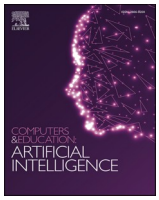
Computers and Education: Artificial Intelligence 7 (2024) 100253

Contents lists available at ScienceDirect

Computers and Education: Artificial Intelligence

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



Cheating in the age of generative AI: A high school survey study of cheating behaviors before and after the release of ChatGPT

Victor R. Lee [a,*], Denise Pope [a], Sarah Miles [b], Rosalía C. Zárate [c]

[a] Stanford University, 485 Lasuen Mall, Stanford, CA, 94305, USA
[b] Challenge Success, 531Lasuen Mall #20053, Stanford, CA, 94309, USA
[c] DRZ Strategic Consulting Group LLC, San Francisco, CA, USA

ARTICLE INFO

Keywords:
Cheating
Academic integrity
ChatGPT
Generative AI
Chatbots
Secondary school

ABSTRACT

The public release of ChatGPT and other generative AI chatbot technologies has been accompanied by questions about how academic integrity and student cheating behaviors will be impacted. We analyzed anonymous survey data from three high schools to see if self-reported cheating numbers changed following the introduction of ChatGPT and similar technologies. This survey data set is unique in that data on cheating had been collected with this set of schools both before and after November 2022, when ChatGPT was publicly released and drew attention to these educational concerns. The results suggested that cheating behaviors remained relatively stable after the introduction of this current generation of generative AI chatbot technology. However, some changes in reported behaviors differed depending on the type of cheating (social cheating, AI-related cheating, etc.). Additional survey questions about high school students' AI chatbot usage and the perceived allowability of such technology revealed mixed opinions on the acceptability of using AI for various academic-related tasks. Most students did not think that using a chatbot to produce an entire paper or complete an entire assignment should be allowable. However, there was support for using AI chatbots to help students to start on assignments and papers and to help explain new concepts to them.

1. Introduction

Within months of the November 2022 release of ChatGPT, the artificial intelligence (AI) chatbot technology that can produce human-like text on-demand, educators expressed concerns that students would be using this technology to complete assignments that would previously have been assumed to be completed without AI assistance (Tilli et al., 2023). Plagiarism, cheating, and ethics of its use for classroom assignments, examinations, and other assessments continue to be discussed in the news, preprints, and academic commentaries. Various surveys have provided some data about ChatGPT use, such as 30% of surveyed undergraduate students using it for their assignments (Intelligent.com, 2023), that more K-12 teachers (51%) than students (33%) reported using it for school related activities (Walton Family Foundation, 2023), or that more students (58% of 12–18 year olds) have used it than parents (30%) (Common Sense Media, 2023). Educator responses in the US to ChatGPT have ranged from initial full bans of the technology (New York, see Elsen-Rooney, 2023)[1] to the enthusiastic embrace of AI-generated writing detection technologies (Yan et al., 2023). While numerous questions abound regarding how educational practice will need to change with the increased availability of ChatGPT and other new generative AI technologies, there are also some unanswered questions about what initial impact the widespread availability of these technologies had on students with respect to what is often described as their academic integrity. That is, while ChatGPT and other emerging generative AI chatbot technologies could be used by students in ways that some consider "cheating", we do not know the extent to which it actually has been used in that way now that it has been made publicly available and actively discussed in the media.

This paper offers an opportunistic empirical examination of some of these matters. For several years prior to the release of ChatGPT, *Challenge Success* (abbreviated as CS), has been conducting survey research for schools with respect to school climate and student academic integrity. That work has provided snapshots of how prevalent cheating has been in individual schools that have previously been surveyed by CS. Following the release of ChatGPT and other similar technologies, CS

* Corresponding author.
E-mail address: vrlee@stanford.edu (V.R. Lee).

[1] NY has revised this policy.

https://doi.org/10.1016/j.caeai.2024.100253
Received 28 January 2024; Received in revised form 8 June 2024; Accepted 10 June 2024
Available online 12 June 2024

collected additional data in 2023, including with schools that had been surveyed in previous years. Given that those data were available, and in consideration of the many questions that educators raise about academic integrity in an era where such technology will be widespread, this paper provides a unique before-and-after look at students reported "cheating" behaviors with respect to the release of ChatGPT and flood of attention to generative AI's impact on the educational system in the first half of 2023.

As the release of ChatGPT was not anticipated, certain ideal conditions could not be established. For instance, while CS has obtained large amounts of data over several years and with many schools repeatedly, the exact schools and intervals of sampling for what would be ideal comparisons cannot be retroactively included. With those known limitations, the data do show consistency and are useful for analysis, particularly given current concerns about a new rash of "cheating" being triggered by generative AI technologies at this point in time.

2. Background

2.1. Chatbots and generative AI

While ChatGPT has been named specifically, it is important to recognize that is just one of many generative artificial intelligence technologies. ChatGPT, developed by OpenAI, is a chatbot technology – an application that receives text input from a user and produces a response. Others included Bard (now Gemini) from Google, Claude from Anthropic, as well integrations into Microsoft's search engine, Bing. These are considered instances of "generative" AI which is able to produce new content rather than just detect patterns or relationships in data or input. Text – whether it is in a response to a query or an essay on request – is one of the most recognized outputs of chatbots. However, some generative AI and chatbots are capable of producing images and computer code on demand (Marcus et al., 2022; Nguyen & Nadi, 2022).

Although impressive relative to precursor technologies given the ability to produce novel-seeming responses, current generative AI chatbot technologies are known to produce "hallucinations" (statements that are incorrect or refer to events or objects that do not exist) lack adequate responses to requests (such as about current events), and to encode or amplify biases within the training data such as stereotypes about groups of people (Noble, 2018). Because they have been trained on predominantly online texts, chatbot technologies are also exhibiting the ability to only engage with specific registers, languages, and sociolects as participation in producing online texts is not equal due to larger systemic factors (e.g., Hargittai & Shaw, 2015). Moreover, there are valid concerns about copyright, intellectual property, accessibility, privacy, surveillance, transparency, and environmental impact related to generative AI and chatbot technology (Bender et al., 2021).

2.2. Academic integrity and "cheating"

Cheating is a valenced term that presumes deviation from expected behaviors is negative. It is often invoked when standards of academic integrity are not being met, usually with the assumption that deviating from those standards is an intentional decision on the part of the student. For decades prior to widespread public awareness of generative AI technology and its possible use by students, academic cheating has been documented as prevalent among secondary school students in the US. In a 2012 survey of more than 23,000 high school students in the US commissioned by the Josephson Institute, 75% had reported copying someone else's homework and 52% reported cheating on a test in the previous year (Josephson Institute, 2012). Other research also suggested high occurrence of cheating behaviors. Wangaard and Stephens (2011) surveyed 3600 students from six economically and ethnically diverse high schools in the Northeastern United States and found 95% of respondents had reported engaging in at least one form of academic cheating in the past academic year. *Challenge Success* surveyed 1400 students with 97% of students admitting to cheating at least once in the past year, and 26% reporting multiple instances (Conner, Galloway, & Pope, 2009).

