ACE up your Sleeve: An Analysis of Student Generative AI Usage in an Engineering Statics Course

Jacklyn Wyszynski, University of Pittsburgh
David Adam DeFrancisis, University of Pittsburgh
David Pabst, University of Pittsburgh
Mr. Lee Allen Dosse, University of Pittsburgh

Lee A. Dosse is a PhD student working with the Engineering Education Research Center at the University of Pittsburgh.

Dr. Matthew M. Barry, University of Pittsburgh

Dr. Barry is a Visiting Associate Professor in the Mechanical Engineering and Materials Science department at the University of Pittsburgh. His research interests include space power propulsion systems and Engineering Education.

ACE up your Sleeve: An Analysis of Student Generative AI Usage in an Engineering Statics Course

Abstract

Rapid technological advancements, including the emergence of computer-aided design and simulation, have had a significant impact on the engineering industry. This, in turn, extends to engineering education, demonstrating a similar influential effect. The latest development to have such reverberations is the launch of a generative artificial intelligence (AI) chatbot known as ChatGPT. ChatGPT utilizes a large language model (LLM) that trains the platform to understand and generate human-like responses. This LLM comprises numerous neural networks trained using the vast amount of information available online, including research papers. As this new technology is widely accessible to students, the questions that arise regarding its role in education are almost always related to academic integrity. ChatGPT can answer questions, compose and revise papers (like this one), and complete collegiate course evaluations. Though ChatGPT can be misused, like many tools, when used as intended, it can assist students in their educational efforts. That is, instead of asking AI to answer homework prompts, one can inquire for clarification to further their knowledge about nuanced engineering phenomena.

This paper attempts to understand students' perspectives and engagement with generative AI in a university-level introductory engineering course. Self-reported student data were collected through a survey and six focus group interviews, which were then thematically coded to elucidate any common trends. Unsurprisingly, students admitted to copying and pasting assignment questions into AI but soon discovered that the answers were, more often than not, unreliable. This feedback forced them to use it more productively. Interestingly, many students viewed AI as an around-the-clock tutor that was conveniently always available, asking it to clarify complex topics or provide definitions or equations needed for assignments. A substantial number of responses indicated that students found AI helpful when preparing for exams, as it helped formulate study guides by synthesizing student-inputted equation sheets or created practice problems that mimicked exam questions.

Introduction

Artificial intelligence (AI) emerged in computer science in the 1950s and has since undergone significant development. There are many distinct flavors of AI, such as fuzzy logic, genetic algorithms, knowledge-based systems (rule-based), inductive learning (automatic knowledge acquisition rule-based), and neural networks, to name a few [1]. Many of these AI programs are rule-based (i.e., follow "if-then" logic based on a knowledge base) with the provision to update

the knowledge base in an automatic and/or informed manner. Others, such as artificial neural networks, learn patterns from training datasets.

Developed by OpenAI, ChatGPT was released to the public in November 2022, garnering widespread attention and fanfare. In less than a week, over one million users had accessed GPT, and by the end of January 2023, that number had surpassed 100 million [2]. ChatGPT is a large language model (LLM) chatbot capable of simulating human conversation [3]. It is built upon a generative pre-trained (GPT) LLM model that has undergone a series of significant improvements in capability. The LLM being utilized in ChatGPT initially started in 2018 with GPT-1, the first LLM based on transformer deep learning architecture [4]. ChatGPT was not released to the public, however, until the creation of GPT-3.5, and is currently built on GPT-4. The GPT models received supervised learning and human feedback as part of their training. They were taught using information found both online and offline [5], resulting in a chatbot with excellent abilities to understand human prompts and respond in a very conversational fashion [6]. These advanced LLMs have enabled ChatGPT to possess numerous advanced natural language processing capabilities. Besides simulating basic conversation, it is capable of summarizing a section of text or a complete work, reviewing a given topic, editing or generating text, debugging and creating computer code, transcribing an audio file, and more [7].

As this technology continues to mature, ChatGPT has begun to appear in various settings, with potential applications in fields such as medicine [8], law [9], and economics [10]. Already, ChatGPT acts as a disruptor in higher education [11]. Experts and educators alike disagree on whether, and how, this new technology should be used. There are many ways in which the software can be used in academia. Its text generation capabilities enable the answering of questions of varying degrees of complexity when prompted by users, as well as creating assessments to test students' knowledge. ChatGPT can provide students with a summary of any assigned reading and help explain any sections they may have found challenging to understand [11]. The program can edit and provide suggestions about students' written work, helping them become better writers. Initial studies have indicated that ChatGPT can increase student curiosity and improve their ability to ask meaningful questions [12]. It has demonstrated a positive effect on several aspects of education, including learning interest and achievement, knowledge retention, and explicit reasoning [13]. It is believed that ChatGPT, if implemented correctly, can enhance student creativity and critical thinking [14].

On the contrary, the software can also be used to violate academic integrity policies. Its text-generation capabilities are especially alarming for writing assignments, including research papers and dissertations [15]. Indeed, a survey sent to college students in the US found 53% used ChatGPT to write papers and 48% even used it during exams [16].

With many policies and decisions being made in higher educational settings, it is crucial to understand how students use generative AI [17, 18]. Specifically, it is essential to understand how generative AI works in disciplines that are more complex and nuanced, such as engineering. To this end, the researchers aim to answer the question, "How are engineering students using AI?" We hypothesize that, for the majority of students, the platform is not being used for dishonest purposes (i.e., attempting to obtain answers from the AI platform), but rather as a learning aid and tool. To accomplish this, the researchers conducted a study on how students use generative AI platforms in an introductory Statics and Mechanics of Materials course. By analyzing survey and

focus group data on student usage of AI within the course, the researchers aim to provide a clearer understanding of how this technology, like past technologies (e.g., CAD), can assist engineers and students.

Methodology

This study focuses on students self-reporting their use of generative AI in an introductory Statics and Mechanics of Materials course. One hundred ninety-three students were enrolled in two sections of this course (n = 131 and 62, respectively). The students had access to a version of ChatGPT, known as Top Hat ACE, within their textbook, Statics and Mechanics of Materials: An Example-based Approach [19]. This text, hosted on Top Hat's platform, has its own version of ChatGPT-4 embedded directly into the browser. ACE uses the text for fine-tuning its responses to student-generated prompts. Students can be prompted by ACE's generic queries, such as "Show examples," "Explain a concept," "Question help," and "Quiz me." Students can also ask ACE any question via a text-entry dialogue box. Upon obtaining university policy approval, this feature was made available to students for the last quarter of the 16-week semester. However, alternative generative AI platforms were freely available throughout the course.

