

RESEARCH ARTICLE

Open Access



Not quite eye to A.I.: student and teacher perspectives on the use of generative artificial intelligence in the writing process

Alex Barrett^{1*}  and Austin Pack²

*Correspondence:
abarrett3@fsu.edu

¹ College of Education, Florida State University, Stone Building, 114 West Call Street, Tallahassee, FL 32306, USA

² Faculty of Education and Social Work, Brigham Young University-Hawaii, 55-220 Kulanui Street, Laie, HI 96762, USA

Abstract

Generative artificial intelligence (GenAI) can be used to author academic texts at a similar level to what humans are capable of, causing concern about its misuse in education. Addressing the role of GenAI in teaching and learning has become an urgent task. This study reports the results of a survey comparing educators' (n = 68) and university students' (n = 158) perceptions on the appropriate use of GenAI in the writing process. The survey included representations of user prompts and output from ChatGPT, a GenAI chatbot, for each of six tasks of the writing process (brainstorming, outlining, writing, revising, feedback, and evaluating). Survey respondents were asked to differentiate between various uses of GenAI for these tasks, which were divided between student and teacher use. Results indicate minor disagreement between students and teachers on acceptable use of GenAI tools in the writing process, as well as classroom and institutional-level lack of preparedness for GenAI. These results imply the need for explicit guidelines and teacher professional development on the use of GenAI in educational contexts. This study can contribute to evidence-based guidelines on the integration of GenAI in teaching and learning.

Keywords: Artificial intelligence, Large language model, GPT, Writing education, Academic integrity

Introduction

Public interest in artificial intelligence (AI) has grown substantially as a result of recent public access to large language models (LLMs; e.g., OpenAI's GPT-3 and 4, Google's PaLM 1 and 2), and chatbots (e.g., OpenAI's ChatGPT, Google's Bard, Microsoft's Bing) that allow users to interface with LLMs. These Generative AI (GenAI) tools afford individuals with the ability to instantly generate writing on any topic by inputting a simple prompt. The public discourse surrounding GenAI has been mostly positive, but in the education sector there is serious concern about academic integrity and plagiarism (Dehouche, 2021; Lampropoulos et al., 2023; Sullivan et al., 2023; Yeo, 2023). Some schools have responded by banning the technology outright (Yang, 2023), a move likened by some to the banning of the pocket calculator when it was perceived as a threat to math education (Urlaub & Dessein, 2022). What

is clear is that this new technology possesses disruptive potential and that institutions which have relied heavily on student writing for education and assessment will need to respond accordingly.

Although a few schools have banned ChatGPT and similar tools, many have not, displaying confidence that their institution's academic integrity policy is robust enough to accommodate the new technology. However, current definitions of plagiarism have been described as medieval (Dehouche, 2021; Sadeghi, 2019), typically including language such as kidnapping, stealing or misappropriating the work of *others* (Sutherland-Smith, 2005), which now leads us to question whether a chatbot counts as one of these others. Generative AI is trained on a selection of diverse natural language data from across the Internet which allows it to string together unique combinations of words and phrases, similar to how humans learn to produce an unlimited amount of novel spoken or written text from the limited language they absorb from their environment, a tenet of generative grammar (Chomsky, 1991). The result is that there is no identifiable *other* whose work is being stolen by a chatbot. To complicate matters, the language of OpenAI's Terms of Use state that it assigns users "all its right, title and interest in and to Output" from ChatGPT, including for purposes of publication (OpenAI, 2023). Any practiced educator would likely agree that submitting an essay written by ChatGPT without disclosure violates academic integrity, but students may not readily see a problem with it.

Although GenAI has multiple applications, its use as an authoring tool in programs like ChatGPT allow for easy misuse. Students who have purposefully violated academic integrity in the past through the use of contract cheating or paper mills will likely not hesitate to use ChatGPT or other GenAI tools to do so now, but other students will need guidance on how to avoid inadvertently cheating. Student perceptions of academic dishonesty have historically been unclear or incomprehensive, and rarely align with teacher expectations (Tatum, 2022), GenAI will only serve to complicate this (Farrokhnia et al., 2023).

Some advocate working towards a coexistence with AI in education by establishing common goals and guided exploration of the limitations of the technology (Godwin-Jones, 2022; Tseng & Warschauer, 2023). Yeo (2023) has specifically recommended the exploration of student perceptions about the ethics of using GenAI tools, and Pack and Maloney (2023a) suggested teacher and researcher use should also be investigated.