The primary methodological approach for ascertaining cheating in these studies is student surveys, typically from a confidential or anonymous self-report. Students are typically informed that their responses cannot and will not be linked to them individually and would not be accessible to any school personnel. The use of self-report, especially with behaviors that are considered in some way to be morally transgressive has generated concerns. First, students may have poor recall, as recollection of actions can be unreliable due to limitations of human memory (Loftus & Loftus, 1976). Second, there is concern about social desirability of responses where giving information that the respondent thinks is the desirable one to give may lead them to respond in different ways. This could contribute to over- or under-counting of cheating.

Overcounting could occur because students may feel the goal of the study is to address certain behaviors in school, and thus the respondents are more sensitized to those behaviors and may consider reporting more than would be typical. For example, if a student was asked if they looked at someone else's test without that person's knowledge and they had caught an accidental glimpse of someone's test while glancing at the clock, the student might treat that as a "yes" response even though it was not intentional and may have had no bearing on what they did on their own test. On the other hand, undercounting is a frequent concern because behaviors that fall under the umbrella of cheating are seen as negative, and while they know the responses cannot be linked back to them, students may not want to actually acknowledge in any way that they have participated in those behaviors. However, students in this line of research, and in this current study, have only asked about specific behaviors and how many times they had engaged in those behaviors in a stated time interval (e.g., the previous month). For example, they were asked how many times in the previous month they had copied someone else's homework. Questions were designed to minimize use of the term "cheat", with two exceptions. One was in the context of a student different from the respondent cheating and another was in the context of "cheat sheets", a colloquial term for a type of written aid familiar to many students.

While outside of the scope of this study, but knowing that survey self-reports (even when anonymous and stating behavior frequencies rather than assuming a negative social label) raise questions about data quality, we do note also that other approaches have been used in the academic integrity literature to measure "cheating" behavior with findings that are consistent and complementary to what accepted survey instruments have yielded. For instance, students have been asked to estimate the frequency of their peers' cheating behaviors rather than their own, and researchers have gathered qualitative data from students through naturalistic observations and focus groups to measure cheating (Stephens & Wangaard, 2013; Waltzer et al., 2023). These methods still find that cheating behavior is quite prevalent across students. Furthermore, recognizing the limits of self-report instruments, given that they are reliable across repeated use and that the numbers stay the same – we also note that they do seem to measure something related to those behaviors, if not precisely those behaviors. If something new is introduced that has the potential to change those behaviors, researchers should see some change in the proxy one way or the other. For instance, if one were to use a food measurement scale for cooking purposes that is reliable in giving the same measurements but is not fully accurate because it consistently reports weight as between two and five ounces less than the actual weight, we still would expect that should the chef add significant additional weight to an object they are measuring, the new weight on the scale would still increase despite the same two to five ounce undercount the true weight. This does not eliminate the need for there to be future triangulation; one should rely on multiple scales and take other measurements with different instruments in order to be confident of exact food weight. But it also does not mean that the instrument, while limited, is incapable of detecting change.

The cheating literature provides limited guidance for how predictive demographic attributes or individual histories are with respect to participating (or reporting) cheating behaviors. Prevalence has been established in both students with high and low academic performance, although studies are mixed with regard to whether students with histories of low or high achievement tend to engage in cheating behavior more (Miller et al., 2007). Anderman and Midgley (2004) have documented that self-reports of cheating increase through high school. Boys have been assumed to cheat more than girls, although the results are inconsistent (Jensen et al., 2002). Similar levels of cheating have been found in both public and private schools (Galloway, 2012). Various factors have been identified with the prevalence of these behaviors. They include competing pressures, the perception that school assignments had low value, and perceived high importance of grades (Waltzer et al., 2023). Perceiving or believing that their peers are cheating is also associated with higher likelihood of cheating (Zhao et al., 2022).

2.3. Two facets of reported cheating

There is an old admonition, "do as I say, not as I do", which is reflective of a common disconnect between espoused and enacted behaviors. This points to a widely held belief that what people may hold or recommend as an ideal may deviate from what they actually do, which can happen knowingly or not. For example, it is common in research on teaching practice to conduct examinations of what teachers' beliefs are and what their actions are in the classroom (Buehl & Beck, 2014; Cross Francis et al., 2014), and while much of that research highlights the disconnect between teacher beliefs and their classroom actions, some arguments have been made that the differences may be illusory (Cross Francis, 2015) or that they can lessen over time (e.g., Polly & Hannafin, 2011). To the extent possible, it is worthwhile to consider both. For use of generative AI in school-related tasks, patterns of use and patterns of preference are still being negotiated. Therefore, in this study, beyond examining change in reported behaviors due to recent highly publicized advances in generative AI, we examine what students believe to be acceptable use in addition to how those who have reported using it. Granted, we do look to self-report of prior use as the proxy in following the established methods in the academic integrity research community and for consistency with existing data obtained by CS against which the post-ChatGPT data had been collected with the expected caveats. The focus of this study is on whether there was a change – but it is not an in-depth examination for what is driving change beyond the release of these new technologies. Given the alarm expressed about the possible student use of these technologies for "cheating" in schools, we consider this sufficient motivation for the current analysis.

3. Research questions

The questions we ask in this study are.

- To what extent have high school students' self-reported cheating behaviors changed in 2023 once chatbot technology became more widely known and available?
- How frequently do high school students report using chatbot AI technology for school-related tasks and how does that vary across task types?
- To what extent and for which activities do high school students believe chatbot technology should be permitted for academically-related student use and how does that compare to reported behaviors?

4. Method

This study is survey-based, using an instrument developed to understand student perceptions of their school experience, and used by the partner organization CS for more than 10 years.

The survey instrument was 151 items long and took students between 20 and 39 min to complete. Students could choose to take the survey in Spanish or English. Parents/Caregivers were given information about the survey and the opportunity to opt their children out of the survey. Students could skip any question they wished and provided their assent before taking the survey with explicit information to the students that the responses were anonymous, not identifiable to school personnel, and had no implications to their grades or class standing.

The survey included a cheating scale (adapted from McCabe, 2001, pp. 38–43) with 12 items and four options for responses (see Table 1). These items were presented in a table with the prompt, "*In the past month, how often have you engaged in any of the following actions?*". Response options were *"Never", "Once", "2 to 3 times", "4 or more times"*. Items covered topics ranging from plagiarism to forms of social cheating involving sharing or obtaining information from peers when it is not permitted, to providing forged or false excuses to get extra time for an assignment or assessment. Most questions were about individual behavior affecting that individual's assignment or assessment. However, consistent with existing models (e.g., McCabe, 2001, pp. 38–43), there were also items about sharing information (e.g., allowing others to cheat) even when it did not directly affect one's own submission. Consistent with methods in prior research on "cheating", questionnaires limit the use of the word "cheating" to avoid asking about identification with being a "cheater". Rather, individual behaviors are presented. Students and their caregivers are given multiple notices that the information is deidentified and confidential to also encourage honest responses.

Beyond the cheating scale and stated above, the entire survey addressed matters of engagement with school, perceived amount of support available at school, perceived pressures and demands from school and parents, factors such as sleep and extracurricular commitments. Some of the predictiveness of these variables are reported in Lee et al. (2024) for the sample subsets where such analyses could be run and are summarized in the discussion.[2] All questions were optional, so

Table 1
Cheating and AI chatbot survey item structure.

