generative AI is growing, and thoughts on academic integrity are far from decided – but there does seem to be an opportunity to teach students the variety of ways it can be used effectively for programming tasks.

Introduction

ChatGPT, a Generative AI product developed by OpenAI, was released in November 2022 and almost immediately, its popularity began to surge worldwide, as illustrated by its steep increase as a search term on Google. Teachers and administrators took notice – "'plagiarism' was ranked in two out of the top five related search queries alongside 'ChatGPT'" [1]. The popularization of AI tools raises concerns about plagiarism rates and the ethical use of technology in academic settings. Educational institutions are actively crafting policies to navigate the complexities of GenAI usage while maintaining academic integrity [2], [3].

Recent advancements in GenAI have ushered in a new era for educational methodologies, offering innovative tools for learning and teaching. Integrating GenAI tools such as ChatGPT and MidJourney into educational practices is becoming increasingly common, with these tools predicted to become as ubiquitous as traditional software like Microsoft Excel in the near future [4]. The emergence of GenAI necessitates reevaluating pedagogical strategies, suggesting a shift towards technology-integrated learning environments that promote adaptive and personalized education [5]. However, this shift also amplifies concerns regarding academic honesty, calling for a nuanced redefinition of academic integrity in the GenAI era [6], [7].

In light of these developments, educational policy-makers and academic institutions are exploring diverse approaches to integrate GenAI tools ethically and effectively. For instance, UNESCO has issued guidance on using GenAI in education and research, highlighting the rapid advancement of these tools and the need for adaptive policy frameworks [8].

Literature Review

Recent research [9] - [10] highlights the already widespread use of Generative AI in higher education, examining both benefits and drawbacks.

Various researchers have noted the ways in which these new Generative AI tools, or GenAI for short, can mimic student achievement in current higher education coursework. Zastudil [9] summarized that GenAI models can:

- Answer multiple choice quizzes slightly worse than students
- Complete basic programming assignments
- Produce code explanations better than students

With continuous improvements to GenAI models (e.g. GPT-4), the differences between students' and GenAI tools' capabilities seem to be disappearing. This leads to administrator and faculty concerns about the potential for plagiarism in higher education institutions, especially given the difficulty human graders have in detecting work written by Large Language Models or LLMs (e.g. [1], [9], [11]).

Other research has begun to address how computing education needs to change to reflect the new professional landscape graduates are entering where employees are expected to seamlessly integrate GenAI tools into their workflows for improved efficiency. Some faculty are providing GenAI tools to be used during the course, such as Harvard's CS50 Duck Debugger, allowing students to practice leveraging such tools. Others are diving into the deeper pedagogical implications, such as Agarwal and colleagues [12], who highlight that teachers might need to shift focus from students' ability to write code from scratch to students' ability to critique code, potentially through the use of refute-style assessments.

Turning to the students themselves, researchers conducted surveys to get a sense of adoption rate and uses. Unfortunately, large-scale, general surveys about AI in education often focus on writing essays and have limited reporting on coding assignments. When research is focused on computing students, we see much more robust usage. For example, Amoozadeh and team [10] surveyed over 100 computing students at a large US university and found 76% reported having used GenAI tools, and 65% reported using them to complete programming tasks.

Notably, these student-facing surveys and interviews have not dug into what exactly students are doing with GenAI, their reflections on its usefulness, and their perceptions on how it is changing their educational journey in terms of academic integrity and career prospects.

Methods

We recruited survey participants via SurveyMonkey's proprietary audience database and network, "SurveyMonkey Audience," which consists of over 175 million individuals around the world who have completed other SurveyMonkey instruments. SurveyMonkey balances their US database of survey takers according to census data of age and gender. Survey takers received either donations to a charity of their choice, a chance to win a sweepstakes prize, or credits which they can redeem for gift cards.

The eligibility requirements were being a current university student and being enrolled in a computing course in the United States at the time of the survey. A total of 371 complete, eligible responses were received. Only the respondents who completed the survey were considered in the subsequent analysis.

Survey Design

Though our survey design drew on past research such as literature reviewed, it does not seek to exactly replicate any one study as, at the time of investigation, there was no existing literature that surveyed exactly what we were interested in. Our survey was designed to capture three informational themes - awareness and use of AI in education, AI in computing education, and academic integrity. The survey instrument questions are listed by theme in Table 1.

The questions in the awareness and use of AI in education theme aim to capture general usage patterns both inside and outside of school. We build on Zastidul [9] semi-structured interviews evaluating awareness of AI by asking about specific details of usage. Inspired by Amoozadeh and colleagues' [10] finding that 76% of students use Generative AI tools, we specifically investigate why students would choose not to use these tools for schoolwork.

Next, the questions in the AI in computing education aim to capture students' usage, specifically around programming. We chose to ask the general perspective questions (i.e. helpfulness, accuracy and time savings) in this section since the context of computing education is our main focus. We aim to reproduce Amoozadeh and colleagues' [10] finding that 65% of students use them to complete programming tasks and dive into what specifically they are using AI tools for.

We also explore student perceptions found in Zastidul [9] and Amoozadeh [10] around usefulness at a larger scale.

Finally, the questions in the academic integrity theme aim to capture students' experiences and observations on how these tools are used and framed in academic institutions. We aim to replicate Zastidul [9] finding that students believe the amount of plagiarism will increase as AI tools increase in popularity. We expand this work by capturing the context of the institutions and course policies around AI usage.

Table 1: List of Survey Questions by theme

Theme	**Questions**
Awareness and Use of AI in Education	● Do you currently use an AI product? ● What AI product(s) do you currently use? ● How much do you pay for the AI products you use? ● What do you use AI for? ● When did you start using AI? ● In school, what types of courses do you use AI in? ● What do you use AI for when working on schoolwork? ● When you choose not to use AI for school work, what are the main reasons?
AI in Computing Education	● Which of the following were you aware that AI could help you within your computing courses? ● During your Computing course, did you use AI for any of the following? ● In your computing coursework, how much time per week do you think AI saved you? ● How do you find the helpfulness of AI tools when working on computing coursework? ● How do you find the accuracy of AI tools when working on computing coursework?
Academic Integrity and AI	● When using AI tools to help you with computing coursework, how do you ensure the academic integrity of what you submit? ● From your experiences or discussions with other students in your computing courses, how does AI affect academically dishonest actions? ● How does AI affect the detection of academically dishonest actions? ● How does your computing course frame the use of AI tools? ● How concerned are you about AI replacing software developers and other programming jobs?

The SurveyMonkey software reported that participants completed the survey in an average of 4 minutes.

The survey consisted of multiple choice questions that either allowed a single answer (e.g. "How much do you pay for the AI products you use?") or multiple answers (e.g. "What AI product(s) do you currently use?"), so the number of responses for each question varied.

The descriptive results section below lists individual question responses' results.

Respondent Demographics

Overall gender of respondents was 55% female and 45% male. While non-binary, a gender not listed here or prefer not to answer options were provided, no respondent selected them.

Interestingly, most respondents would be considered “non-traditional” students (22+). The largest group, aged 45 to 60, comprised 43% of the total, followed by the 18 to 29 years age group at 29%, the 30 to 44 years group at 23%, and those aged 60 or older at 5%.

The most common major of respondents was Business (144) followed by STEM (134) and Computing (125). Slightly less common were Arts (73) and Social Sciences (62) with 10 respondents reporting Other.

Respondents self-reported years of programming experience ranging from 0 to 61 years. The average number of years was 5 while the median was 3. This indicates that while there were several students with a lot of years of experience, half of respondents had 3 or fewer years of programming experience.

Limitations

As we are using an unstandardized survey instrument, there is a risk for bias. Additionally, as we are recruiting a disparate community, there is a risk of not having a representative set of respondents. Notably, as described in the demographics section above, our sample does not reflect the typical computing student population in terms of gender and age distributions. It is worth noting that this survey was voluntary, and may have skewed the respondent pool to more heavily represent students interested in Generative AI.

Statistical Analysis

We use two statistical tests throughout the Statistical Results section to explore the relationship between various dimensions of student’s GenAI perspectives and use.