To date no consensus has arisen regarding what constitutes appropriate use of GenAI in higher education. Therefore, with the goal of identifying some common expectations, the purpose of this study is to explore student and teacher perspectives of using GenAI for various tasks in the writing process, including brainstorming, outlining, writing, and revising done by students, and evaluation and feedback done by teachers. The research questions guiding the study are:

1. What are undergraduate students' and teachers' perspectives on using GenAI in the writing process (brainstorming, outlining, writing, revising, evaluation, and feedback)?

2. How do student and teacher perspectives on the use of GenAI in the writing process compare?

Related literature

Writing instruction and assessment have made use of AI for some time in the form of programs like Grammarly or spell checker that identify grammatical and lexical errors in writing (Godwin-Jones, 2022). Yet, with recent advances in machine learning and natural language processing, AI-integrated language tools now pose a considerable challenge for educational systems that have relied heavily on writing to develop and assess the cognitive and linguistic competencies of learners. Students who do not use English as a first language can now use machine translation programs to accurately render their native language writing into a target language (Godwin-Jones, 2022); students can use programs like Quillbot or Wordtune to paraphrase, summarize, or adjust the tone of a paragraph (Yeo, 2023); and they can use ChatGPT to instantly generate entire essays. Authoring tools such as ChatGPT are prevalent and affordable (or even free) to anyone with an internet connection. ChatGPT garnered over 100 million users within three months of its launch (Hu, 2023), many of which are likely students and teachers. How GenAI, like ChatGPT, is being used in academic settings and how much reform will be required of educational systems as a result of this use is yet to be determined.

Looking at how individuals and institutions have used AI-integrated writing and language tools in the past might inform predictions of how GenAI tools like ChatGPT will be used in the future. For instance, automated essay scoring with AI has been the status quo for many large testing services, such as Pearson's Intelligent Essay Assessor and ETS's e-Rater (Gardner et al., 2020), which administer university placement exams, and other major exams such as the GRE and TOEFL. Receiving millions of summative writing samples each year, AI can greatly reduce the workload of manually scoring each essay (Hockley, 2018). However, the programs are typically limited to assessment of grammar, usage, mechanics, and style and are not able to detect more complex features like the presence of a thesis statement or overall coherence (Gardner et al., 2020). GenAI might be used to strengthen these tools to assess more complex discourse elements.

For formative writing, on the other hand, tools like Grammarly or Criterion are commonly used to provide corrective feedback during the revising and editing phases of the writing process. This includes basic grammar, spelling, and punctuation, or more advanced analytics such as word counts and readability (Fitria, 2021), all of which can be procured instantly and at any time. Multiple studies have researched automated writing evaluation (AWE) tools such as these with mixed results as to their efficacy and reliability (Huawei & Aryadoust, 2023; Wang & Han, 2022; Zhang, 2020), however, students tend to appreciate the feedback and respond positively to it (O'Neill & Russell, 2019). ChatGPT may be used as a cheaper and more robust AWE tool by students.

A recent experimental study by Fan (2023) looked at the effects and perceptions of university EFL students' use of Grammarly for corrective feedback on their writing. Although no differences were noticed in terms of writing improvement, most students in the Grammarly group found the feedback to be understandable and useful. For those that did not, Fan (2023) noted that students' low proficiency may have prevented them

from effectively understanding Grammarly's feedback. One potential use of ChatGPT by EFL students is to revise commercial AWE feedback to be more comprehensible and accessible for them.

Although GenAI tools can be useful as a virtual tutor that offers individualized writing feedback, the threat of misuse sours these benefits. AWE tools like Grammarly and Criterion tend to be useful for evaluation and revision stages of the writing process, but GenAI tools like ChatGPT can further be used in the brainstorming, outlining, and writing stages of the writing process. The concern of authorship and plagiarism arises when students are using GenAI to give them ideas or to draft writing assignments (Ingle & Pack, 2023). This is the concern voiced by many in academia since the release of ChatGPT in November of 2022 (Sullivan et al., 2023).

Writing is a fundamental skill that is necessary for learner academic and professional development. In a joint publication, the Council of Writing Program Administrators, the National Council of Teachers of English, and the National Writing Project described the benefits of writing practices to develop rhetorical knowledge and critical thinking, which are in turn supported by habits of mind such as curiosity, creativity, persistence, and responsibility (CWPA, NCTE, & NWP, 2011). It has been shown that weak writing ability can result in less learning in all school subjects and negatively impact professional success (Graham, 2019; Graham et al., 2020). Therefore, it stands to reason that students outsourcing writing to AI will likely incur a negative impact on their cognitive ability and prospects for future success.