Two methods were used to collect data on student usage: a survey and focus group interviews. Each of these methods of data collection will be discussed below.

Survey

A dynamic survey was created using the university-approved platform Qualtrics. The survey questions and logical flow are provided in Fig. 1. The survey consisted of a total of 13 possible questions. The first question asked students if they had used generative AI within the course, followed by which AI platform they had used. Thereafter, the students could indicate how they used generative AI within this course: for "Clarification," "Assistance," "Review," or "Other" purposes. Two questions would appear for each of the usages indicated. The first question (i.e., Q4/Q6/Q8/Q10) in this series would ask the students to further elaborate on how they used generative AI for providing clarification, assistance, review, or other purposes, respectively. The second question asked about the frequency of use for seeking "Clarification," "Assistance," "Review," or "Other" purposes (i.e., Q5/Q7/Q9/Q11, respectively). Thereafter, the students were asked if generative AI had impacted their understanding and were allowed to elaborate. The survey's final question asked the students if they would participate in a focus group interview.

Due to the branching nature of the survey, students could answer anywhere from one (they did not use generative AI and were thanked for their time) to 13 questions. If the students used generative AI for only one purpose in the course (e.g., clarification), they would be prompted with seven questions; for two purposes, they would be prompted with nine questions; for three purposes, they would be prompted with 11 questions; and for four purposes, they would be prompted with all 13 questions.

Focus Groups

Six focus group interviews were conducted to gain a deeper understanding of how students utilized generative AI in the course (n = 9, 10, 7, 8, 5, and 9). The focus groups were moderated by two researchers at any given time. Within the semi-structured 50-minute period, the

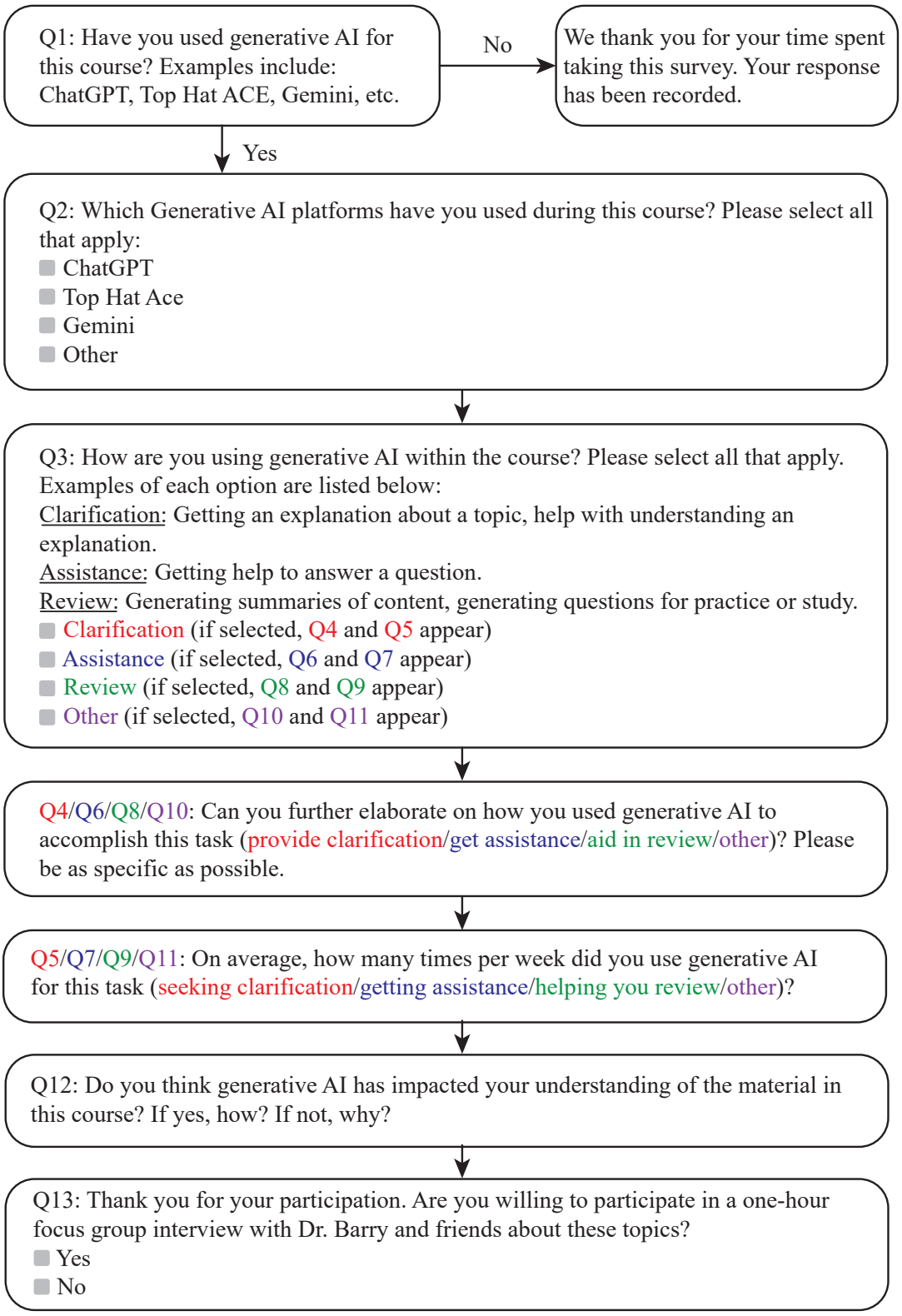

Figure 1: Qualtrics survey flow-chart.

moderators asked the students pre-determined questions, allowing them to answer the questions and provide additional comments or feedback. The interviews were recorded using Panopto and were transcribed using Microsoft Word; all transcriptions were subsequently verified by the researchers. During the interviews, no visual data was collected. The questions, which were provided to each student, asked during the interviews are listed below.

1) How did you find yourself using generative AI, and why?
2) Did you prefer using generative AI to get help instead of seeking help via office hours, the course email, or peers?
3) What do you think are the pros and cons of using generative AI in this course?
4) (For students who used ACE and ChatGPT): Which generative AI do you prefer? Top Hat ACE, which is limited to the textbook, or ChatGPT, which has access to the whole internet?