When both of the two dimensions consist of ordinal data such as student’s perception of helpfulness or frequency of AI tool use, Spearman regression analysis is used. First, we process data by mapping the textual responses to integers reflecting the ordinal nature of the data. Then we use the Spearmanr method from python’s SciPy library to calculate both ϱ, the Spearman regression coefficient and the p-value. We then use the mean, standard deviation and square root functions from python's numpy library to manually compute the Cohen's d value to represent effect size.

If either or both of the dimensions is not ordinal, Chi-Square is used. We use pandas crosstab function to create frequency tables which we then passed into SciPy's chi2_contingency function to retrieve the p value. We used Chi-Square tests mostly for multiple select items which were considered as a series of Yes/No answers based on whether the respondent selected that option or not. This meant we were often conducting repeated testing, which necessitated the use of a Bonferroni corrected alpha value.

Results

We break the results into two sections. In the first section, we directly build on previous work by replicating prior work at scale and adding nuances to these findings. In the second section, we use the statistical methods described above to explore relationships between various dimensions of student's GenAI perspectives and use.

Descriptive Results

Looking at general usage data (Figure 1), the majority of learners are using AI products. Most use it weekly, though a large number of learners also use it monthly. These results show us that while AI product use is widespread, many learners are not relying on it for everyday usage.

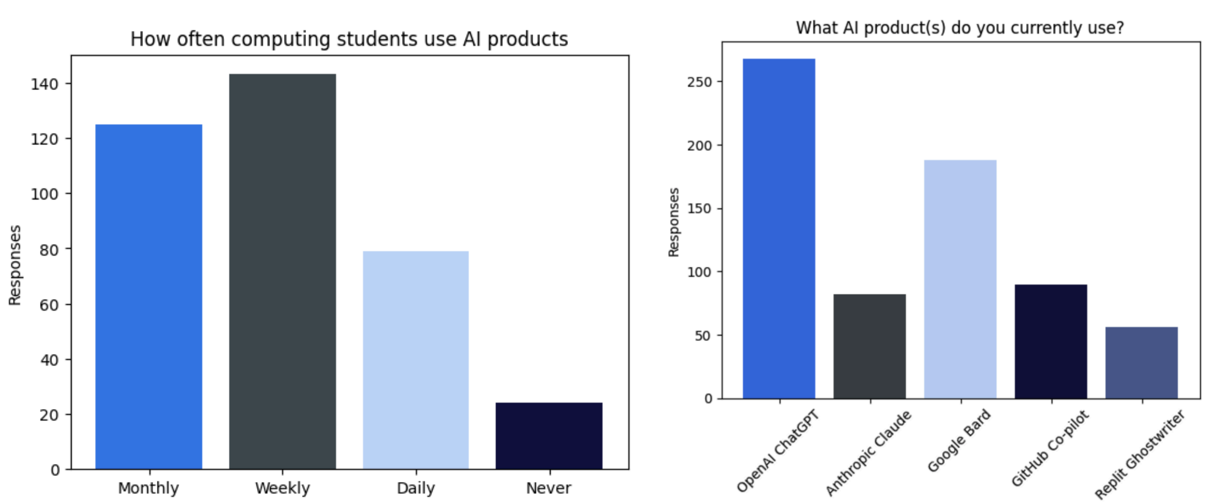

Figure 1: How often computing students use AI products and which AI products they use.

As shown in Figure 1, learners use various AI products, but OpenAI ChatGPT and Google Bard received the vast majority of responses. Those two tools received 456 responses, while the other three choices, Anthropic Claude, GitHub Co-pilot, and Replit Ghostwriter received 248 all together. This question allowed participants to select as many as applied, as many other questions do, so response numbers vary throughout. This is particularly interesting because both ChatGPT and Bard are standalone, chat-like tools, whereas the integrated tools, like Co-pilot and Ghostwriter, were much less popular. ChatGPT and Bard's popularity might indicate a desire for creativity and freedom to use GenAI in a variety of ways and in different subject areas.

In Figure 2 we see learners reported using GenAI pretty evenly for personal, professional, and educational use. This potentially explains the preference for free-standing tools, as respondents seem to be using GenAI in a variety of settings and so a programming-based tool might not always be applicable to their use cases.

Cost might be another consideration participants made when choosing some tools over others. While many learners do not pay anything for their GenAI products (Figure 2), the vast majority of our respondents do. In fact, there are almost just as many learners who pay nothing as learners who pay up to $10/month. While most of the products we asked about are free, some, like ChatGPT, include paid, premium versions users can upgrade to. Learners may feel that some tools offer more value for price over others and might be more willing to pay for such tools.

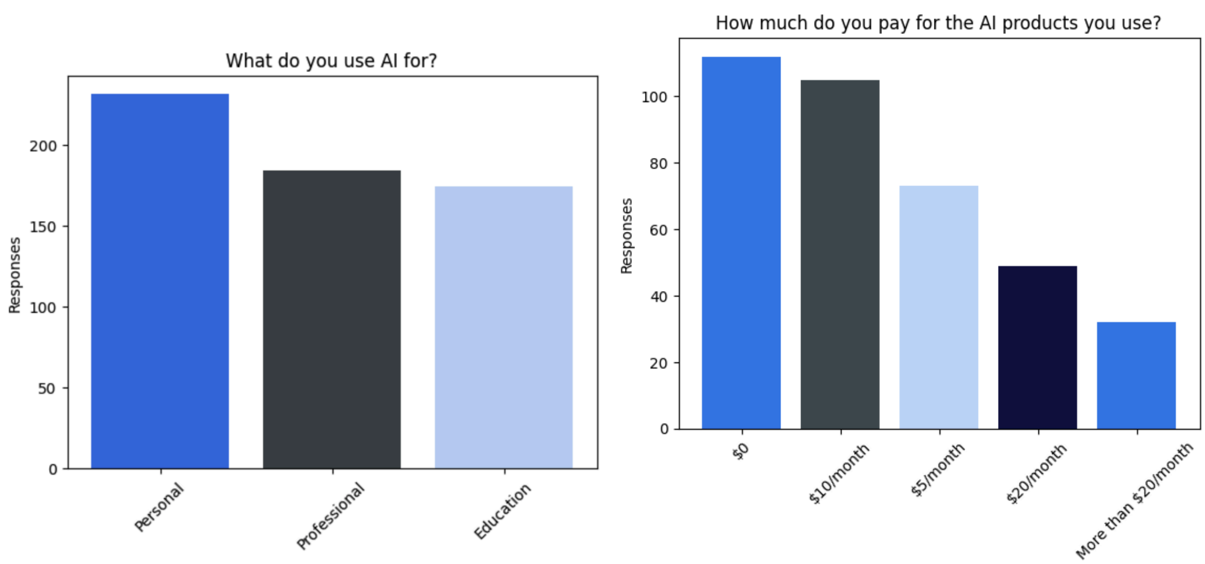

Figure 2: What students use AI for and how much they pay for it.

When examining Figure 3, it is clear that adoption of GenAI tools was highest when they were first released, in November 2022. Though there are some spikes, like around September 2023, adoption rates never matched those at the time of release.

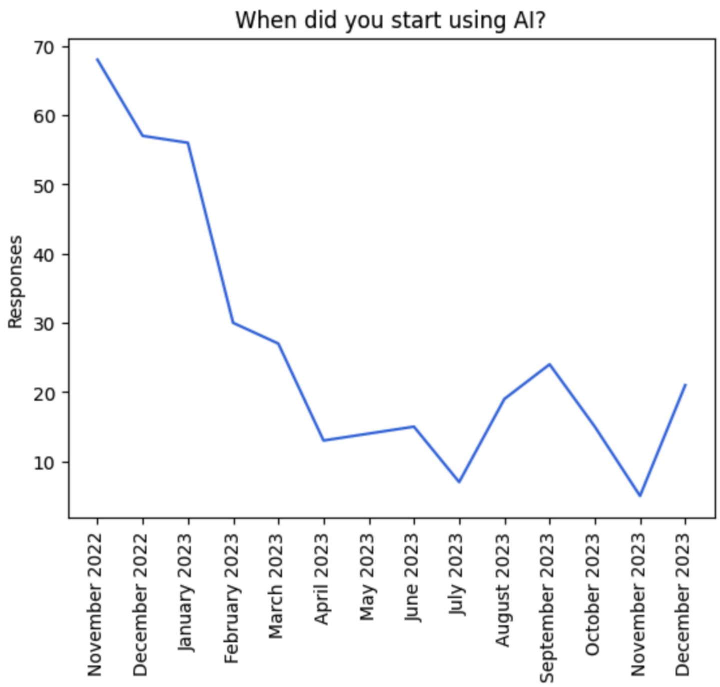

Figure 3: When students started using AI tools.