Concerns have also been voiced in how AI may be used by educators (Pack & Maloney, 2023a, b; Carlson et al., 2023; Lo, 2023). Educators have used GenAI to create course material and assessment tasks, adapt materials to be more suitable for specific students, and generate lecture notes (Bonner et al., 2023; Lo, 2023). However, one of the most frequently suggested uses of GenAI for teachers is to grade and provide feedback on student writing (Chiu et al., 2023; Kaplan & Haenlein, 2018; Weigle, 2013; Yeo, 2023). Accompanying some of these suggestions are characterizations of grading writing as burdensome and tedious, and GenAI is seen as a way of reducing teacher workload.

Teachers who seek to escape the 'tedious and burdensome' process of essay grading by using GenAI may inadvertently be signaling that it is ok for students to use the tool to "take the pain out of...the writing process" (Yeo, 2023, p. 2). In a position paper on machine essay scoring, the National Council of Teachers of English (2013) highlighted the social aspect of writing and that machine scoring sends students a message that writing is not worth the time because reading it is not worth the time. It will be imperative to incorporate GenAI in the writing process in a way that alleviates the "pain" and "burden" of the process without diminishing the social nature of writing.

Teachers might also be misled by the seeming impartiality of AI tools when the truth is these instruments are susceptible to influences unwittingly included in them by the developers who created them, a phenomenon known as algorithmic bias (Jackson, 2021). In educational contexts, algorithmic bias in GenAI can manifest when LLMs are trained on a convenience sample of language from the Internet which tends to be majority English language and majority western (Graham et al., 2015), resulting in underrepresentation of different languages, dialects, philosophies, ethnicities, and

multiple other demographic divisors (Baker & Hawn, 2021). A study by Bridgeman et al. (2012), for example, showed that the e-Rater essay scoring AI produced inaccurate scores for populations along ethnic lines.

The use of GenAI for education involves other risks such as functional opacity, data privacy, and reliability (Yu & Guo, 2023) which need to be accounted for, but the usefulness of the technology for assisting in the writing process will likely lead to its adoption by teachers and students regardless of these limitations. To what degree do students and teachers feel it is acceptable to use GenAI in writing? Addressing this research question will provide a foundation from which common goals can be established that might assist with the successful adoption of GenAI in education.

Methodology

This study utilized a cross-sectional questionnaire design to make within- and between-group comparisons of students' and teachers' perceptions of the use of GenAI in learning and teaching.

Participants

A total of 226 participants (158 students and 68 teachers) completed the questionnaire. Students from across a variety of disciplines were recruited via non-probability voluntary response from a public research university in the United States with an incentive of course credit. Teachers were recruited using a purposeful sampling and snowballing method whereby the questionnaire was sent out to contacts of the authors at multiple institutions worldwide for distribution. The teacher questionnaire also provided a self-reflexive URL link for those who completed it with a request to forward the link to relevant individuals.

Second language teachers that teach a language different than the students' native language were the primary target of the teacher survey due to their experience with ground-up writing instruction and the corresponding need to address plagiarism and academic dishonesty regularly. Insights from English as a Second Language (ESL) and English as a Foreign Language (EFL) teachers may prove useful as they are exposed to non-western writing traditions which hold diverse views on intellectual property and originality (Pennycook, 1996; Sutherland-Smith, 2005). The student population was chosen for its homogeneity and typicality of public university undergraduates. A minority of surveys were collected from graduate students ($n = 5$) and non language educators ($n = 3$).

Informed consent from all participants was acquired and all ethical procedures were adhered to according to the standards of the university institutional review board. Participant demographics can be found in Tables 1 and 2.

Instrument

Divisions of use of AI in the writing process

A quantitative questionnaire instrument was developed to measure participants' perspectives on ways of including GenAI in the writing process. According to Seow (2002), process writing includes four basic stages: planning (including brainstorming and outlining), writing, and revising or editing. Planning includes prewriting activities that assist students in generating and organizing ideas. During the writing stage

Table 1 Student demographics (n = 158)

Category	Value	Frequency	% (rounded)
Sex	Male	57	36
	Female	99	63
	Non-binary/transgender	2	1
Age	18–24	151	96
	25–34	7	4
Degree level	Undergraduate	153	97
	Graduate	5	3
Highest degree obtained	High school	89	56
	Associate's	53	34
	Bachelor's	16	10
Language	English native	151	96
	English non-native	7	4

Table 2 Teacher demographics (n = 68)