Survey Scale	Survey Item Code	Survey Item Text
In the past month, how often have you engaged in any of the following actions? (Response Options: "Never", "Once", "2 to 3 times", "4 or more times")		
Cheating	C33	Working on an assignment with others (peers, parents, etc.) when the teacher asked for individual work.
Cheating	C34	Getting questions or answers from someone who has already taken the assessment.
When asked for usage:		
In the past 3 months*, have you used an AI chatbot (such as ChatGPT) for any school-related tasks?*		
(Response Options: "Never", "Once", "2 to 3 times", "4 or more times")		
When asked for allowability:		
To what extent do you think an AI chatbot (such as ChatGPT) should be allowed for the following school-related tasks?		
(Response Options: "Not allowed", "Allowed sometimes or under some circumstances", "Allowed all of the time", "I am not sure")		
AI	AI1	To generate ideas for a paper, project, or assignment
AI	AI2	To complete (or edit) a portion of paper, project, or assignment
AI	AI3	To write all of a paper, project or assignment

Note. For students who had reported prior use of an AI Chatbot, AI items were asked as both a usage question and then later as an allowability question. For those who did not report prior use of an AI Chatbot, only the allowability version of the questions were asked.

[2] Note that due to timing, slightly different versions of questions were asked and students could skip questions or sections, meaning not all respondents to one section completely responded to other sections.

students could skip any single question or section they wished. They could also opt to leave the survey incomplete. The phrasing of questions related to the questions of interest for this paper is provided in Table 1. The various questions asked are provided in Tables 3–7.

In 2023, the survey was amended twice. The first amendment, approved in February 2023, modified the wording of one question on the cheating scale to more explicitly address the use of AI as potentially unapproved potential technology use and to be more inclusive of different academic work products. Previously, the wording of that item asked about frequency of recently "Using an electronic/digital device as an unauthorized aid during an assessment." The new wording in 2023 became "Using an Artificial Intelligence or digital device (*e.g.* Chat GPT, smart phone) as an unauthorized aid during an assessment, school assignment, or homework" (emphasis added to show additions and changes). This item change was known to have methodological consequences, although CS's priority in this change was to give desired information requested by schools who would ultimately receive aggregated reports of survey responses. Supporting academic research was a consideration but secondary to responding to participants' needs and interests.

The second amendment, approved in mid-April 2023, was to add a new set of questions specific to students' self-reported AI chatbot use and whether students believed such technology should be allowable for use with school-related tasks. With timing of the human subjects review and approval, school testing, and the end of the school year, this version could only be administered to a few schools. Furthermore, human subjects approval timing was such that at one of the few schools in our sample that could complete the survey, students were in the midst of finals week. Thus, the response rate was overall lower relative to previous years although still reasonable for survey research – their sample to population ratio provides a 6% margin of error for a 95% confidence interval. The focus of this paper is only on questions from the Cheating and AI scales (shown in Tables 3–7), with analyses of other scales ongoing for future reports.

For this study, criteria for inclusion included whether (1) for schools located in the US, (2) the survey with the first amendment – to include wording inclusive of AI and Chatbot technology – had been administered in the spring of 2023, after the release and public awareness of ChatGPT and other generative AI technologies, and (3) data from 2019 (the last year pre-remote learning) were available. Three schools fit these criteria, designated as Private High (surveyed on 4/24/2023), Public High (surveyed on 5/12/2023) and Charter High (3/21/2023). Two of those (Private and Charter High) also had data from 2022. Two (Private and Public High) had completed questions from the second survey amendment, which had questions about specific uses of AI chatbot technology. Public High was the school that could only administer the survey during finals week before summer break, with a shortened response window. These schools and the demographics that can be reported with consideration for formal confidentiality agreements are described in Table 2. These schools came from two different states in the continental United States across multiple time zones. As we will discuss below, while there are forms of heterogeneity in this sample (across race, gender, geography, and school type), we urge readers to consider carefully what generalizations should and should not be made to what other populations.

5. Analysis

Responses to the cheating questions were recoded for those who reported never engaging in the described item behavior and those who had, regardless of whether they had done it one or more than four times. The 2023 responses served as a baseline for comparisons to 2022 and 2019. Chi-Squared tests with an alpha level of 0.05 were used, plus Bonferroni's correction if significance had been found at the 0.05 alpha and multiple pairwise tests were used for multiple years. Cell sizes were not smaller than 5 so our team decided Fisher's exact test was not necessary. The comparisons reflected times before and after concerns about AI-enabled cheating, including the release of ChatGPT. The 2019 data represented pre-remote learning, while data from 2020 to 2021 were not used (nor necessarily collected per school constraints) due to school disruptions and many schools operating at least partially online rather than through face-to-face meetings. Some recent studies document how some populations of students exhibited specific behaviors related to academic integrity that differed substantially during remote learning periods at the onset of the pandemic (e.g., Jenkins et al., 2023; Yazici et al., 2023), further motivating the exclusion of the 2020 and 2021 years. The schools in this article had returned to in-person instruction during the latter portion of the 2020–2021 academic year and was in person for the entire 2021–2022 and 2022-2023 school years. Analysis included only students who responded to all cheating scale questions; students who only responded to only one or a few of the cheating questions were excluded from the current analysis.

For AI chatbot allowability items, a "not sure" response was treated like a non-response, excluding less than 10% of responses from each school for items AI1 to AI4, and less than 19% for items AI5 and AI6 (see Tables 6–7), the latter two considered 'frontier' uses of chatbots and not as well-known to students as of early 2023.

6. Results

6.1. Responses on cheating scale

Across schools for the years analyzed, between 59.9% and 69.5% reported engaging in at least one academically dishonest behavior in the previous month. These numbers were relatively consistent for Private High (2019: 61.3%, 2022: 62.4%, 2023: 59.9%). The range was higher for Charter High (2019: 82.7%, 2022: 64.0%, 2023: 64.4%) and Public High (2019: 69.5%, 2023: 62.7%). Conner, Galloway, and Pope (2009)

Table 2
School survey response and respondent demographic information.

School Pseudonym	School Type	2019 Survey Data	2022 Survey Data	2023 Survey Data	Detailed AI Questions Included?	2023 Respondents' Gender	2023 Respondents' Race
Private High	Small Private, All Girls (<250 students)	Yes (181 Respondents, 81.9% Response Rate)	Yes (213 Respondents, 92.2% Response Rate)	Yes (202 Respondents, 87.8% Response Rate)	Yes	92% Female, 5% Nonbinary	37% White, 35% Asian, 5% Hispanic/Latinx, 17% Multiracial
Public High	Large Public, Co-ed (>1500 students)	Yes (1319 Respondents, 65.8% Response Rate)	No	Yes (250 Respondents, 14.0% Response Rate)	Yes	46% Female, 44% Male, 5% Nonbinary	28 % White, 46% Asian, 7 % Hispanic/Latinx, 11% Multiracial
Charter High	Medium Charter, Co-ed (<1500 students)	Yes (468 Respondents, 72.6% Response Rate)	Yes (356 Respondents, 54.8% Response Rate)	Yes (419 Respondents, 61.5% Response Rate)	No	50% Female, 36% Male, 9% Nonbinary	White 32%, Asian 38%, Hispanic/Latinx 11%, Black/ African-American 9%, Multiracial 7%

Note. Gender and Race/Ethnicity are self-identified by students and can be skipped. Race/Ethnicity categories with less than 5% are withheld per human subjects and school privacy agreements.