We asked what types of courses our respondents used GenAI in. Figure 4 shows AI is used in English courses more than twice as often as in other courses, though almost every other subject received about 100 responses indicating widespread usage.

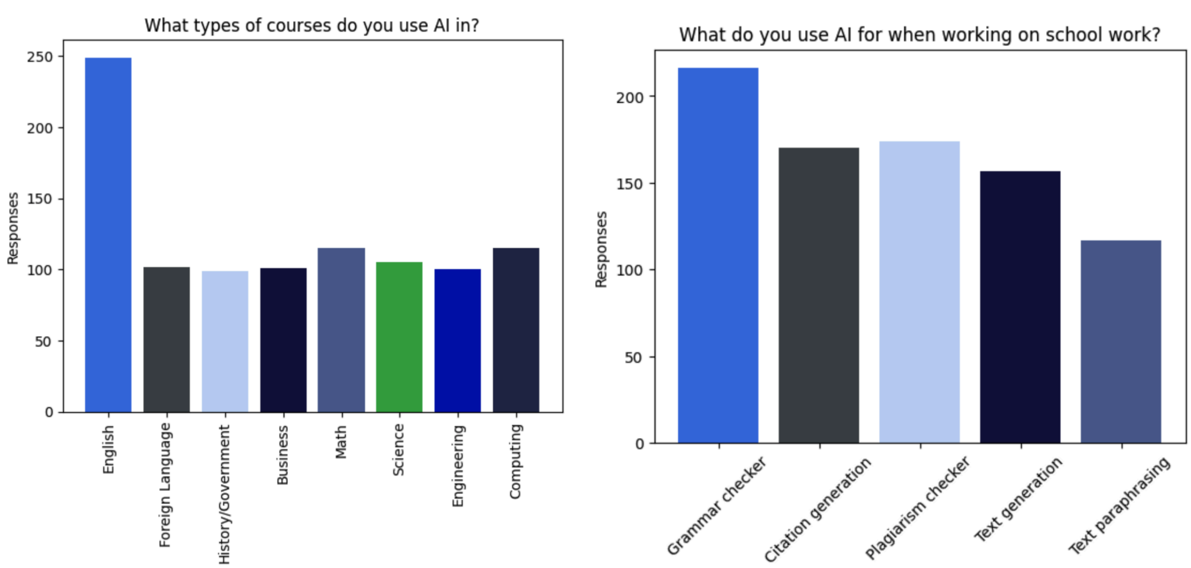

Figure 4: What courses is AI used in and how is AI used in general schoolwork.

Learners report using GenAI in lots of different ways when completing school work (Figure 4). Here, using a grammar checker received the most responses, which aligns nicely with both computing students using GenAI in English courses and the different usage possibilities of tools like ChatGPT and Bard. The popularity of other categories of schoolwork also indicates the widespread usage of these tools.

Respondents indicated they were aware of many different uses of GenAI in their computing courses (Figure 5). However, knowledge of understanding/summarizing home or project prompts received the most responses. This result suggests that project or assignment prompts might be poorly written or overly verbose to the point that students struggle to even understand them. Additionally, considering this usage of GenAI tools does not register as one that might break academic integrity rules, students could be the most aware of it because they either understand themselves or have been told by an instructor that this is an accepted way to use Gen AI for school.

We then asked users if they used GenAI for the same areas as the awareness question and found a similar response pattern (Figure 5). Most items received more responses indicating awareness than usage.

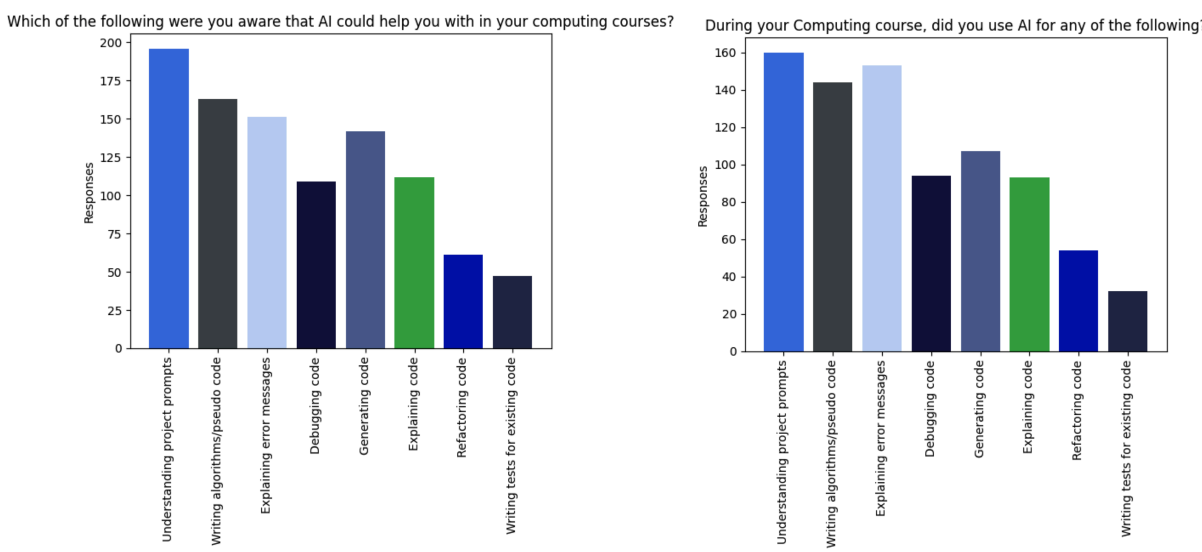

Figure 5: Students' awareness and actual usage of different GenAI uses in computing courses.

In Figure 6, responses show that when grouped together, more participants reported net time saved than time wasted, however, almost just as many responses that indicated GenAI saved noticeable amounts of time reported that they wasted time using GenAI.

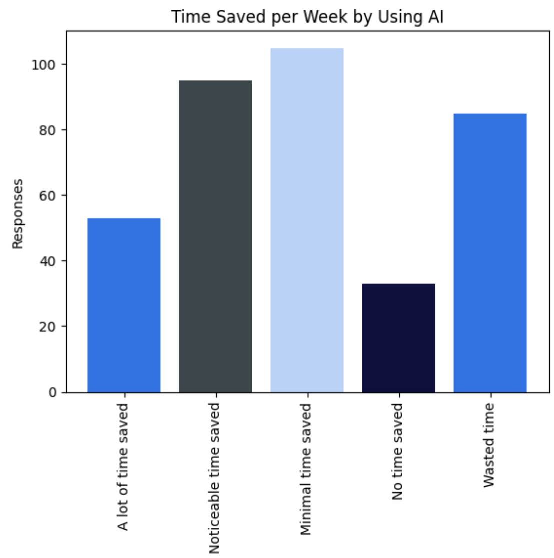

Figure 6: Computing students self-reported time savings when using AI

We asked participants to rate their perceptions of the helpfulness of GenAI tools first, and responses followed a bell curve, from rarely helpful to often helpful (Figure 7). Most participants felt GenAI tools were sometimes or frequently helpful, which aligns well with most respondents saving some amount of time. However, this curve also shows the variability in helpfulness, which again could be attributed to the user rather than the tool.