Category	Value	Frequency	% (rounded)
Sex	Male	25	37
	Female	43	63
Age	18–24	1	2
	25–34	23	34
	35–44	22	32
	45–54	13	19
	55–64	5	7
	65+	4	6
Education level	Associate's (2-year degree) or lower	2	3
	Bachelor's	3	4
	Master's	46	68
	Doctoral	17	25
Years of experience	Less than 1	2	3
	1 to 3	5	7
	4 to 6	13	19
	7 or more	48	71
Country currently teaching in	United States	38	56
	China	14	21
	Other	15	23
Level currently teaching	Adult/university	63	93
	K-12	2	3
	Other	3	4
Subject	English as a new language (ENL)	55	81
	Modern languages	10	15
	Other (i.e., Health, Psychology, Law)	3	4

students focus on communicating their ideas to a specific audience in an initial draft. Based on feedback from the initial draft, revising occurs as students reexamine their writing and rewrite areas where their intent, style, tone, mechanics, or organization was identified as needing improvement. In addition to these steps that students

complete, Seow highlights the importance of evaluation and feedback from teachers throughout the writing process. The questionnaire used in this study explored participants’ perspectives on potential uses of GenAI for brainstorming, outlining, writing, revising, evaluating, and providing feedback.

For each of the writing process steps participants were asked to read an example GenAI prompt and response produced by OpenAI’s ChatGPT (GPT-3.5 turbo) obtained in February of 2023. They were then presented with four divisions of use (or misuse) and asked to rate the appropriateness of each division on a five-point Likert scale (strongly disagree to strongly agree). These divisions and their explanations can be found in Table 3. The internal consistency of this instrument was measured with the alpha coefficient for each division of use of AI in the writing process. All alpha coefficients were above 0.7 which is considered satisfactory (Bland & Altman, 1997). These divisions of use were conceptualized by considering the common suggested uses of GenAI in recent literature (e.g., Kasneci et al., 2023), which include things like cognitive offloading of trivial tasks and idea generation, and also incorporating theory of plagiarism degrees of severity (Evering & Moorman, 2012; Yeo, 2007; Yeo & Chien, 2007), which differentiates plagiaristic behaviors.

Each division of use represented in the instrument was accompanied by an example of GenAI output that could be generated with simplistic prompting (using ChatGPT 3.5-turbo, February 2023 version). The simplistic approach taken when creating the examples was adopted to portray the capabilities of the technology from a layperson perspective, who may approach prompting naively. No sophisticated or iterative prompt engineering was used when generating the output for these examples but each prompt was submitted to ChatGPT separately so that the chatbot would not have a direct memory resource of previously submitted prompts. All complete prompts and output can be found in the appendix.

We divided the examples of GenAI use in the writing process between students and teachers so that brainstorming, outlining, writing, and revising were student actions and evaluating and providing feedback were teaching actions. In doing this we do not intend to suggest students cannot or should not use GenAI for evaluation or feedback purposes, but to demonstrate to participants that GenAI use in the writing process is

Table 3 Divisions of use of AI for the writing process as worded in the questionnaire, with explanations

A	B	C	D
“This use of AI is appropriate if the person already knows how to perform the task.”	“This use of AI is appropriate if the person uses the output to generate ideas about the task, but completes the task without further assistance from AI.”	“This use of AI is appropriate if the person submits the output for the task assignment and discloses the use of AI.”	“This use of AI is appropriate if the person submits the output for the task assignment but does not disclose the use of AI.”
Explanation: If a student or teacher is tasked with an assignment that they can already perform with a high degree of competency, then it is acceptable to offload that task to AI	Explanation: It is acceptable for students or teachers to use AI to generate models or ideas for task assignments, considering the submitted assignment is composed of their own language and not the language of the AI	Explanation: It is acceptable for students or teachers to submit task assignments that contain, in part or in whole, language produced by AI, considering that any AI language is clearly identified as such	Explanation: It is acceptable for students or teachers to submit task assignments that contain, in part or in whole, language produced by AI without disclosing the use of AI

not restricted to only students, and to elicit broader perceptions on the use of GenAI in education by multiple stakeholders.

Looking at the student-oriented prompts, for brainstorming we prompted ChatGPT to come up with ideas for an essay on the topic of urban challenges and global warming and the output provided 8 relevant ideas. The outlining prompt requested an outline for a 5-paragraph essay on the same topic and returned a bulleted outline including a thesis statement, paragraph topics, and supporting points. For the writing prompt we requested a fully written 5-paragraph essay on the same topic and it returned a coherent and cohesive essay. For revision we provided ChatGPT with a four-sentence paragraph that was written in an informal tone and asked for it to be revised to be more academic; the resulting output was more formal and academic.