Data Analysis

Survey and focus group responses were thematically coded according to the methods and procedures outlined by Creswell et al. [20]. The researchers implemented an inductive coding scheme throughout this process [21]. That is, they only constructed their themes after first reading through all the student responses. A summary of this process is provided in the following; if the reader is familiar with this data analysis technique, they are encouraged to proceed to the next section.

Two researchers independently reviewed the survey results and focus group transcripts and individually assigned themes (i.e., codes) to the responses. These themes captured recurring ideas and sentiments. After coding each item (i.e., questions in the survey or focus group interview), the researchers would meet to discuss their identified codes. During this discussion, they decided on a final set of codes to apply to the survey data and focus group transcripts. Then, each researcher would independently reanalyze the data, applying (or binning a response into) the agreed-upon codes. Upon re-coding the data, the researchers would meet again to compare which code(s) were applied to which responses. In the case of disagreement regarding code assignment, a third researcher would arbitrate and ultimately decide which code(s) were applicable.

The researchers determined the inter-rater reliability percentage (i.e., the percentage of agreement on assigned final codes between the researchers) for each survey question and focus group when applicable. The percentages for the survey questions can be found in Tab. 1, and the focus group questions resulted in values of 97.35%, 97.62%, 100%, and 100% for Q1, Q2, Q3, and Q4, respectively.

Table 1: Student Survey Participation by Question

Question	n	Course Response Rate	Usable Survey Response Rate	Inter-Rater Reliability
Q1	158	81.87%	100%	N/A
Q2	94	48.70%	100%	N/A
Q3	94	48.70%	100%	N/A
Q4	80	41.45%	98.75%	100%
Q5	80	41.45%	98.75%	N/A
Q6	52	26.94%	94.23%	95.74%
Q7	52	26.94%	94.23%	N/A
Q8	38	19.69%	97.37%	100%
Q9	38	19.69%	100%	N/A
Q10	3	1.55%	33.33%	N/A
Q11	3	1.55%	100%	N/A
Q12	90	46.63%	97.78%	98.88%

Results & Discussion

The researchers jointly assessed survey and focus group responses to gain a holistic understanding of student usage and opinions regarding the implementation of generative AI in engineering coursework. Fundamentally, the researchers were curious about what types of generative AI students predominantly utilized within the course. Of the 158 student responses to the survey, 107 (67.72%) students confirmed that they used generative AI within the course, and 51 students (32.28%) denied using generative AI (Q1).

Students who responded affirmatively were prompted to list the generative AI platform(s) utilized (Q2). The results are shown in Fig. 2, with 74.47% of student respondents opting to use ChatGPT and 69.15% of students reporting the use of ACE. Only 5.32% of student responses were noted using Gemini (Google AI), Copilot (Bing AI), and Perplexity, which were all uniquely noted by single responses (1.06%). The course and survey response rates are detailed in Tab. 1. The course response rate was calculated by dividing the number of student responses for a specific question by the total number of students enrolled (n = 193).

Furthermore, the usable survey response rate was calculated by dividing the number of valid responses by the sample size (n) for a particular question. Oftentimes, invalid responses were simply questions left unanswered by students. The researchers also deemed responses that did not directly address the survey question as invalid. For example, when asked how many times a week they utilized generative AI for assistance, a student replied, “dgjhgd.” The varying response rate and trend of decreasing n values observed in the table may be attributed to the dynamics of the questionnaire through the unique paths taken by students, or survey fatigue.

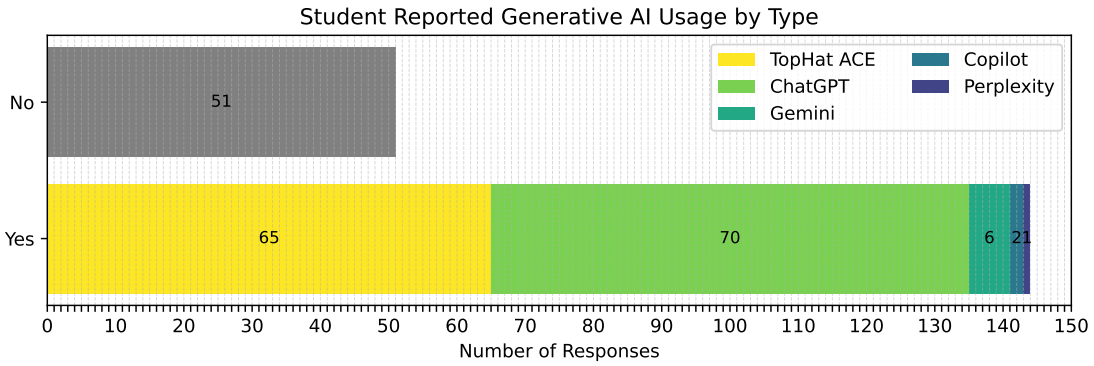

Figure 2: Number of students reporting using AI, and which AI they used.

The survey respondents were then asked how they utilize generative AI (Q3), with the options "Clarification," "Assistance," "Review," and "Other" presented with defined examples of each usage. Students were capable of making as many selections as they felt applied. The examples given for each option can be seen in Fig. 1. Out of the 94 students who used generative AI during the course, 85.11% reported doing so for "Clarification." "Assistance" followed, appearing in 56.38% of responses, with "Review" trailing in 40.43% of responses. Few students reported using generative AI for "Other" purposes, accounting for only 3.19% of responses. According to Fig. 3, while students primarily express using generative AI one to four times per week for "Clarification" (though this range is omnipresent in all usage types), the most significant percentage of responses about usage greater than five times per week fell within the "Assistance" category.

One student in the survey described using ChatGPT almost daily, "like a 24/7 tutor" to better understand the concepts taught in class. Even students who did not use generative AI as often noted increased usage before tests, especially when using AI for "Review." A student writes, "I use this task no more than once a week, since I use it to help prepare for the exams."

Student availability to complete coursework appeared to influence AI usage; some students opted to use AI depending on their weekly schedule. Survey responses suggest periodic usage spikes when time management becomes difficult during the semester, for instance, during exams and term projects.