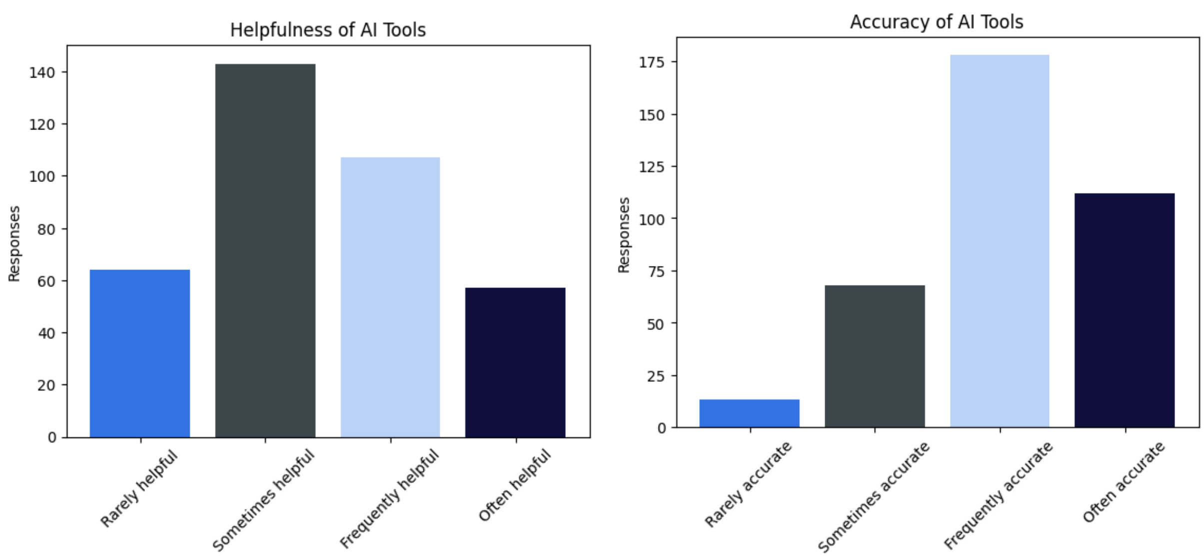

Figure 7: Students' perceptions of AI helpfulness and accuracy.

Then, there is a somewhat strange discrepancy between the previous results and perceived accuracy (Figure 7). Respondents' answers skewed heavily towards the tools being frequently or often accurate. We propose that some users have learned how to immediately ask a good question, one that will get them the right answer right away. However, it might take other users a longer time to get to that question that yields the answer they are looking for, and they would have reached the answer on their own with the time it took to get to that golden question. It could also be that some tools are more efficient and help the user ask the right questions better than others.

When looking at why students choose not to use GenAI for schoolwork (Figure 8), it firstly becomes clear that most people are aware of and know how to use these tools, and secondly that intrinsic motivators, like preferring to do your own work or not trusting the tools' accuracy are deeply important. That being said, these tools being forbidden by institutions also factors heavily into decision-making, showing the varying levels of influence on learners.

We examined how learners ensure the academic integrity of their work when they use GenAI tools to help in computing coursework. The results in Figure 8 suggest that students seem to be comfortable submitting so long as they can personally check and edit their work. While a substantial amount of responses show learners do not even consider academic integrity, the majority indicated that they read over everything to make sure they understand it, edit what gets generated, or only use AI for small pieces. These results show that learners seem to be making responsible decisions and attempting to maintain academic integrity when using GenAI to help them out with their work.

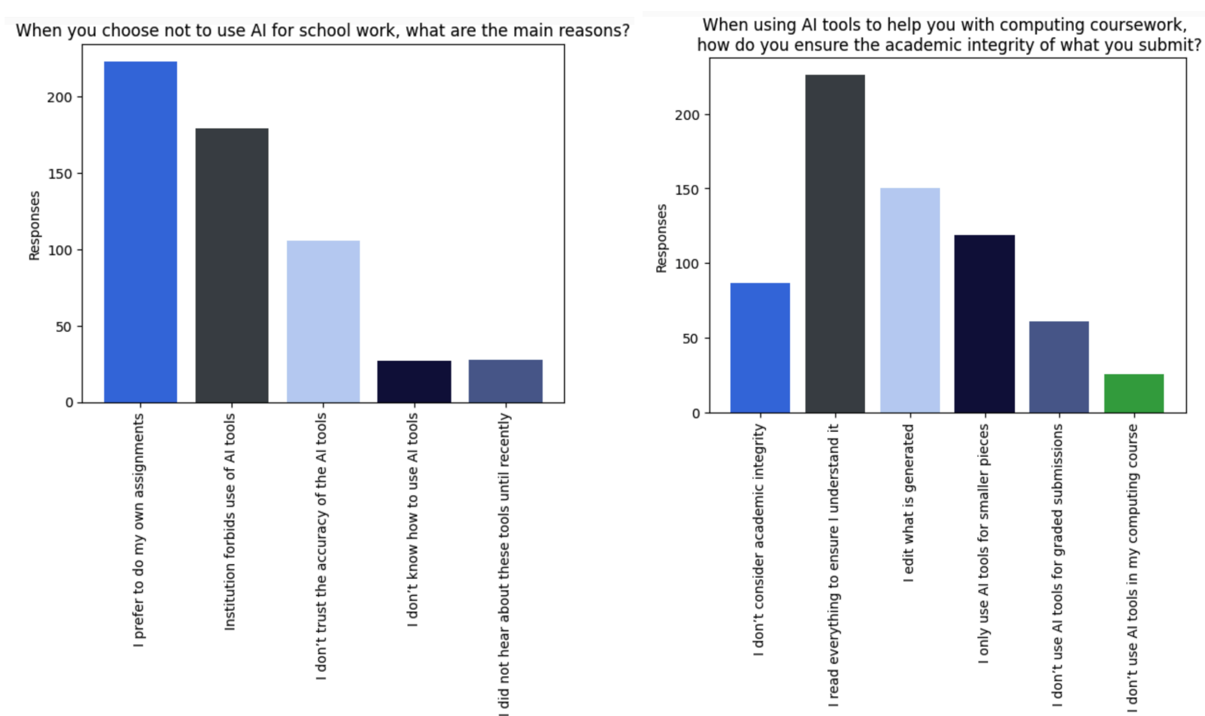

Figure 8: Students' perceptions about their use of AI in schoolwork, including academic integrity.

We asked if GenAI affects academically dishonest actions at all, and the vast majority of respondents were split between GenAI increasing or enabling dishonest actions and GenAI not changing the amount of dishonest actions (Figure 9). This split could be attributed to some students feeling that academic dishonesty was already rampant, given Chegg and other answer sites, so GenAI is merely replacing these answer sites, while other learners believe GenAI to be adding to the problem, enabling further dishonesty.

We then asked how GenAI affected the detection of academically dishonest actions and found predictable responses (Figure 9). Most learners feel GenAI makes it harder to detect dishonest actions, though many still feel GenAI does not affect detection.