Table 3
Private High students' reported cheating behaviors.

Cheating Behavior Items	Private High					
	2019 (N = 181)		2022 (N = 213)		2023 (N = 202)	
	Yes	No	Yes	No	Yes	No
C33: Working on an assignment with others (peers, parents, etc.) when the teacher asked for individual work.	47.51%	52.49%	46.95%	53.05%	44.06%	55.94%
C34: Getting questions or answers from someone who has already taken the assessment.	11.05%	88.95%	10.33%	89.67%	11.39%	88.61%
C35: Helping someone else cheat on an assessment.	4.97%	95.03%	9.86%*	90.14%*	3.96%	96.04%
C36: Copying from another student during an assessment with their knowledge.	3.31%	96.69%	4.23%	95.77%	6.44%	93.56%
C37: Copying from another student during an assessment without their knowledge.	4.42%	95.58%	5.16%	94.84%	3.47%	96.53%
C39: Paraphrasing or copying a few sentences of material from a written source without referencing/citing it in a paper or on a project.	20.99%	79.01%	26.29%	73.71%	28.22%	71.78%
C40: Using unpermitted cheat sheets or sources of information during an assessment.	2.21%	97.79%	6.10%	93.90%	3.47%	96.53%
C41a (pre 2023 wording): Using an electronic/digital device as an unauthorized aid during an assessment.	1.66%*	98.34%*	4.23%	95.77%		
C41b (2023 wording): Using an Artificial Intelligence or digital device (e.g. Chat GPT, smart phone) as an unauthorized aid during an assessment, school assignment, or homework.					6.44%	93.56%
C42: Copying material, almost word for word, from any source and turning it in as your own work.	1.66%	98.34%	1.88%	98.12%	2.48%	97.52%
C44: Using a false or forged excuse to obtain an extension on a due date or delay taking an assessment.	10.50%	89.50%	12.21%	87.79%	11.88%	88.12%
C45: Turning in work done by someone else.	1.10%	98.90%	1.41%	98.59%	1.98%	98.02%
C46: Copying someone else's homework.	33.15%	66.85%	27.70%	72.30%	24.26%	75.74%

*$p < 0.05$ relative to 2023.

Table 4
Public High students' reported cheating behaviors.

Cheating Behavior Items	Public High			
	2019 (N = 1319)		2023 (N = 250)	
	Yes	No	Yes	No
C33: Working on an assignment with others (peers, parents, etc.) when the teacher asked for individual work.	52.46%*	47.54%*	45.20%	54.80%
C34: Getting questions or answers from someone who has already taken the assessment.	31.99%	68.01%	25.60%	74.40%
C35: Helping someone else cheat on an assessment.	17.29%	82.71%	14.40%	85.60%
C36: Copying from another student during an assessment with their knowledge.	13.27%	86.73%	10.40%	89.60%
C37: Copying from another student during an assessment without their knowledge.	7.43%	92.57%	7.20%	92.80%
C39: Paraphrasing or copying a few sentences of material from a written source without referencing/citing it in a paper or on a project.	28.81%	71.19%	31.20%	68.80%
C40: Using unpermitted cheat sheets or sources of information during an assessment.	7.81%	92.19%	6.40%	93.60%
C41a (pre 2023 wording): Using an electronic/digital device as an unauthorized aid during an assessment.	7.28%*	92.72%*		
C41b (2023 wording): Using an Artificial Intelligence or digital device (e.g. Chat GPT, smart phone) as an unauthorized aid during an assessment, school assignment, or homework.			15.20%	84.80%
C42: Copying material, almost word for word, from any source and turning it in as your own work.	9.25%	90.75%	5.20%	94.80%
C44: Using a false or forged excuse to obtain an extension on a due date or delay taking an assessment.	12.89%	87.11%	9.60%	90.40%
C45: Turning in work done by someone else.	5.91%	94.09%	3.20%	96.80%
C46: Copying someone else's homework.	35.41%	64.59%	27.20%	72.80%

*$p < 0.05$ relative to 2023.

would be the first year in which responses could be provided after the rollout of ChatGPT and other similarly powerful AI chatbots, it is worth noting that the overall percentages from responses stayed about the same or decreased overall relative to prior years (i.e., Public High 2022; 2023; Charter High 2019; 2023). Responses to cheating items are provided for the three schools in Tables 3–5

6.1.1. *Cheating scale items without significant difference*

Across all three schools, five items on the cheating scale (C37, C39, C40, C44, C45; refer to Tables 3–5) showed no significant difference compared to 2023, indicating no major changes in reported behaviors. Pertinent to AI-generated writing, there were no significant changes for items C39 (paraphrase or copy a few sentences from a written source without attribution) and C45 (turning in work done by someone else). It may be that students do not consider AI technology a written source, the sentence level is not where paraphrasing or copying happens, or that students did not consider AI to be 'someone else'.

With respect to potentially using an AI chatbot as an aid during an assessment, the lack of significant differences for C40 suggests there has not been a major overall proportional change driven by the introduction of AI.

6.1.2. *Cheating scale items less connected to AI with some significant differences*

There were five items where some significant differences relative to 2023 responses appeared at least within one school (Tables 3–5). These included C33, C34, C35, C36, and C46 (see Tables 3–5 for item text). Of those, there was only one item that showed a difference between 2022 and 2023 (which only could be examined for Private High and Charter High). That one item was C35: helping someone else cheat on an assessment. This is a social form of cheating that is not expected to relate to AI directly. This difference appeared only for Private High between 2022 and 2023 ($\chi^2 = 4.679$ df = 1 p = 0.031), which if considered with a Bonferroni connection, would not be significant. Charter High had a significant difference with Bonferroni correction applied for 2019 to 2023 ($\chi^2 = 5.905$ df = 1 p = 0.015) but not between 2022 and 2023 ($\chi^2 = 0.005$ df = 1 p = 0.941).

The other differences were for 2019 to 2023 comparisons. Public High and Charter High had significant differences for items C33: working on an assignment with others when it was supposed to be done individually (Public High: $\chi^2 = 4.153$ df = 1 p = 0.042; Charter High: $\chi^2 = 20.221$ df = 1 $p < 0.001$) and C46: copying someone else's homework (Public High: $\chi^2 = 5.937$ df = 1 p = 0.015; Charter High: $\chi^2 = 15.56$ df =

Table 5
Charter High students' reported cheating behaviors.