Referring to Fig. 1, based on their selected usage, students were asked to detail how generative AI was used in the task's completion. The coding schemes developed for each response are shown in Tab. 2. As indicated, the upper section of the table presents the categorical descriptors generated for usage questions (Q4, Q6, Q8, and Q12), which consist of various tasks that students associated with their utilization of generative AI. The researchers categorized every response to these questions by applying the nine codes presented. The lower segment of the table contains codes exclusively associated with student opinions towards the helpfulness of generative AI in the context of the Statics course (Q12). The Q12 codes contain a "Yes" and "No" indication for the impact of AI on student learning, in addition to the previous usage codes associated with the student's response, and "Convenient" and "Unreliable" as modifiers to AI usage. The researchers assessed the Q12 responses using a total of twelve codes.

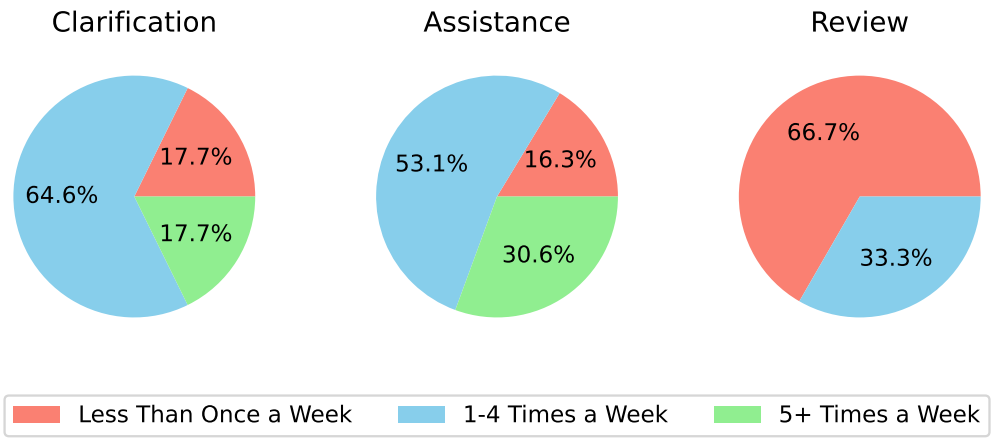

Figure 3: Student Generative AI usage frequency by usage type as reported by responses to Q5/Q7/Q9.

Table 2: Survey Coding Schemes by Question (Q4, Q6, Q8, Q12)

	CODE	CATEGORICAL DESCRIPTION
Q4, Q6, Q8, Q12	Conceptual Understanding (C)	Student asks generative AI for a combination of conceptual help to deepen understanding, to explain, define, clarify, or to rephrase with the intention of furthering comprehension.
	Methodology (M)	Student asks generative AI to determine a method for completing a Statics problem.
	Practice Problems (PB)	Student uses generative AI to generate course-related practice problems.
	Summarization (S)	Student uses generative AI to synthesize or compile information for review.
	Find Solution (FS)	Student directly asks the given coursework question to be answered by generative AI.
	Reference Textbook (RT)	Student utilizes generative AI in conjunction with the textbook to find formulas, equations, or definitions within the Statics textbook, or to further aid literary comprehension.
	Error Checking (ER)	Student uses generative AI to analyze student-completed work for accuracy or to check for possible mistakes.
	Study Content (SC)	Student uses generative AI to study or create material to support studying, including flashcards, review sheets, or study schedules.
	Unreliable (UR)	The student negatively comments that generative AI is not always correct, or fails to understand a problem.
Q12	Yes (Y)	Student thinks that generative AI has been helpful to their understanding of the course material.
	No (N)	Student thinks generative AI has not been helpful to their understanding of the course material, perhaps due to the belief that the course had sufficient material for understanding.
	Convenient (CN)	Student positively comments on the ability to immediately get assistance remotely due to generative AI.

No codes were generated for Q10 or Q11 due to the combined effect of low and inadequate response rates. The three responses collected for Q10 included two omitted responses where students revealed that they did not use generative AI for its intended purposes. One student admitted that they "asked Top Hat ACE to write a poem about almonds," and another student, "asked ACE if it was 'geeked'." The researchers were not geeked by these responses and consequently decided to disregard them in the coding process.

Figure 4 presents the codes generated from usage responses related to "Clarification," "Assistance," and "Review." Student usage for "Clarification" (Q4) showed a significant partiality for the category "Conceptual Understanding," denoted by (C), and appearing in 59.49% of responses. This percentage was determined by taking the 47 responses containing the (C) theme and dividing by the number of usable responses (79 total) for Q4. The researchers computed all additional percentages using this method, adjusting the denominator to account for different n values associated with the question. "Clarification" was also firmly supported by the "Methodology" descriptor, (M), accounting for 41.77% of responses. Subsequent descriptors were intermittently employed, though no other code was seen to occupy more than 10% of responses.

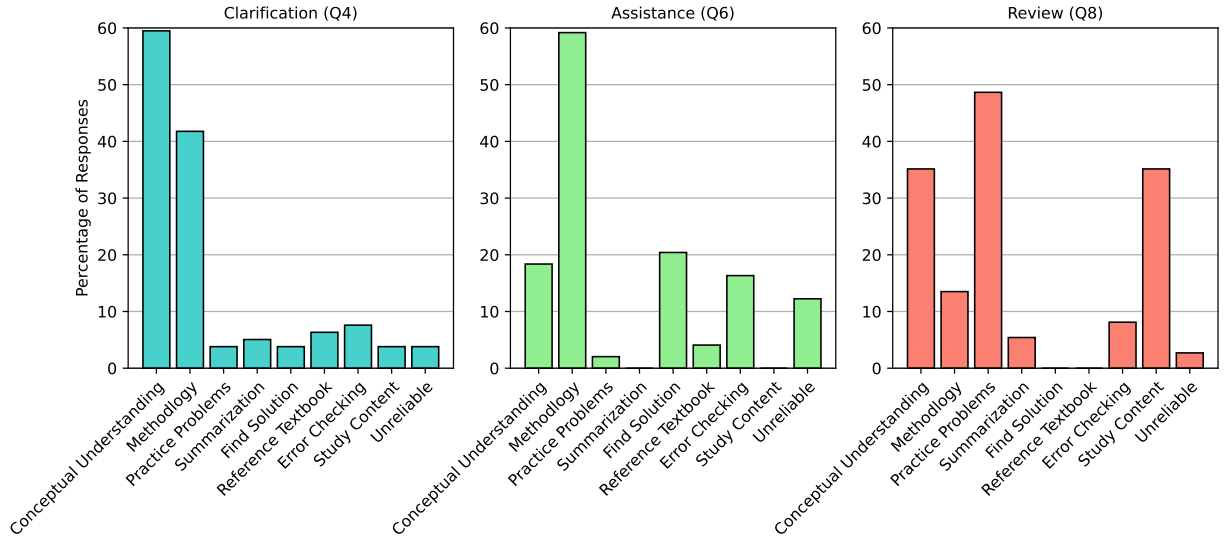

Figure 4: Categorical Descriptions by Survey Question (Q4, Q6, Q8)