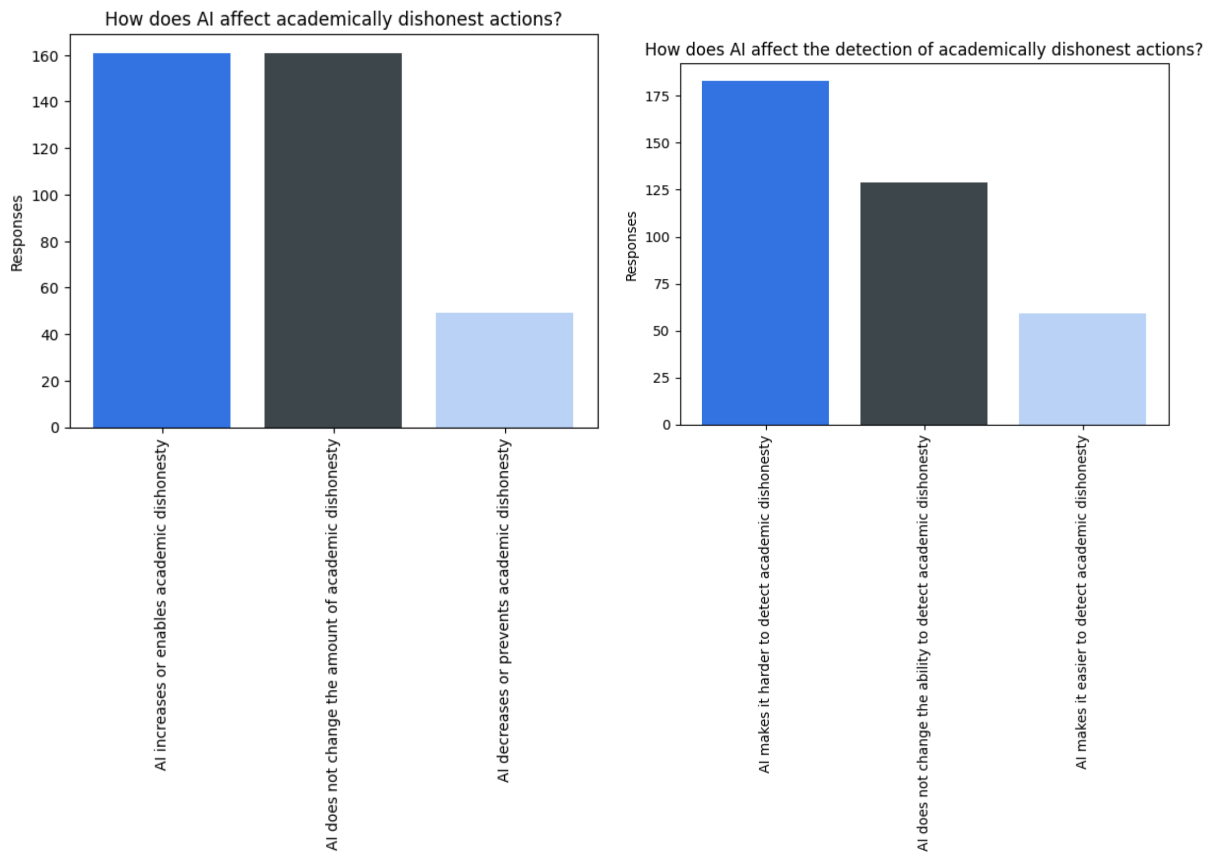

Figure 9: Students' perceptions about AI's influence on academic integrity.

We asked how their computing courses framed the use of GenAI tools, and found that most courses seem to be accepting of GenAI usage (Figure 10). Most respondents were taught how GenAI tools work, taught how to use them effectively, or asked to cite these tools when used. These responses indicate that instructors or institutions have begun to accept GenAI and are hoping to at least help students use GenAI in responsible, effective ways. That being said, a large portion of learners indicated these tools were forbidden or were not mentioned at all. When comparing this to usage responses, however, it seems clear that, forbidden or not, learners are using these tools to aid in coursework.

Finally, we asked participants if they were worried about GenAI replacing jobs, and the majority were concerned on some level (Figure 10). Interestingly, there was a slight drop off from minimally to somewhat concerned, but responses increased for very concerned, showing that participants have strong opinions on this topic.

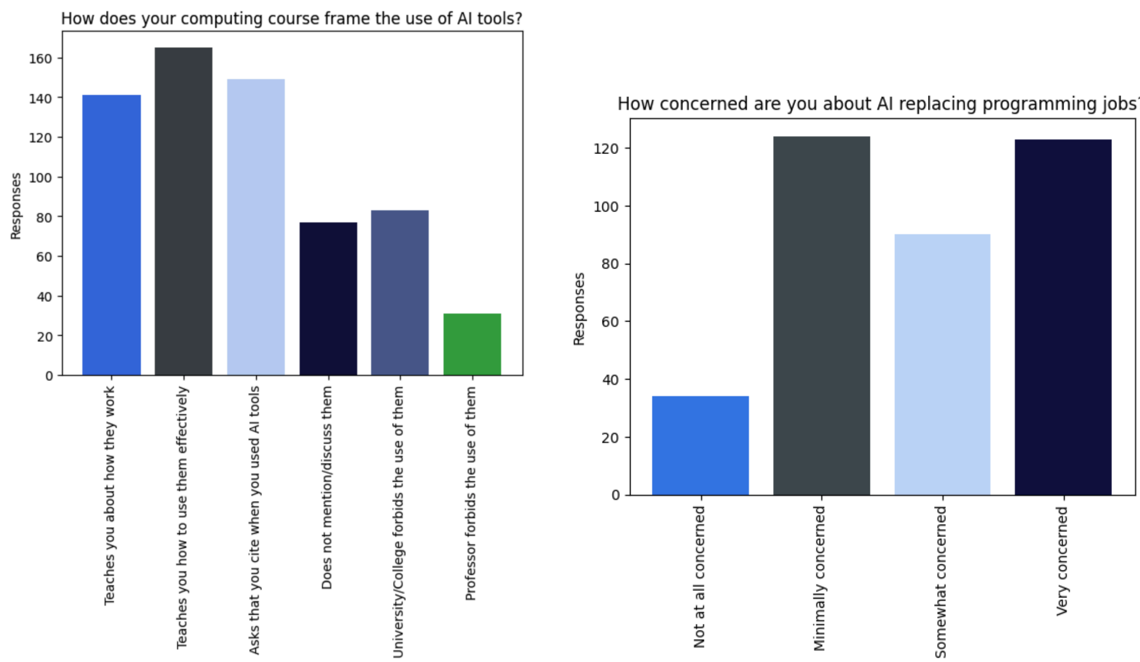

Figure 10: Students' experiences of AI in courses and perceptions of AI in industry.

Statistical Results

To examine the different ways students are using AI and how they view the output of GenAI tools in terms of time savings, helpfulness and accuracy, we performed a number of chi-squares and Spearman correlations to evaluate any relationship present within the data. We conducted Bonferroni tests to adjust the alpha levels throughout the statistical analyses.

Chi-square tests were run repeatedly to examine the relationship between how students generally used GenAI tools (e.g. Grammar checker, citation generation) and how helpful they perceived the tools to be. Because 5 chi-squares were conducted, the adjusted alpha was .05/5=0.01. Significant correlations were found between helpfulness and text paraphrasing ($p = .0002$), and helpfulness and text generation ($p = 6.63 \times 10^{-7}$). These findings suggest that the GenAI tools being used might be most adept at generating or paraphrasing texts.

Chi squares were also run to examine the relationship between accuracy and general use task, with an adjusted alpha of .05/5=0.01. A significant correlation was found between accuracy and citation generation ($p = .002$). The discrepancy between significant relationships relating to helpfulness and accuracy is intriguing, as it suggests that just because a tool is accurate does not mean it is necessarily helpful. When considering these results contextually, text generation and paraphrasing are more subjective tasks so they might be helpful, but accuracy could be more difficult to conceptualize, whereas citation generation is an objective task, with a clear distinction between an accurate and inaccurate answer.

Finally, we examined the relationship between time savings and general use tasks and found three significant relationships at the adjusted alpha level of .05/5=.01. There were significant correlations between time savings and text paraphrasing ($p = 5.26 \times 10^{-5}$), time savings and grammar checkers ($p = .001$), and time savings and text generation ($p = 7.31 \times 10^{-12}$). Text paraphrasing and text generation both had significant relationships with time savings and helpfulness, suggesting that students might struggle the most in these areas or again, that the tools are most adept in these areas.

We then examined the relationships between helpfulness, accuracy, and time savings with GenAI use in computing courses (CC) specifically. Each test was 8 times, and so the adjusted alpha level used was .006. There were two significant relationships between CC GenAI use and helpfulness, being helpfulness and understanding/summarizing homework/project prompts ($p = .001$), and helpfulness and writing tests for existing code ($p = .0004$).

We then examined relationships between CC GenAI use and accuracy and found no significant relationships.

Finally, there was one significant relationship between CC GenAI use and time savings, being between time savings and explaining code ($p = 6.13 \times 10^{-9}$). The significant relationships here seem to mostly surround interpretation tasks, like understanding project prompts or explaining code, suggesting that students might struggle in those areas more than others. Understanding/summarizing homework's relationship with helpfulness is especially interesting, as it suggests that students might need assistance even before starting a project.

Next, we examined how usage rates relate to accuracy and helpfulness by conducting chi-squares (Figure 11).