Cheating Behavior Items	Charter High					
	2019 (N = 468)		2022 (N = 356)		2023 (N = 419)	
	Yes	No	Yes	No	Yes	No
C33: Working on an assignment with others (peers, parents, etc.) when the teacher asked for individual work.	67.52% ***	32.48% ***	52.25%	47.75%	52.51%	47.49%
C34: Getting questions or answers from someone who has already taken the assessment.	50.43%	49.57%	33.99%	66.01%	33.65%	66.35%
C35: Helping someone else cheat on an assessment.	26.71% **	73.29% **	19.10%	80.90%	19.57%	80.43%
C36: Copying from another student during an assessment with their knowledge.	24.36%	75.64%	15.45%	84.55%	15.51%	84.49%
C37: Copying from another student during an assessment without their knowledge.	12.82%	87.18%	5.62%	94.38%	9.31%	90.69%
C39: Paraphrasing or copying a few sentences of material from a written source without referencing/citing it in a paper or on a project.	30.56%	69.44%	19.94%	80.06%	24.82%	75.18%
C40: Using unpermitted cheat sheets or sources of information during an assessment.	10.26%	89.74%	8.99%	91.01%	11.22%	88.78%
C41a (pre 2023 wording): Using an electronic/digital device as an unauthorized aid during an assessment.	10.47% ***	89.53% ***	8.99% ***	91.01% ***		
C41b (2023 wording): Using an Artificial Intelligence or digital device (e.g. Chat GPT, smart phone) as an unauthorized aid during an assessment, school assignment, or homework.					24.11%	75.89%
C42: Copying material, almost word for word, from any source and turning it in as your own work.	10.68%	89.32%	7.02%	92.98%	9.31%	90.69%
C44: Using a false or forged excuse to obtain an extension on a due date or delay taking an assessment.	13.03%	86.97%	12.08%	87.92%	13.13%	86.87%
C45: Turning in work done by someone else.	7.26%	92.74%	7.30%	92.70%	8.83%	91.17%
C46: Copying someone else's homework.	46.15% ***	53.85% ***	29.21%	70.79%	32.94%	67.06%

*p < 0.05 relative to 2023.
**p < 0.01 relative to 2023.
***p < 0.001 relative to 2023.

Table 6
Students' reported prior use for specific school-related tasks with an AI chatbot.

AI Previously Used to Perform Task	Private High (N = 69)		Public High (N = 140)	
	Yes	No	Yes	No
T1: To generate ideas for a paper, project, or assignment.	39.13%	60.87%	50.00%	50.00%
T2: To complete (or edit) a portion of paper, project, or assignment.	5.80%	94.20%	31.65%	68.35%
T3: To write all of a paper, project or assignment.	1.45%	98.55%	19.68%[a]	80.32%
T4: To explain a new concept.	42.65%	57.35%	39.29%	60.71%
T5: To generate computer code.	7.25%[b]	92.75%	16.43%	83.57%
T6: To generate images or pictures.	17.39%	82.61%	20.00%	80.00%

[a] N = 137.
[b] N = 68.

1 p < 0.001). Only Charter High showed a significant difference from 2019 to 2023 for C34: getting questions or answers from someone who had taken the same assessment earlier ($\chi^2 = 24.777$ df = 1 p < 0.001). Public High exceeded the alpha of 0.05 for both C34 ($\chi^2 = 3.725$ df = 1 p = 0.054). Only Charter High showed significant differences between 2019 and 2023 for item C36 ($\chi^2 = 10.197$ df = 1 p = 0.001), which was copying from another student during an assessment with that other student's knowledge, another social cheating transaction.

These are all social cheating transactions where someone else who should not have been involved in the work was involved in that their work was copied, advance knowledge was shared, or they provided assistance on what was supposed to be an individual assignment.

6.1.3. Cheating scale items connected to AI with some significant differences

As noted above, there were five items with no significant differences in any comparison with 2023. That included items C39 and C45, which could be pertinent to using AI for providing text to use in writing assignments. Two other items seemed to have relevance to AI: C41 and C42. C41 was the most direct question about artificial intelligence use. Note that the language of C41 changed in 2023 as AI was not as prominent of a concern until the end of 2022. The former language involved asking about use of a device in an unauthorized way during an assessment. The newer language explicitly named artificial intelligence and broadened to include assessments, school assignments, and homework. This newer language was used as it was unclear how students recognized chatbot technology and if its use was for them bound to mobile devices. Moreover, with heightened public concern around chatbot technology, explicitly naming various writing products beyond assessments was deemed appropriate in order to directly consider

Table 7
Students' perceived allowability of use for specific school-related tasks with an AI chatbot.

AI Allowability Item	Private High (N = 202)			Public High (N = 250)		
	Always	Sometimes	Never	Always	Sometime	Never
AI1: To generate ideas for a paper, project, or assignment.	14.44%[a]	54.01%	31.55%	22.61%[b]	58.70%	18.70%
AI2: To complete (or edit) a portion of paper, project, or assignment.	6.56%[a]	29.51%	63.93%	9.82%[c]	41.52%	48.66%
AI3: To write all of a paper, project or assignment.	1.55%[a]	3.09%	95.36%	4.33%[b]	9.09%	86.58%
AI4: To explain a new concept.	45.50%[d]	41.27%	13.23%	59.66%[e]	32.19%	8.15%
AI5: To generate computer code.	4.68%	42.11%	53.22%	16.06%[b]	54.40%	29.53%
AI6: To generate images or pictures.	24.86%	56.22%	18.92%	37.68%	45.51%	16.91%

[a] N = 201.
[b] N = 249.
[c] N = 247.
[d] N = 200.
[e] N = 248.

several different types of assigned writing that students may be asked to do for school. This showed a significant difference for Charter High (2019–2023: $\chi^2 = 10.197$ df = 1 p = 0.001, 2022–2023: $\chi^2 = 29.882$ df = 1 p < 0.001) and Public High (2022–2023 $\chi^2 = 15.886$ df = 1 p < 0.001), but not for Private High when a Bonferroni correction is applied (2019–2023: $\chi^2 = 4.316$ df = 1 p = 0.038, 2022–2023: ($\chi^2 = 0.617$ df = 1 p = 0.432).

Only Public High showed significant differences between 2019 and 2023 for C42($\chi^2 = 3.883$ df = 1 p = 0.049), "copying material, almost word for word, from any source and turning it in as your own work", and reports of this behavior actually decreased, which seems to go in the opposite direction than one would expect with AI chatbot availability. However, appropriate limits on generalizability should be considered given sampling limitations and that this is one school's student population.

As noted above, the other significant differences – namely, those associated with C41 about mobile device and AI use - might be partly explained by the broadened contexts of use as well as the mention of AI. This was a known risk in the modification of this item and its wording. However, it was deemed acceptable given additional questions had been added. To know more about AI use, we turn to the two schools that completed additional questions specifically about generative AI chatbot technology. These two schools were able to provide these responses given the timing of human subjects approval for new questions.

6.2. Responses on AI chatbot scale

6.2.1. Prior experience with chatbot technology

Respondents at Private High and Public High were asked, as part of the detailed AI Chatbot usage scale, whether they had prior experience with an AI Chatbot. The wording for that question was: In the past three months, have you used an AI chatbot (such as ChatGPT)? There were three response options: *"Have Never Used"*, *"Have Used"*, and *"I don't know what an AI Chatbot is and/or I don't know how to use it"*. Responses at Private High were 57.7% (Have Never Used), 35.6% (Have Used), and 6.7% (Don't Know What It Is). Responses at Public High were 36.8% (Have Never Used), 54.0% (Have Used), and 9.2% (Don't Know What It Is). Overall, a larger percentage of Public High respondents had reported prior AI chatbot use experience than Private High.