Student usage for "Assistance" (Q6) painted a familiar scene, though "Methodology" prevailed with an appearance in 59.18% of responses, likely due to student usage for homework problems and classwork assistance. "Conceptual Understanding" was observed less frequently than "Find Solution" codes at 18.37% and 20.41%, respectively. "Error Checking" was observed in 16.33% of responses related to usages for assistance. Several students in the survey utilized generative AI as a tool to identify incorrect partial solutions or approaches to given quantitative problems. A student wrote:

> *I would ask about parts of homework problems that I was having a hard time figuring out on my own. Sometimes, I would ask it if I was getting a correct value or solving*

for a part of the question correctly so I knew if I was on the right track.

Students also noted that ChatGPT did not always provide accurate feedback on submitted work. Students primarily felt that while generative AI could, in some but not all cases, provide the correct framework for quantitative problem-solving, it failed when analyzing involved and complex problems. Generative AI software was exceedingly poor in solving and assisting in problems that contained multiple variables or required students to infer data through figures−problems that emulated realistic engineering applications and demanded more from students than simply "plugging and chugging." One student discussed this difference in accuracy with respect to a homework problem dependent on a figure, which included multiple measurements.

> *Just going back to ACE, those ...questions that have like a 3.5 km drill hole. I was using the 12.19 km long length of drill bit as the length for the torsion angle calculation, and I kept getting it wrong because that's the wrong length. I kind of fed the numbers I had and what I was using those numbers to do into ACE and it was not able to help me with that one, [it] was not even telling me that my length was wrong. It kinda just stated that it looks like I'm moving in the right direction, but here's the formula again. So well, I know that's the formula. That's not what I was confused about. It's which one of these variables do I have the wrong value for and it was not able to help me cross that line.*

A new trend emerges when analyzing students' use of generative AI for "Review" (Q8). The theme most often coded, "Practice Problems," was applied to 48.64% of responses, followed by "Study Content" and "Conceptual Understanding," which both appeared in 35.14% of responses. The majority of students used AI to fortify their understanding of course material in preparation for upcoming exams. One student stated, "I used AI to create study sets and review cards for me on Quizlet. It can create study sheets using input terms." Interestingly, one student combined the ideas of the three aforementioned codes, stating:

> *Particularly helpful for exam review and generating practice problems. I wasn't able to have it generate problems that included a figure, but useful for creating basic problems to ensure retention of broader concepts.*

Evidently, students are not exclusively using generative AI to violate academic integrity policies. As seen in the responses to Q8, generative AI can play a beneficial role when used appropriately as a study tool. However, in the face of more complexity, for example, with the implementation of figures, AI begins to struggle in the recreation or interpretation of visual information.

The final question of the survey investigated whether students felt that generative AI influenced their comprehension of course material. The researchers categorized the responses as either "Yes" or "No." The findings reveal that 78.41% of students reported that generative AI was impactful, while 21.59% disagreed. In other words, approximately four times as many students felt generative AI enhanced their understanding compared to those who did not. Furthermore, the researchers applied an additional 10 codes to the responses for deeper insight into the students' reasoning behind their answers. Figure 5 presents the frequency of descriptors appearing based on the student responses. The most frequent descriptor (C) was identified in 54.55% of student responses. The remaining descriptors listed in descending order of frequency of response, are as follows: (CN)

26.14%, (M) 13.64%, (UR) 13.64%, (SC) 3.41%, (ER) 1.14%, and (S) 1.14%.

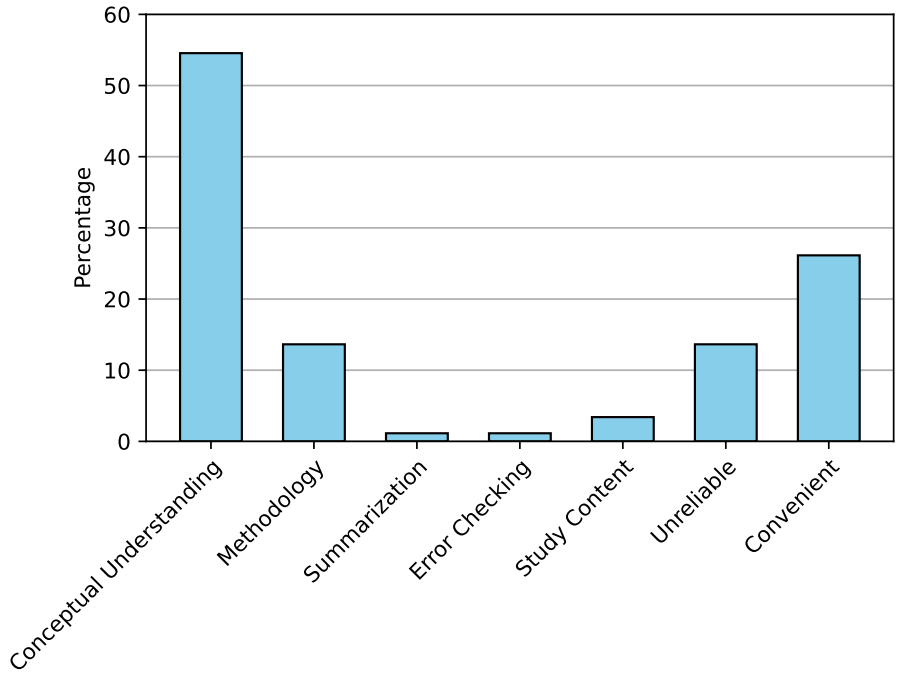

Figure 5: Categorical descriptions by survey question (Q12).