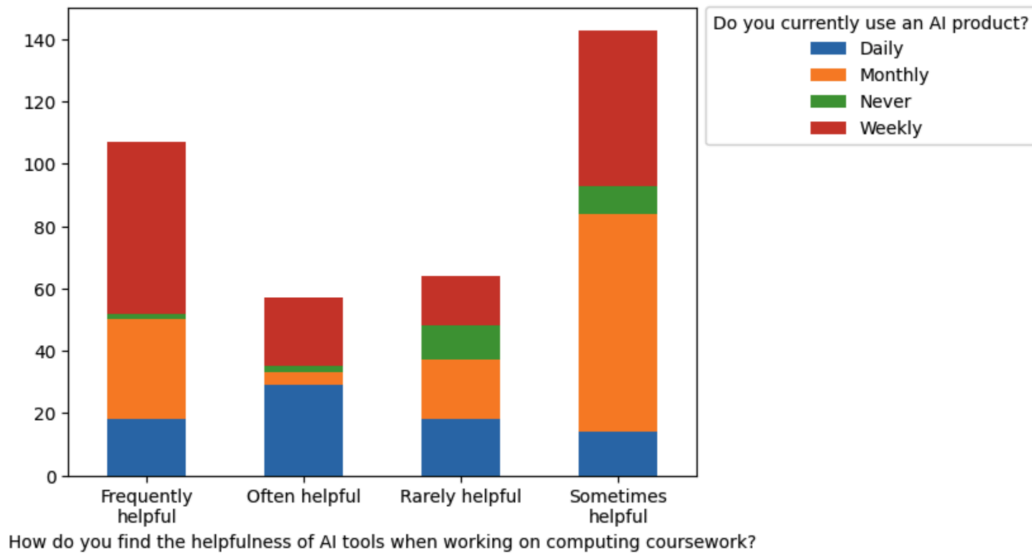

Figure 11: GenAI usage rates cross-cut with perceived helpfulness.

A statistically significant, weak correlation was found between computing students' self-reported usage rates of GenAI tools (i.e. Never, Monthly, Weekly, Daily) and students' perception of the GenAI tool's helpfulness ($p = 4.26 \times 10^{-8}$ and $\varrho = 0.278$). The effect size of this relationship was small ($d = -0.357$). It is interesting that there is not a stronger relationship between perceived usefulness and amount of use, indicating that other factors such as novelty, curiosity, and accepted learning curve may cause students to persevere with GenAI tools despite lackluster results.

A non-significant relationship was found between computing students' self-reported usage rates of GenAI tools (i.e. Never, Monthly, Weekly, Daily) and students' perception of the GenAI tool's accuracy ($p = 0.170$, $\varrho = 0.0715$, $d = 0.554$). This finding suggests there is no relationship between a tools' accuracy and how often computing students are using that tool. We would expect that as perceived accuracy increases, so too would usage rates, and the existing discrepancy is troubling as it suggests that students might use these tools no matter how accurate they believe them to be.

To examine how students understand the shifts of academic integrity in computing education with the introduction of generative AI tools, we again performed a number of chi-square analyses, adjusting alpha levels through Bonferroni corrections. Firstly, we examined the relationship between why students choose not to use GenAI and their general use task, with an adjusted alpha level of .05/24 = .002. We found significant relationships between preferring to do their own assignments and grammar checkers ($p = 8.88 \times 10^{-6}$), between institutions forbidding the use of GenAI and citation generation ($p = .001$), and between not trusting the accuracy of GenAI tools and text paraphrasing ($p = .0005$). It is worth noting that these uses are all non-blatantly academically dishonest. Additionally, preferring to do their own work was the most popular response given by participants, and it is unsurprising that it would have a significant relationship with grammar checkers, which require original work to already be completed.

We also examined the relationship between how students ensure academic integrity when using GenAI and how students use GenAI in computing courses and adjusted the alpha level, .05/48=0.001. There was a significant relationship between not considering academic integrity when using GenAI tools and understanding/summarizing homework/project prompts ($p = 7.69 \times 10^{-5}$). This relationship is not surprising, as using GenAI to understand a homework assignment would probably not violate academic integrity rules. Next, we evaluated the relationships between how use of GenAI is framed to students and how students ensure academic integrity, adjusting the alpha level .05/36=0.001. There were significant relationships between not considering academic integrity and being taught how GenAI tools work ($p = 3.34 \times 10^{-5}$), between reading everything to ensure they agree with and understand what they've received from GenAI and being how to use GenAI tools effectively ($p = 8.56 \times 10^{-7}$), and only using the tool for creating or debugging smaller pieces which they then assemble and the university or college forbidding the use of GenAI tools ($p = 7.16 \times 10^{-5}$). These results suggest that students who are taught to use GenAI have different views on academic integrity than those who are not.

We then examined how framing is related to students' perceptions of how GenAI affects academically dishonest actions, adjusting the alpha level, .05/6=0.008, and found a significant relationship between understandings of academic dishonesty and being taught how GenAI tools work ($p = .0002$). We also found one significant relationship between framing and how GenAI impacts the detection of academically dishonest actions, adjusted alpha .05/6=0.008, again between being taught how they work and perceptions of detection ($p = .0007$). When considering these results with the previous results, it becomes clear that teaching students how they work is an important factor in how they understand academic integrity and dishonesty, as those who are taught have significantly different perceptions of academic dishonesty than those who are not.

Finally, we examined how general use tasks relate to ensuring academic integrity, with an adjusted alpha level .05/30=0.002. We found three significant relationships, between reading everything to ensure they agree with and understand it and citation generation ($p = .0002$), between editing or tweaking what is generated and plagiarism checker ($p = .0006$), and only using the tool for creating or debugging smaller pieces which they then assemble and plagiarism checker ($p = 4.00 \times 10^{-6}$).

We examined relationships between the month of adoption and how GenAI was used in computing courses and found no significant relationships. However, there was a significant relationship between the timeline and general use task, with an adjusted alpha level .05/5=0.01, between month of adoption and text paraphrasing ($p=.006$). These findings suggest that when learners started using GenAI is not related to how robustly they use the tools or for what tasks they complete with them, perhaps subverting the assumption that users would gradually increase or decrease what they use the tools for as time goes along.

We also examined how cost related to usage rate (Figure 12), performing a spearman correlation test.

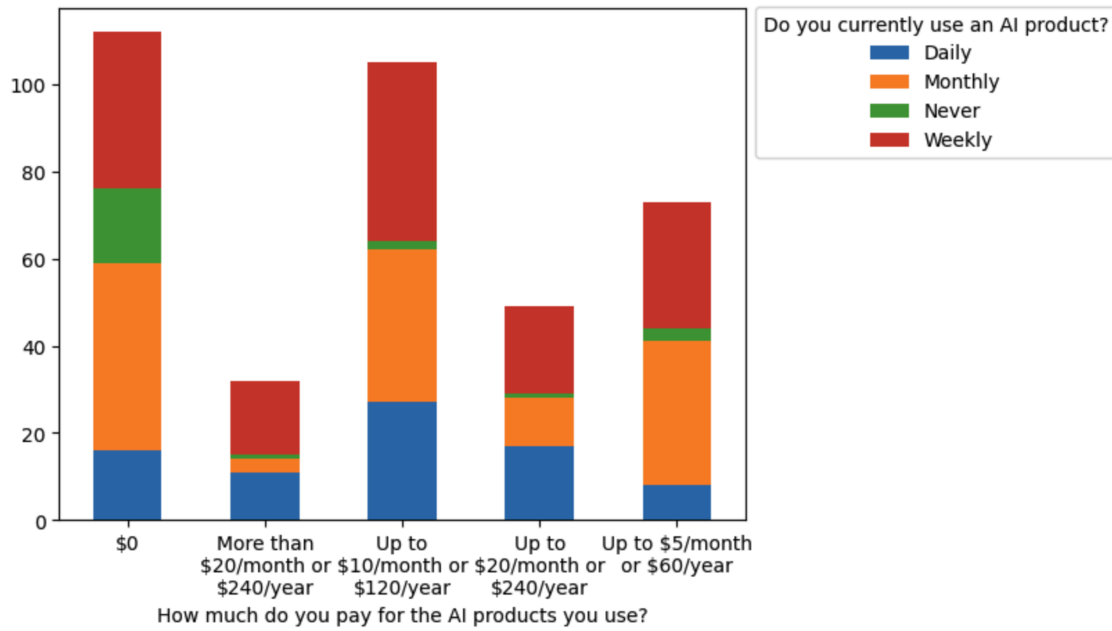

Figure 12: GenAI usage rates cross-cut with cost.