6.2.2. AI scale chatbot usage items

Only those respondents who had reported previous use of an AI Chatbot were asked about how they had used it in the preceding 3 months, shown in Tables 6 and 7 Three months, rather than one month, was selected as the range of time for this to be more inclusive of isolated early use of the technology when it was first publicized. Most of these respondents had not used an AI chatbot for the purposes listed in the survey. For Private High, the majority of respondents reported that they had not used an AI chatbot for any of the surveyed tasks in the preceding 3 months (61.7% responded *"Never"* for the behavior in AI1, 95.6% for AI2, 100% for AI3, 60.9% for AI4, 95.5% for AI5, and 83.8% for AI6). For Public High, the majority of respondents reported they had not used an AI chatbot for 5 out of the 6 surveyed tasks in the preceding 3 months (54.3% responded *"Never"* for the behavior in AI1, 72.0% for AI2, 94.2% for AI3, 86.0% for AI5, and 84.2% for AI6). The one item where less than a majority had reported that they had not used an AI Chatbot was AI4 "to explain a new concept" (48.2% had responded "Never" for AI4).

6.2.3. AI scale chatbot allowability items

All survey respondents at Private High and Public High, beyond those who had reported previous use or familiarity with AI Chatbot technology, were asked about AI chatbot allowability (shown in Tables 6 and 7). Excluding those who had not provided a response or selected *"I am unsure"* as their response, survey respondents at both Private High and Public High displayed similar response patterns with respect to AI use with completing papers, projects, and assignments. More felt that sometimes allowing the use of a chatbot to generate ideas for a paper, project, or assignment should be allowable (item AI1; Private: 54.0%, Public 58.7%). The largest percentage of respondents at each school reported that they felt that an AI chatbot should never be allowed for use to complete (or edit) a portion of a paper, project, or assignment (item AI2; Private: 63.9%, Public 48.7%), although several respondents felt it should sometimes be allowed (item AI2; Private: 29.5%, Public 41.5%). However, most respondents felt that an AI chatbot should never be allowed to complete all of a paper, project, or assignment (item AI3, Private 95.4%, Public 86.6%).

The one item where most respondents selected that an AI chatbot should always be allowed for use was related to using it to explain a new concept (item AI4: Private, 45.5%, Public 59.7%). The response pattern was mixed for whether AI should be used to generate computer code (item AI5, at Private High, 53.2% selected *"never"* and at Public High, 54.4% selected *"sometimes"*). However, the top responses to whether an AI chatbot should be allowed for use to generate images or pictures was the same between the schools (item AI6, "*sometimes allowable"*, Private: 56.2%, Public 45.4%).

7. Discussion

While these surveys were done with three different types of schools (private, public, charter), there were a number of similarities in response patterns. Among these are that, consistent with other studies which found that cheating behavior is common in US schools (Josephson Institute, 2012), more than half of respondents reported engaging in at least one academically "dishonest" behavior in the month prior in 2019, which was before the November 2022 release of ChatGPT and other chatbot technologies (such as Claude and Bard in late March of 2023). In the subsequent sections, we discuss what has been observed in this sample from the final months of the 2022–2023 academic school year.

7.1. Changes from before and after the release of ChatGPT

ChatGPT was only one of many generative AI chatbot technologies that have been released and continue to appear since November 2022, but has gained notoriety. We use its release as a reference point, although we worded items to be inclusive of multiple chatbot technologies. With that in mind, the results show the prevalence of reported cheating behavior as represented in this sample stayed similar across this time period.

It is possible that this may change as AI chatbot familiarity and prevalence increases. Only a little more than half of respondents at Public High (54%) had reported previously trying an AI chatbot and a little over a third at Private High (35%). Consistent with other surveys that report 20–50% of students having used ChatGPT (Common Sense Media, 2023; Walton Family Foundation, 2023), it appears that not all students were knowingly using AI chatbot technology by the end of that school year. However, a Pew survey conducted in the fall of the following school year (2023–2024) with teens aged 13–17 from September 26 to October 23, 2023 showed that only 32% of teen respondents have not heard of ChatGPT (Sidoti & Gottfried, 2023). That Pew survey suggests that familiarity with ChatGPT increases as family income increases, and it is better known by White teens compared to Black and Hispanic teens.

There does seem to be some increase in the use of unpermitted digital devices for school assignments, papers, and assessments, a category under which AI chatbots are considered. However, the wording of this item was modified from its previous versions, so we are limited in inferring whether it was the AI chatbots that were used more or if the inclusion of more types of assignments and/or the increased availability and use of technology by teens (Haddock et al., 2022) contributed to the different responses. We do note that the majority of respondents at all three schools (ranging from 75.8% to 84.7%) did not report using any

unpermitted technologies in the three months prior to the survey. None of the schools had yet established a school-wide policy related to AI use, according to school administrators. Indeed, many school leaders and teachers have sought data such as what is provided in this study to determine how to prioritize this concern amidst many others (including staffing shortages, student mental health needs from the pandemic, and budgetary changes due to temporary relief funds expiring).

There were not significant differences for the reported plagiarism-related behaviors (items C39, C42, and C45) at Private High and Charter High. Public High similarly did not show significant differences except for item C42. However, the change is in the direction opposite of what one would expect if AI was being used to generate writing for students to submit as their own. It decreased from 9.25% to 5.2%. Increased awareness and integration of chatbot technology into common online tools may change this.

7.2. Reported uses of chatbot technology after ChatGPT's release

A second question was how students who used chatbot technology report using the technology. The heaviest use appears to be for generating ideas for papers and assignments (Private High: 39.13%, Public High: 50.00%) and for explaining new concepts (Private High: 42.65%, Public High: 39.29%). The use of this technology to do more than generate ideas appeared to decrease inversely to the amount of the contribution of the chatbot. That is, fewer students reported using a chatbot to complete or edit their papers. Fewer still used chatbots to write entire papers. It is worth recognizing that there were still some students who did. However, even prior to ChatGPT's release, students would obtain complete papers from other sources such as online services or copying from a website. We can infer that students have a sense that there are gradations of use with which they are willing to engage.

Reported chatbot use for generating code and generating images was modest, occurring less than using the technology to explain concepts or generate ideas but more than having a chatbot write an entire paper. This could be because at the time, these were frontier uses as image generation was not integrated in initial versions of ChatGPT or Bard (now Gemini). Also, writing computer code is not yet as common of an expectation for high school students in the US as writing papers.

7.3. Relationships between beliefs about chatbot allowability and reported use

Students who used chatbot technology for writing reported greater use for idea generation and less for having it write an entire paper. This same trend appears in how students view the allowability of the technology. However, this question was asked of all students, not just those who reported prior chatbot use. We could assume some segment of nonusers consider chatbot use as impermissible, so that would decrease the percentage who viewed chatbot use as allowable.

Regardless, these results are consistent with previous research on which activities students perceive as "serious cheating," (Murdock et al., 2004). Using chatbot technology for writing a full paper is more serious and egregious than using it to get ideas. The aforementioned fall 2023 Pew survey (Sidoti & Gottfried, 2023) showed similar results in that nearly 75% of respondents who chose between ChatGPT use being acceptable or not acceptable for writing essays responded that it was not acceptable. On the other hand, Pew found (similar to our findings) that the vast majority who had an opinion on the matter (over 80%) felt ChatGPT was acceptable to use in order to research new topics.