Many students advanced their understanding of nuanced concepts introduced in the course by utilizing generative AI. However, some students expressed that the use of AI was wholly unnecessary. One student wrote,

> *I largely did not use it because I don't see the point of it. Everything you need is in the textbook, and if you need help understanding something, there are approximately 40 hours of office hours every week. In my cynical eyes AI is just the latest dot-com bubble that every company is trying to get their share of cash out of. Why does a textbook need an AI? Everything you need to know is just in the textbook!*

Unquestionably, some students did not see the utility of implementing generative AI into course material. This disinclination, mated with the observed variance in AI-generated solutions, raised student doubts about the software's reliability. One student noted, "I don't really trust it as much as the book or TAs." and another stated, "...it struggled to give me reliable answers and contradicted the textbook in places." These sentiments reveal that students did not unanimously benefit from generative AI.

The first focus group question accentuated the similarities and differences between responses held in-person and through the survey. Similar to the survey, focus group responses were inductively coded, though not all questions were considered relevant to the major trends elucidated in this study. All responses, however, were considered to substantiate the discussions and add depth to the findings. The compiled focus group codes can be seen in Tab. 3.

Table 3: Focus Group Coding Schemes by Question (Q1, Q3, Q4)

			CODE	CATEGORICAL DESCRIPTION
Q1			Conceptual Understanding (C)	Student asks generative AI for a combination of conceptual help to deepen understanding, to explain, define, clarify, or to rephrase with the intention of further comprehension.
			Methodology (M)	Student asks generative AI to determine a method for completing a Statics problem.
			Practice Problems (PB)	Student uses generative AI to generate course-related practice problems.
			Summarization (S)	Student uses generative AI to synthesize or compile information for review.
			Find Solution (FS)	Student directly asks the given coursework question to be answered by generative AI.
			Reference Textbook (RT)	Student utilizes generative AI in conjunction with the textbook to find formulas, equations, or definitions within the Statics textbook, or to further aid literary comprehension.
			Error Checking (ER)	Student uses generative AI to analyze student completed work for accuracy or to check for possible mistakes.
			Unreliable (UR)	Student negatively comments that generative AI is not always correct, or fails to understand a problem.
			Convenient (CN)	Student positively comments on the ability to immediately get assistance remotely due to generative AI.
Q3	Pros		Convenient (CN)	Student positively comments on the ability to immediately get assistance remotely due to generative AI.
			Informative (I)	Student positively comments on their ability to gain information or guidance through the use of generative AI.
	Cons		Unreliable (UR)	Student negatively comments that generative AI is not always correct, or fails to understand a problem.
			Dependency (D)	Student negatively comments on an over-reliance of generative AI.
			Less Stimulating (LS)	Student negatively comments that using AI is less stimulating than in-person learning.
Q4			ChatGPT (CH)	Student prefers using ChatGPT.
			Top Hat ACE (A)	Student prefers using Top Hat ACE.
			Neutral/Neither (NE)	Student does not have a preference on which generative AI they use, or seldom uses generative AI.
			Other (O)	Student uses an alternative generative AI.
	Justification		Convenient (CN)	Student positively comments on the ability to immediately get assistance remotely due to generative AI.
			Unreliable (UR)	Student negatively comments that generative AI is incorrect a lot of the time, or fails to understand a problem.
			Informative (I)	Student positively comments on their ability to gain information or guidance through the use of generative AI.

Figure 6 illustrates the consistency in frequency of responses categorized under (C), differing by only 1.46%. Descriptor (M) appeared with a higher frequency in the focus group than survey data, as did (S), (RT), and (ER). There was a significant difference in the frequency of response of the (FS) descriptor, which was more commonly used in the focus groups. Unlike the survey, focus groups allowed students to elaborate extensively on their usage. These augmented explanations may account for the inflation of thematic codes observed in the focus group data. Another explanation may lie in the researcher's involvement in the focus group interviews. Immediately prior to initiating the focus groups, the researchers reassured students that their comments would have no bearing on academic performance or perceptions of the individual's character. A higher level of comfort was likely established throughout the focus groups due to their innate communal structure compared to the survey.

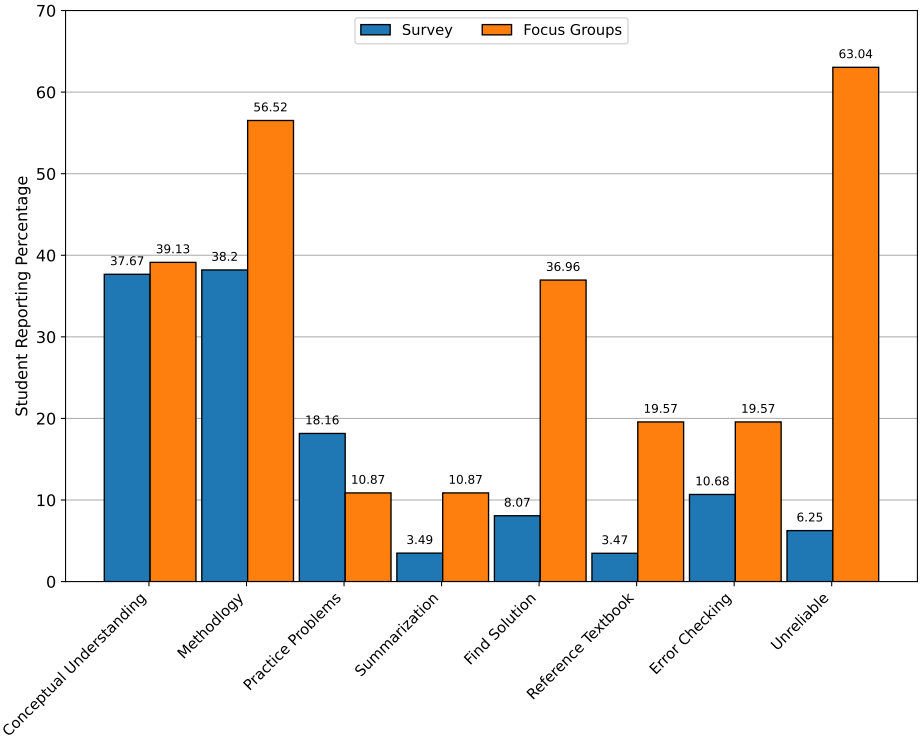

Figure 6: Overlapping survey and focus group (Q1) categorical descriptions colored blue and orange, respectively. Frequencies of survey codes were averaged across all responses to Q4, Q6, and Q8.