Next, we examined how cost relates to perceived accuracy (Figure 13). A statistically significant, weak correlation was the reported cost of the GenAI tool and students' perceptions of accuracy ($p = .045$ and $\varrho = 0.104$). The effect size of this relationship was large ($d = 0.800$). We might expect a more expensive option to be more accurate, resulting in a stronger relationship between these two variables. The relative weakness yet strong effect size of this relationship suggests that spending more money on a GenAI tool, like through upgrades, might not largely impact the value of results it will generate. Practically, this relationship also suggests that the skill of the user might be more impactful than the tool in the value of results.

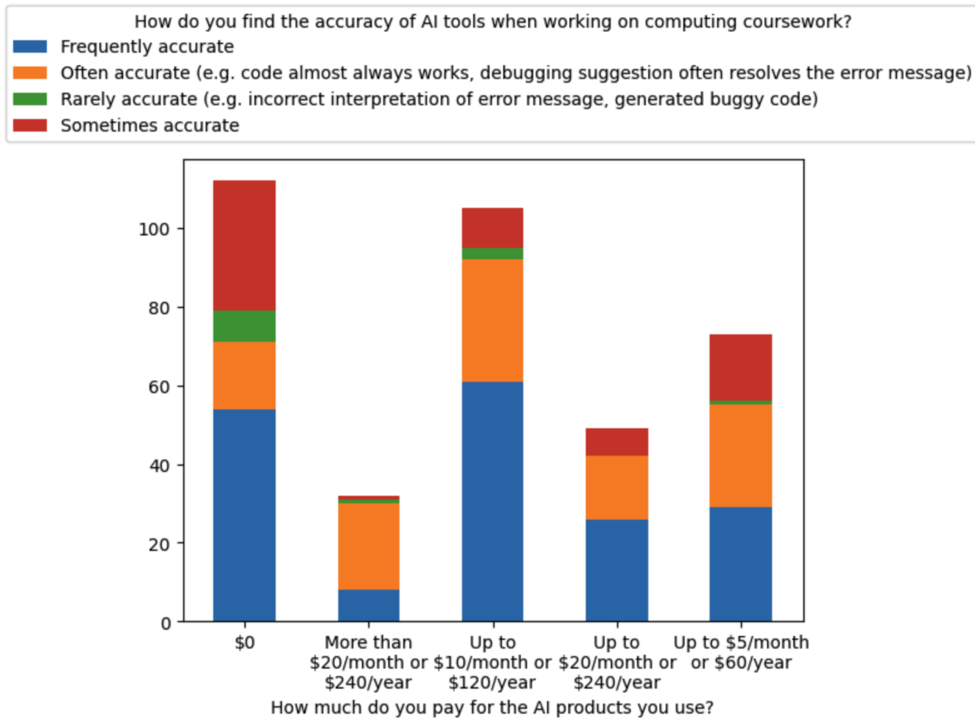

Figure 13: Cost for GenAI tools cross-cut with perceived accuracy.

Next, we examined the relationship between cost and perceived helpfulness (Figure 14). A statistically significant, weak correlation was the reported cost of the GenAI tool and students' perceptions of accuracy ($p = .0003$ and $\varrho = 0.187$). The effect size of this relationship was small ($d = 0.005$). Similarly to the cost x accuracy findings, a stronger relationship between helpfulness and cost might be expected compared to our findings. The relative weakness, both in terms of correlation coefficient and effect size, suggests that cost and helpfulness are minimally related to each other, again indicating that the skill of the user might be more impactful on the tool's helpfulness than how much the user pays for it.

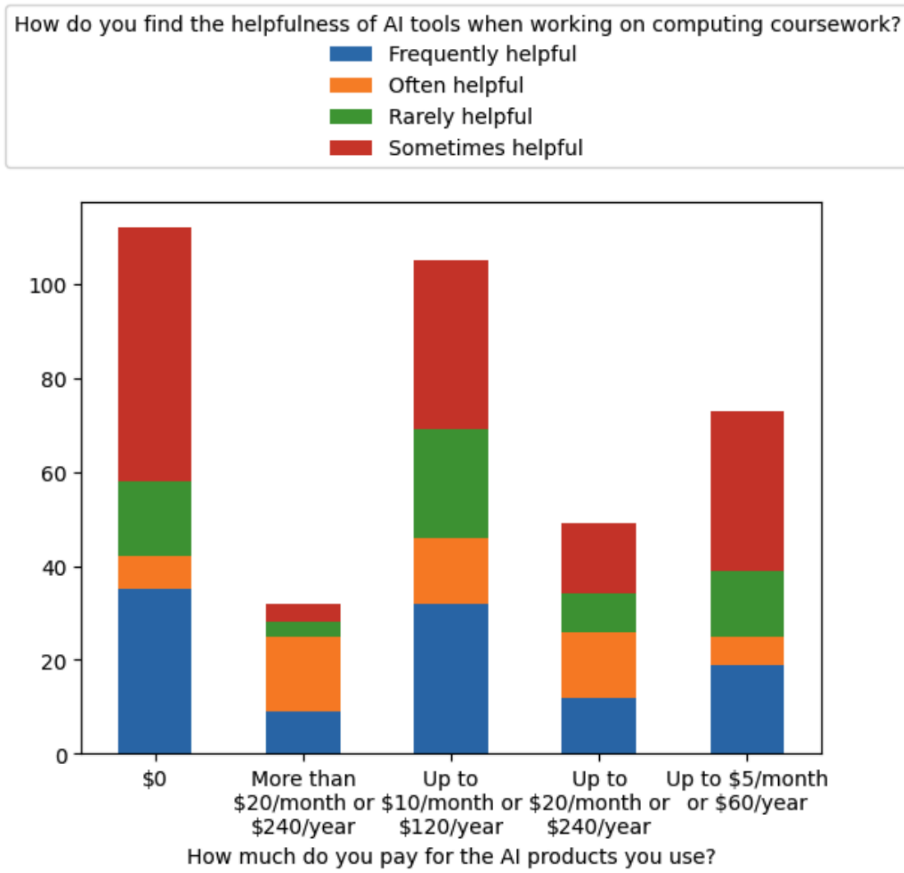

Figure 14: Cost for GenAI tools cross cut with perceived helpfulness.

Finally, to examine how students' views on GenAI tools (i.e. time savings, helpfulness, accuracy, academic integrity concerns, and job replacement concerns) affect the amount of time and depth of usage, we performed a number of chi-squares, adjusting alpha levels with Bonferonni corrections. First, we analyzed how perceived helpfulness relates to a students' major, with an adjusted alpha of .05/6=0.008, and found one significant relationship between helpfulness and SE/CS/CE or other Computing majors ($p = .0009$). This result suggests that computing majors have different perceptions of the helpfulness of GenAI than other majors, which may indicate that computing students make better, or worse, use of GenAI than students in other majors do. Next, we looked at major and accuracy, with an adjusted alpha .05/6=0.008, and found one significant relationship between perceived accuracy and Social Sciences ($p = .001$). Next, we examined job concerns with majors, with an adjusted alpha level .05/6=0.008. We found two significant results between job replacement concern and SE/CS/CE or other Computing majors ($p = 3.15 \times 10^{-7}$) and between job replacement concern and Social Science majors ($p = 2.10 \times 10^{-5}$). Considering the question surrounds programming jobs being replaced, it is unsurprising that computing majors would have a significant relationship, as they might be closest to the issue. However, social sciences majors having a significant difference is slightly surprising but could be attributed to an over representation in our population. Finally, we performed a number of chi-squares to examine how time of adoption related to job concern, accuracy, helpfulness, and time savings. We found significant relationships, with an alpha of .05, between adoption timeline and accuracy ($p = .003$) and adoption timeline and time savings ($p = .0004$). These results suggest that time of adoption relates to how accurate the GenAI tool is perceived to be and how much time is saved by using the GenAI tool. Both of these implications support the assumption that the user might be more important than the tool, as the more time they spend with the tool, the better they might be at quickly asking helpful questions.