In Lee et al. (2024), respondents who provided adequately complete responses were analyzed given known predictors such as students' academic worry, amount of time required for homework, and engagement with school along cognitive, affective, and behavioral dimensions (Fredricks et al., 2004). Only behavioral engagement was predictive of student cheating. That is, the more students reported completing the assigned work that was given to them, the less likely they were to report cheating behaviors. When each AI chatbot usage behaviors were each examined as a possible dependent variable, students who considered chatbot usage to be allowable for a specific task (such as generating ideas for an assignment or paper) were more likely to have used it for that same task. This is not altogether surprising and suggests a tendency toward consistency between beliefs and behaviors. However, when demographic variables for race, gender, school, and other perceptions are considered, much of the predictive capabilities of beliefs about allowability disappears. This suggests more about the school and circumstance are involved in student behavior than exclusively belief about what should be allowable. However, that analysis is necessarily smaller and requires larger samples for more robust modeling. That analysis is preliminary and details and caveats are in Lee et al. (2024). We recommend waiting for larger samples given the large number of variables that could be included and looking at the larger literature on academic integrity for guidance given that we are seeing limited evidence that chatbot adoption in high schools is the driver of new cheating behaviors. Cheating has been around before and was motivated for other reasons beyond the availability of this technology. However, behaviors and beliefs may change as large language models and chatbot technologies are integrated into more tools and platforms and norms change. What we provide here should be considered one measure at a specific time point so we can better understand the changes as they happen in the future.

8. Implications for schools

Although much alarm has been expressed at how quickly chatbot technology has progressed and that students will potentially be using AI chatbot technology for cheating (Dunnigan et al., 2023), there does not yet seem to be reason to assume that this technology has caused major changes nor sparked increases in cheating. This was examined in our study at a time when schools have yet to establish policies and fully understand the implications that generative AI will have on education. At the same time, other data, such as that which has been obtained by Pew, suggests there is not yet even widespread familiarity with ChatGPT among teens. For those who actively follow news about generative AI developments, this may seem counterintuitive but can be understood as reflective of a diverse and geographically distributed populace that places its attention elsewhere. Unique to our research is the ability to conduct pre- and post-ChatGPT release comparisons to provide further evidence that cheating itself is not increasing – in part because it was already high. However, the behaviors that would be closer to plagiarism or "having the AI do all the work" did not exhibit significantly detectable changes.

We are hesitant to exert our value judgments on what should be policies for schools in the context of this specific study. Our view is that schools should determine how they wish to proceed, but gaining student perspective and data can be valuable. From this sample, students do not seem to be advocating for using such technology in ways that they consider to be inappropriate for schools, but their perspectives on when and how chatbot technology could be productively used may differ from their teachers and administrators. Students may think it is fine to use ChatGPT to explain something they do not understand, but that might be a cause for concern given propensities for the technology to "hallucinate" or be reliant on old information. At the same time, the ability for a student who is struggling with a topic introduced in their class to understand something better because they are able to get a new explanation from AI without being publicly visible about their confusion to their peers or teachers could be a boon for students. If anything, our view is that recent AI developments are forcing us to ask questions about why we require or limit the use of certain tools in schools and why we rely on certain assessment techniques in various circumstances. In some cases, we do value students being able to think through something without the aid of any digital technology present – and in situations like those, handwritten products or oral examinations might be useful. However,

there are times when we have generally accepted the use of technology as an aid – such as the use of spell checkers and grammar tools to ensure we have expressed our points clearly or so that we can focus on ideas rather than technical details about spelling and grammatical forms.

Increasing AI literacy – some understanding of how these AI technologies work, how to use them effectively, and awareness of risks and harms that can come from them – is something we do endorse for schools and educators. This technology appears to be here to stay, and as a matter of enabling participation and advancement given this will be increasingly prominent in many sectors of work and life, it is an imperative that we encourage frank conversations about this technology, our responsibilities, and our social expectations given its availability.

9. Limitations and future research

This survey study relied on pre-existing data to analyze changes before and after generative AI chatbot tools like ChatGPT became available. Some limitations include dependence on schools that contracted with CS and impacts from the shift to remote learning affecting the 2019–2020 and 2020–2021 school years. While additional data exist and are incoming, the goal of focusing on schools that could be compared on similar grounds before- and after-ChatGPT limited which schools could be included in this analysis.

Because of these self-imposed constraints, only some schools could participate in surveys at this time and in a timeframe given necessary ethical approvals. We stress that generalizations should be made cautiously as schools have different demographics and circumstances. Moreover, some decisions about instrument revision were made that come with trade-offs. For example, our notions of what might be unapproved technology in schools and what way those would likely be used necessitated changes in item wording that limit abilities to make broad generalizations. Because this was also part of a much larger survey, we were very mindful and restrained about adding new items. More qualitative research would be beneficial in addition to larger sampling and continued measurement at future time points, as behaviors and beliefs about allowability could change over time. Indeed, we anticipate they will, but this study provides data from one timepoint to aid the interpretation of future studies and samples.

10. Conclusion

Although AI chatbot technology has been a cause for alarm, especially with respect to students potentially using it for cheating and to produce writing to pass as their own, we do not see indications in this study that it has dramatically changed the prevalence of cheating – including specifically in behaviors that are ones that are thought to be most likely to be encouraged by AI chatbots. While this may change over time, this is an important reminder that concerns voiced in news headlines and individual anecdotes may look outsized relative to other approaches for examining prevalence and change. Also, it is important to note that this is in the context of high schools in the US. Things may be quite different with other populations, such as postsecondary students and in other nations. Future research, perhaps even using or adapting some of this instrumentation or referring to this study as a comparison, can help us collectively do the work to get a better sense of what is happening across these different settings and as we seek to have generative AI be an aid, rather than a harm, in education.

Ethics statement

The study was approved by a university institutional review board with ID: 14198. Informed consent was obtained from all participants, and their privacy rights were strictly observed. The participants were protected by hiding their personal information during the research process and be restricting information that can identify individuals or institutions. They knew that the participation was voluntary and they could withdraw from the study at any time, as well as skip any items. There is no potential conflict of interest in this study. The data, with necessary redactions for human subjects protections, can be requested via e-mail to the corresponding author.

CRediT authorship contribution statement

Victor R. Lee: Writing – review & editing, Writing – original draft, Supervision, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Denise Pope:** Writing – review & editing, Project administration, Methodology, Data curation. **Sarah Miles:** Writing – review & editing, Project administration, Methodology, Data curation. **Rosalía C. Zárate:** Writing – review & editing, Visualization, Software, Methodology, Formal analysis.

Declaration of competing interest

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

Acknowledgments

Funds supporting this work were provided by the Stanford Accelerator for Learning. Samantha Selby provided valuable assistance. We also thank Jeremy Roschelle and Bill Damon for earlier comments.

References

Anderman, E. M., & Midgley, C. (2004). Changes in self-reported academic cheating across the transition from middle school to high school. *Contemporary Educational Psychology, 29*(4), 499–517. https://doi.org/10.1016/j.cedpsych.2004.02.002

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of stochastic parrots: Can language models Be too big?. *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency, virtual event, Canada*. https://doi.org/10.1145/3442188.3445922. Paper presented at the.