A similar trend is observed in the (UR) categorical descriptor. Students frequently emphasized the unreliability of generative AI, particularly when searching for a direct answer. However, after recognizing that generative AI was not a consistently accurate source for answers, students were compelled to use generative AI to benefit their methodological understanding of assigned problems. A student notes in the focus group:

> *Once it... was getting the answers wrong [...] I stopped just plugging in the straight problem and then asked step by step, can you help me this way?*

Aside from using generative AI for methodology as an alternative to finding a direct solution, students more notably showed a propensity to use AI solely to help them understand various problem-solving methods.

Figure 6 shows that the (M) categorical descriptor repeatedly appeared in the survey and focus group responses alike. One student remarks on AI's ability to relay steps for solving sample exam questions, using generative AI to prep for all of the course exams:

> *Same thing for the final, I gave the practice final and I asked the generative AI to explain each question, different ways of solving it, and how can I approach these types of questions, and it actually helped a lot.*

As seen in Fig. 7, students emphasized the convenience of generative AI in the focus group interviews, accounting 48.48% of responses. When asked to describe the pros and cons of generative AI, students benefited from the ability to gain time-saving assistance remotely (Q3). One student describes this behavior:

> *I think office hours are great, but I don't want to have to go to the old engineering building and get lost in that maze trying to find... some TA's office hours. But I know I could just plug it into ChatGPT, and [...] it's not gonna be the best response ever, but it'll help me get that basis of learning that I need.*

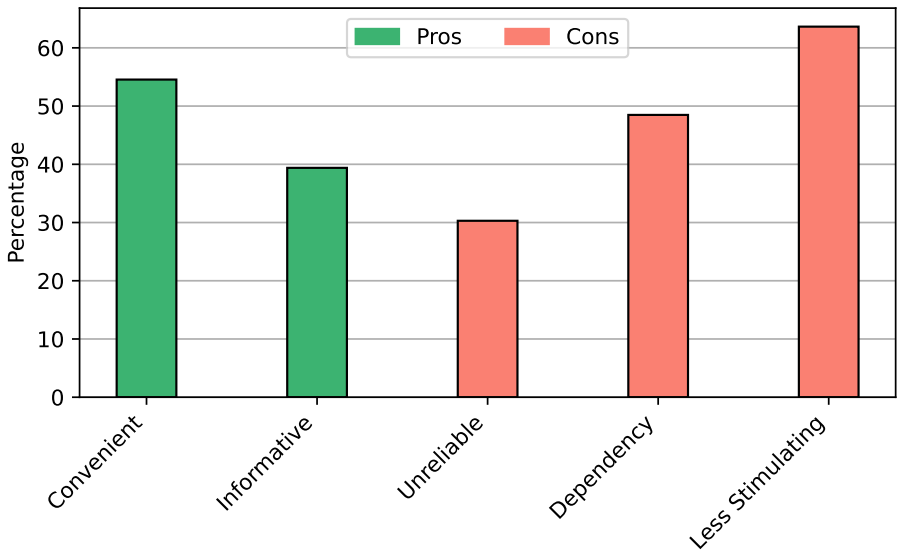

Figure 7: Categorical Descriptions of Student Opinions on Pros and Cons of Generative AI (Q3)

A key relationship emerged between student responses that were sorted with both the "Convenient" and "Dependency" codes. Students repeatedly expressed a sentiment that the risks of being dependent on generative AI were exacerbated by poor time management skills, a trend substantiated by an increase in usage frequency during periods of heavier workload during the semester. The convenience of AI makes dependency all the easier—why spend hours memorizing content when the answers can be found online in seconds?

However, generative AI is not a proxy for instruction; many students believed that generative AI still fell short of other forms of educational aid. 54.55% of students found generative AI to be unreliable, leading them to pursue either the course textbook or their peers for further assistance. Another prevailing view was that memorizing content through generative AI detracted from the learning experience. Many students underscored the value of personal initiative in the learning process. A student commented:

> *I feel like that active process of either going out and seeking someone or going to the textbook and writing things down is a lot better than just like being a zombie and just copy and pasting into different web browsers and just like looking at it that way.*

Another student says:

> *I think learning comes in steps. So you have to take small steps up in your ladder to fully understand a subject. And ChatGPT does not have a good enough understanding of where you are there [...] But in a sort of way, a teacher, say, in comparison to ChatGPT or AI in general, would know where you are in your course, could give you the next step instead of giving you ten steps ahead of where you are to finalize your learning of that subject. [...] because specifically ChatGPT has such a big subset of data, it has no clue what I'm learning in particular.*

It would appear AI lacks the ability to recognize when and where a student is struggling with content, merely based on their provided prompts. Without recognizing physical and verbal cues (i.e., humanity) during conversational turns, it cannot currently provide the nuanced instruction and assistance necessary in complex disciplines.

Conclusion

This study investigated how students used generative AI in an introductory engineering course. Through a survey and six focus group interviews, the researchers gathered extensive self-reported data on student usage of various AI platforms within the course. By categorizing the data into common themes, the study aimed to uncover both how students were using generative AI and their motivations. An in-depth analysis of the self-reported data revealed unexpected findings: student usage does not align with what university policy-makers believe AI models are used for in the classroom. Generative AI's availability to the public raises fears among policy writers at the university level. Many universities have continued to impose a ban on implementing generative AI in coursework based on the belief that students will only use it unethically. Decisions as significant as complete restriction should not be based on fear of the unknown or assumptions. At some universities, professors have implemented AI models in class for additional tutoring, and to teach students how to use AI for their future careers in industry. Like any other tool, there exists both pros and cons with generative AI. Hasty generalizations should not be made without considering potential benefits; AI requires further investigation beyond predetermined judgments on the software.

AI still clearly has its limitations. Students frequently commented on the unreliability of generative AI, specifically when attempting to use it to solve assigned problems. Since AI models are seemingly not yet capable of successfully completing complex engineering problems, students

turned to more productive uses of the AI software. These beneficial usages include helping with problem-solving methods, reminding students of formulas and definitions, and summarizing course material. Not only can AI aid with coursework, it was also noted to be particularly useful for exam preparation. Students often used AI to help them make practice problems resembling potential exam questions, or create study guides to review content.