Discussion

Results show that, on a general level, the use of GenAI products is widespread. Many participants indicated that they use GenAI for various tasks, both in computing and non-computing courses and in personal and professional applications, which aligns clearly with and replicates research conducted by Amoozadeh and team [10]. Most participants indicated using OpenAI ChatGPT or Google Bard monthly or weekly, and most paid nothing for their tool. Interestingly, ChatGPT and Bard are standalone chat tools that could be used in an array of applications, as opposed to tools like Git-Hub Co-Pilot, which center around coding, and were much less popular in our population. ChatGPT's popularity is unsurprising and aligns with the literature, like that done by Poole College of Management Business Analytics Initiative [4], which predicts these tools will become as ubiquitous in education as traditional software like Microsoft Excel in the near future.

While many participants paid nothing for their tools, the majority reported paying sizable amounts of money for these tools. However, we found only weak correlations between cost and perceived accuracy, perceived helpfulness, and time savings, which suggests that the tool's capabilities as they relate to cost most likely make little difference in terms of actual usefulness to the user. Specifically, we found a weak correlation between cost and perceived accuracy with a strong effect size, which emphasizes the practical applications of this result. This weakness can be interpreted in multiple ways. The first being that upgrades, which usually cost extra money, make only a small difference in terms of the tools' perceived accuracy, and they are probably not worth the extra expense. A second interpretation might be that those people who perceive the tool as more accurate are slightly more likely to pay for an upgrade.

Throughout our results, there were discrepancies between accuracy and helpfulness, which might suggest that an individual user's expertise is more important than the tool in having a productive and efficient experience. We also found that adoption time significantly relates to perceived accuracy and time savings, which, when taken contextually with other results, furthers the implication that user experience and skill might matter more than tool choice. Importantly, the results suggest that no matter the amount of money a beginner computing student spends on a

GenAI tool, it might not be particularly useful to them until they have gained more of their own skills and experience to harness GenAI in a helpful way.

Though use of GenAI products by students is widespread, it is mostly used in unproblematic ways, like interpreting homework assignments and error messages. When looking specifically at how certain tasks related to accuracy, helpfulness, or time savings, we found that tasks like text paraphrasing, text generation, grammar checking, understanding homework prompts, and writing tests for existing code were most frequently truly useful to students. Interestingly, except for text generation, most of these uses seem unproblematic in terms of academic integrity, and more problematic uses like generating code were less popular and less useful. Considering findings that human graders struggle to distinguish between work produced by humans and LLMs, this finding is promising (e.g. [1], [9], [11]). However, it is worth noting that the sample was self selected, and so might skew towards academic curiosity and integrity.

Looking more specifically at academic integrity, perhaps the most interesting finding surrounds whether or not students are taught to use GenAI. We found that variable to be the only one consistently significantly related to students' perceptions of how GenAI impacts academic integrity and how they ensure their own academic integrity when using GenAI. Considering again the difficulty in distinguishing between human work and work done by an LLM, these results suggest that perhaps the only way out is through: the most effective way to impact student's perceptions and behaviors might be to teach them to use GenAI in efficient and honest ways, rather than banning GenAI altogether ([1], [9], [11]).

Though this research presents intriguing and hopeful results, it must be noted that the sample is not representative and participants chose to participate, which hinders the ability to make broad generalizations. We hope that going forward, more representative samples will be used for research on GenAI, perhaps replicating the kind of questions asked in this study. However, this research still paints a broad picture of the use patterns and perceptions of GenAI which may be helpful in illuminating the nuance and adoption trends of GenAI to educators and administrators. Additionally, though the sample is not representative, it still replicates Amoozadeh and team's finding that over half of students were using GenAI [10].

Further, as the sample over-represented older students who are more likely interested in GenAI, as they chose to participate, these results might be particularly applicable to non-traditional learning environments, like upskilling courses or tech boot camps. Administrators of these educational spaces might consider these findings more heavily and might consider enacting clear and adequate GenAI rules and advice for students more quickly than say university administrators.

Future research will certainly be conducted on students' use of GenAI, and should specifically consider evaluating how users' skills progress over time in relation to helpfulness and accuracy. Researchers could build upon our findings to investigate the interplay between helpfulness and accuracy, and look more indepth at what makes a tool accurate and what makes it helpful. Investigating these variables could help us understand students' thought processes and decision making around GenAI and equip educators with more nuanced data. Additionally, different survey designs, like more immediate surveys after using a GenAI tool or longitudinal studies following students over a longer period of time, will be crucial to unpacking their perceptions of GenAI and tracking skill progression. Future research could also examine more closely professors', administrators', and students' understandings of academic integrity, and how they are impacted by GenAI. Findings of such an investigation might shape how GenAI is integrated in class settings, helping educators to approach GenAI in an informed way.

Conclusion

The majority of students who utilize GenAI pay for access to such tools, use GenAI for tasks like paraphrasing text or summarizing homework assignments instead of generating code, and prefer to do their own work when choosing not to use GenAI. Discrepancies throughout our findings suggest that an individual's skill level and knowledge of using GenAI may impact the usefulness of their experience more than the tool chosen or price. Additionally, teaching students how to use GenAI seems the most related to their considerations and perceptions of academic integrity. Our findings suggest that while students are already using GenAI tools, teaching them to use them effectively might benefit not only their productivity, but also increase academic integrity.

References

[1] D. R. E. Cotton, P. A. Cotton, and J. R. Shipway, "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT," *Innovations in Education and Teaching International*, vol. 61, no. 2, pp. 228–239, Mar. 2023, doi: https://doi.org/10.1080/14703297.2023.2190148.

[2] C. M. University, "Examples of possible academic integrity policies that address student use of generative AI tools," *www.cmu.edu*. https://www.cmu.edu/teaching/technology/aitools/academicintegrity/index.html

[3] "AI & Academic Integrity | Center for Teaching Innovation," *teaching.cornell.edu*. https://teaching.cornell.edu/generative-artificial-intelligence/ai-academic-integrity

[4] "Generative AI: Integrating Tools into Teaching and Research," *Business Analytics Initiative.* https://bai.poole.ncsu.edu/event/generative-ai-integrating-tools-into-teaching-and-research/ (accessed May 01, 2024).

[5] M. A. Cardona, R. J. Rodriguez, and K. Ishmael, "Artificial Intelligence and the Future of Teaching and Learning Insights and Recommendations," *Office of Educational Technology*, May 2023. Available: https://www2.ed.gov/documents/ai-report/ai-report.pdf

[6] C. Dede, "What is Academic Integrity in the Era of Generative Artificial intelligence?," *Silver Lining for Learning*, Aug. 06, 2023. https://silverliningforlearning.org/what-is-academic-integrity-in-the-era-of-generative-artificial-intelligence/

[7] G. Chami, "Artificial intelligence and academic integrity: striking a balance," *THE Campus Learn, Share, Connect*, Oct. 23, 2023. https://www.timeshighereducation.com/campus/artificial-intelligence-and-academic-integrity-striking-balance

[8] UNESCO, "Guidance for generative AI in education and research," *Unesco.org*, 2023. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

[9] C. Zastudil, M. Rogalska, C. Kapp, J. Vaughn, & S. MacNeil, "Generative ai in computing education: Perspectives of students and instructors", in *2023 IEEE Frontiers in Education Conference (FIE)*: IEEE, 2023. pp. 1-9.

[10] M. Amoozadeh, D. Daniels, S. Chen, D. Nam, A. Kumar, M. Hilton, M. A. Alipour, and S. S. Ragavan, "Towards Characterizing Trust in Generative Artificial Intelligence among Students", in *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*: ACM, 2023. pp. 67-73.

[11] D. Baidoo-Anu, and L. Owusu Ansah. "Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning." *Journal of AI* 7, no. 1: 2023. pp. 52-62.

[12] N. Agarwal, V. Kumar, A. Raman, and A. Karkare. "A Bug's New Life: Creating Refute Questions from Filtered CS1 Student Code Snapshots", in *Proceedings of the ACM Conference on Global Computing Education Vol 1*: ACM, 2023. pp. 7-14.