Buehl, M. M., & Beck, J. S. (2014). The relationship between teachers' beliefs and teachers' practices. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 66–84). Routledge.

Common Sense Media. (2023). New poll finds parents lag behind kids on AI and want rules and reliable information to help them. Retrieved from https://www.commonsensemedia.org/press-releases/new-poll-finds-parents-lag-behind-kids-on-ai.

Conner, J., Galloway, M., & Pope, D. (2009). Student integrity in high-performing secondary schools. *Paper Presented at the 2009 Annual Meeting of the American Educational Research Association (AERA), San Diego, CA*.

Cross Francis, D. I. (2015). Dispelling the notion of inconsistencies in teachers' mathematics beliefs and practices: A 3-year case study. *Journal of Mathematics Teacher Education, 18*(2), 173–201. https://doi.org/10.1007/s10857-014-9276-5

Cross Francis, D., Rapacki, L., & Eker, A. (2014). The individual, the context, and practice: A review of the research on teachers' beliefs related to mathematics. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 336–352). Routledge.

Dunnigan, J., Henriksen, D., Mishra, P., & Lake, R. (2023). "Can we just please slow it all down?" School leaders take on ChatGPT. *TechTrends, 67*(6), 878–884. https://doi.org/10.1007/s11528-023-00914-1

Elsen-Rooney, M. (2023). *NYC education department blocks ChatGPT on school devices, networks*. New York: Chalkbeat. Retrieved from https://ny.chalkbeat.org/2023/1/3/23537987/nyc-schools-ban-chatgpt-writing-artificial-intelligence.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Galloway, M. K. (2012). Cheating in advantaged high schools: Prevalence, justifications, and possibilities for change. *Ethics & Behavior, 22*(5), 378–399.

Haddock, A., Ward, N., Yu, R., & O'Dea, N. (2022). Positive effects of digital technology use by adolescents: A scoping review of the literature. *International Journal of Environmental Research and Public Health, 19*(21), Article 14009. https://doi.org/10.3390/ijerph192114009. Nov.

Hargittai, E., & Shaw, A. (2015). Mind the skills gap: The role of internet know-how and gender in differentiated contributions to wikipedia. *Information, Communication & Society, 18*(4), 424–442. https://doi.org/10.1080/1369118X.2014.957711

Intelligent.com. (2023). One-third of college students used ChatGPT for schoolwork during the 2022-23 academic year. Retrieved from https://www.intelligent.com/one-third-of-college-students-used-chatgpt-for-schoolwork-during-the-2022-23-academic-year/.

Jenkins, B. D., Golding, J. M., Le Grand, A. M., Levi, M. M., & Pals, A. M. (2023). When opportunity knocks: College students' cheating amid the COVID-19 pandemic.

Teaching of Psychology, 50(4), 407–419. https://doi.org/10.1177/00986283211059067

Jensen, L. A., Arnett, J. J., Feldman, S. S., & Cauffman, E. (2002). It's wrong, but everybody does it: Academic dishonesty among high school and college students. *Contemporary Educational Psychology, 27*(2), 209–228.

Josephson Institute. (2012). *Josephson Institute, 2012 report Card on the Ethics of American youth, resource document*. Josephson Institute. https://charactercounts.org/wp-content/uploads/2014/02/ReportCard-2012-DataTables.pdf. retrieved 7/28/23.

Lee, V. R., Zarate, R. C., Pope, D., Miles, S., & Selby, S. (2024). Examining relationships between engagement and school culture in student cheating behaviors and AI chatbot usage. *Paper presented at the 2024 annual meeting of the*. Philadelphia, PA: American Educational Research Association.

Loftus, G. R., & Loftus, E. F. (1976). *Human memory: The processing of information*. Oxford, England: Lawrence Erlbaum.

Marcus, G., Davis, E., & Aaronson, S. (2022). *A very preliminary analysis of DALL-E 2. arXiv preprint arXiv:2204.13807*.

McCabe, D. (2001). *Cheating: Why students do it and how we can help them stop*. Winter: American Educator.

Miller, A. D., Murdock, T. B., Anderman, E. M., & Poindexter, A. L. (2007). 2 - who are all these cheaters? Characteristics of academically dishonest students. In E. M. Anderman, & T. B. Murdock (Eds.), *Psychology of academic cheating* (pp. 9–32). Burlington: Academic Press.

Murdock, T. B., Miller, A., & Kohlhardt, J. (2004). Effects of classroom context variables on high school students' judgments of the acceptibility and likelihood of cheating. *Journal of Education & Psychology, 96*, 765–777.

Nguyen, N., & Nadi, S. (2022). An empirical evaluation of GitHub copilot's code suggestions. *In the Proceedings of the 19th international Conference on mining software repositories*. https://doi.org/10.1145/3524842.3528470. Pittsburgh, Pennsylvania.

Noble, S. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: New York University Press.

Polly, D., & Hannafin, M. J. (2011). Examining how learner-centered professional development influences teachers' espoused and enacted practices. *The Journal of Educational Research, 104*(2), 120–130. https://doi.org/10.1080/00220671003636737

Sidoti, O., & Gottfried, J. (2023). About 1 in 5 U.S. teens who've heard of ChatGPT have used it for schoolwork [Press release]. Retrieved from https://www.pewresearch.org/short-reads/2023/11/16/about-1-in-5-us-teens-whove-heard-of-chatgpt-have-used-it-for-schoolwork/.

Stephens, J. M., & Wangaard, D. B. (2013). Using the epidemic of academic dishonesty as an opportunity for character education: A three-year mixed methods study (with mixed results). *Peabody Journal of Education, 88*(2), 159–179.

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*(1), 15. https://doi.org/10.1186/s40561-023-00237-x

Walton Family Foundation. (2023). Teachers and students embrace ChatGPT for education. Retrieved from https://www.waltonfamilyfoundation.org/learning/teachers-and-students-embrace-chatgpt-for-education.

Waltzer, T., DeBernardi, F. C., & Dahl, A. (2023). Student and teacher views on cheating in high school: Perceptions, evaluations, and decisions. *Journal of Research on Adolescence, 33*(1), 108–126. https://doi.org/10.1111/jora.12784

Wangaard, D. B., & Stephens, J. M. (2011). *Academic integrity: A critical challenge for schools*. Winter: Excellence & Ethics, 2011.

Yan, D., Fauss, M., Hao, J., & Cui, W. (2023). Detection of AI-generated essays in writing assessments. *Psychological Test and Assessment Modeling, 65*(1), 125–144.

Yazici, S., Yildiz Durak, H., Aksu Dünya, B., & Şentürk, B. (2023). Online versus face-to-face cheating: The prevalence of cheating behaviours during the pandemic compared to the pre-pandemic among Turkish University students. *Journal of Computer Assisted Learning, 39*(1), 231–254.

Zhao, L., Mao, H., Compton, B. J., Peng, J., Fu, G., Fang, F., … Lee, K. (2022). Academic dishonesty and its relations to peer cheating and culture: A meta-analysis of the perceived peer cheating effect. *Educational Research Review, 36*, Article 100455. https://doi.org/10.1016/j.edurev.2022.100455

Glossary

***CS**: Challenge Success*