The reported data highlights another advantage of convenience being a significant factor in students' use of generative AI. For those completing homework late at night, AI serves as a virtual tutor, and students who commute or cannot attend office hours find AI to be a valuable asset. Others made use of AI by creating a study timeline for exams or locating equations without searching through textbooks. While the convenience of AI is exceptional, concerns of dependency emerge. Students who rely on AI to find shortcuts or answers, postponing deeper learning, risk falling into what one student referred to as "learning debt." The convenience that AI provides may encourage poor time management skills or even foster over-reliance. Generative AI was also found to be less stimulating than human interaction. Instructors possess the ability to read facial expressions and interpret questions, making communication of the student's needs more streamlined. Students commented that seeing a problem worked out by the instructional staff is more memorable compared to reading methods or solutions off a screen. Ultimately, students who misuse generative AI may struggle in engineering programs since their reliance on AI leaves them unprepared for exams where it is restricted.

Researchers concluded that AI, when used correctly, can strengthen a students understanding of a curriculum, however, improper use of AI is only detrimental to the student. Those who rely excessively on generative AI without developing their own problem-solving skills may underperform compared to those who are investing themselves in learning the material, thus potentially jeopardizing the former's academic standing. Rather than implementing bans on the use of AI models, the researchers encourage university policy makers and professors to prepare students to use AI models in an ethical and productive manner.

Future Work

Moving forward, more studies will be conducted on student usage of generative AI. The researchers note the imperfection of using self-reported data, although valuable, may lack accuracy. The researchers would like to cross-reference raw input data from TopHat ACE against student claims. Inputs can be systematically compared against assigned questions embedded in online textbook readings, lecture prep questions, in-class problems, and homework to check whether students directly asked generative AI for solutions. Further, students' usages of AI can be reassessed with the TopHat data to develop a comprehensive understanding of AI's role in engineering education. This study also presented questions on generative AI's impact on academic performance. In the future, the researchers may investigate the attempt ratios for graded work in student classes that had access to TopHat ACE with historical results. Another question that arose probes the training of Top Hat ACE. Feeding ACE workflow to additional engineering statics material may widen its range of knowledge on the subject, enabling it to better assist the students in a variety of tasks.

References

[1] D. T. Pham and P. Pham, "Artificial intelligence in engineering," *International Journal of Machine Tools and Manufacture*, vol. 39, no. 6, pp. 937–949, 1999.

[2] D. Milmo, "Chatgpt reaches 100 million users two months after launch," *The Guardian*, vol. 3, 2023.

[3] M. Cascella, J. Montomoli, V. Bellini, and E. Bignami, "Evaluating the feasibility of chatgpt in healthcare: an analysis of multiple clinical and research scenarios," *Journal of medical systems*, vol. 47, no. 1, p. 33, 2023.

[4] A. Radford, "Improving language understanding by generative pre-training," 2018.

[5] T. H. Kung, M. Cheatham, A. Medenilla, C. Sillos, L. De Leon, C. Elepaño, M. Madriaga, R. Aggabao, G. Diaz-Candido, J. Maningo *et al.*, "Performance of chatgpt on usmle: potential for ai-assisted medical education using large language models," *PLoS digital health*, vol. 2, no. 2, p. e0000198, 2023.

[6] T. Rasul, S. Nair, D. Kalendra, M. Robin, F. de Oliveira Santini, W. J. Ladeira, M. Sun, I. Day, R. A. Rather, and L. Heathcote, "The role of chatgpt in higher education: Benefits, challenges, and future research directions," *Journal of Applied Learning and Teaching*, vol. 6, no. 1, pp. 41–56, 2023.

[7] M. M. Rahman, H. J. Terano, M. N. Rahman, A. Salamzadeh, and M. S. Rahaman, "Chatgpt and academic research: A review and recommendations based on practical examples," *Rahman, M., Terano, HJR, Rahman, N., Salamzadeh, A., Rahaman, S.(2023). ChatGPT and Academic Research: A Review and Recommendations Based on Practical Examples. Journal of Education, Management and Development Studies*, vol. 3, no. 1, pp. 1–12, 2023.

[8] A. Gilson, C. W. Safranek, T. Huang, V. Socrates, L. Chi, R. A. Taylor, D. Chartash *et al.*, "How does chatgpt perform on the united states medical licensing examination (usmle)? the implications of large language models for medical education and knowledge assessment," *JMIR medical education*, vol. 9, no. 1, p. e45312, 2023.

[9] P. Hacker, A. Engel, and M. Mauer, "Regulating chatgpt and other large generative ai models," in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, 2023, pp. 1112–1123.

[10] A. Zarifhonarvar, "Economics of chatgpt: A labor market view on the occupational impact of artificial intelligence," *Journal of Electronic Business & Digital Economics*, vol. 3, no. 2, pp. 100–116, 2024.

[11] H. Yu, "Reflection on whether chat gpt should be banned by academia from the perspective of education and teaching," *Frontiers in Psychology*, vol. 14, p. 1181712, 2023.

[12] R. Abdelghani, Y.-H. Wang, X. Yuan, T. Wang, P. Lucas, H. Sauzéon, and P.-Y. Oudeyer, "Gpt-3-driven pedagogical agents to train children's curious question-asking skills," *International Journal of Artificial Intelligence in Education*, vol. 34, no. 2, pp. 483–518, 2024.

[13] X. Deng and Z. Yu, "A meta-analysis and systematic review of the effect of chatbot technology use in sustainable education," *Sustainability*, vol. 15, no. 4, p. 2940, 2023.

[14] J. Rudolph, S. Tan, and S. Tan, "Chatgpt: Bullshit spewer or the end of traditional assessments in higher education?" *Journal of applied learning and teaching*, vol. 6, no. 1, pp. 342–363, 2023.

[15] D. R. 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, 2024.

[16] L. Welding, "Half of college students say using ai on schoolwork is cheating or plagiarism," *Best Colleges*, 2023.

[17] O. P. Yadava, "Chatgpt—a foe or an ally?" *Indian Journal of Thoracic and Cardiovascular Surgery*, vol. 39, no. 3, pp. 217–221, 2023.

[18] C. K. Y. Chan and W. Hu, "Students' voices on generative ai: Perceptions, benefits, and challenges in higher education," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, p. 43, 2023.

[19] M. Barry, S. Wismer, and L. Dosse, "Statics and mechanics of materials: An example-based approach," *Top Hat*, 2020.

[20] J. W. Creswell, *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. pearson, 2015.

[21] V. Elliott, "Thinking about the coding process in qualitative data analysis," *Qualitative report*, vol. 23, no. 11, 2018.