As for the teacher-oriented prompts, we provided ChatGPT with a short paragraph that was replete with grammatical and lexical errors and prompted it to provide suggestions to the student on how to improve their writing. The output provided five somewhat generic suggestions for improvement along with examples, such as using transitional phrases, clearer language, and proofreading for errors. Last, we provided an error-free paragraph and prompted ChatGPT to evaluate the quality of ideas expressed therein. ChatGPT returned several sentences evaluating the argument made in the paragraph, commenting on development and logic.

Survey

Following the questionnaire section measuring perceptions of acceptable use, a short survey was included to better understand the sampled populations and to detect potential covariates or subgroups which might afford further analysis. These included eight survey items covering opinions about AI and technology in general measured on a five-point Likert scale (strongly disagree to strongly agree), and two to three (depending on group) yes/no response items (all items are listed in Table 7 in Sect. "Survey results").

To investigate the validity of the eight survey items, an exploratory factor analysis was conducted with all the completed surveys (n=226) using principal axis factoring and a promax rotation method. The KMO test value was 0.674, showing an adequate proportion of variance in the survey which could indicate underlying factors. Also, Bartlett’s Test of Sphericity was significant at the 0.001 alpha level. The analysis confirmed a four-factor solution (Table 4) which cumulatively accounted for 77.6% of the variance.

Table 4 Exploratory factor analysis loadings

	1	2	3	4
AI will be a useful tool for students	0.927			
AI will be a useful tool for teachers	0.766			
I think AI will have a positive impact on education	0.638			
I am concerned with how students will use AI		0.874		
I am concerned with how teachers will use AI		0.545		
I enjoy using new technologies to help me learn			0.711	
I prefer to stick with tools and methods of learning I am familiar with			- 0.497	
I am familiar with how AI may be used for education				0.574
Factor alpha coefficient (α)	0.82	0.63	0.45	-

However, factors 2 and 3 showed poor internal consistency, with alpha coefficients below the 0.7 benchmark. The low reliability of factors 2 and 3 is likely due to only having two items per factor and also disparate attitudes concerning teachers and students from the surveyed groups. Factor 1 can be described as the perceived utility of AI in education; factor 2 as perceived concern about AI in education; factor 3 as perceptions on technology change and innovation; and factor 4 as familiarity with AI.

Data collection and analysis

The questionnaire was created and distributed using Qualtrics®. Questionnaire responses were collected during the spring semester of 2023. 15 incomplete questionnaires were discarded following the collection period.

Response frequencies were calculated for both groups on the levels of acceptable use of GenAI across each task and for the yes/no survey items. Mean responses and standard deviations were calculated for levels of acceptable use as well as the Likert-style survey items.

Although our teacher and student samples were disparate in size and population, exploratory comparisons were made to infer differences in perceptions. To do this, the Mann–Whitney U test was used to compare means for the teachers' and students' responses on the level of appropriateness of using GenAI, and also for mean responses to the survey items. A key assumption of the Mann–Whitney U test is independent observations from compared groups. This assumption was satisfied as participants completed the questionnaire individually and cross-group contamination was unlikely due to separate questionnaire hyperlinks.

Lastly, a principal component analysis was conducted to better understand teacher and student perceptions of GenAI use in the writing process across the multiple writing process steps and divisions of use.

Results

In general, both teachers and students held similar perceptions on what is appropriate use of GenAI in the writing process. That is, both groups predominantly agreed or disagreed along each division of use for each of the writing process tasks presented to them. Despite this general conformity, there were some significant differences in mean responses for some of the divisions of use as measured by the Mann–Whitney U test. Details of these response frequencies are provided in Table 5, and student–teacher comparisons are presented in subsequent subsections.

Brainstorming

Figure 1 shows response frequencies for teachers and students for the brainstorming task of the writing process. Students and teachers both generally felt that using GenAI to brainstorm ideas was acceptable if the student was already a competent brainstormer or only used the output as a model. Submitting AI-brainstormed ideas in class was seen as acceptable by half of teachers and students who took the survey, with another 10 to 16% uncertain and the remaining against. No significant differences were found between groups on these uses of GenAI. However, there was a larger difference between teachers and students when asked if it was ok to use GenAI to brainstorm ideas without

Table 5 Mean and standard deviation for student (n=158) and teacher (n=68) responses to the use of GenAI in the writing process across four divisions of use (A, B, C, D)

Step in writing process	Group	A	B	C	D
Brainstorming	Student	3.55 (1.17)	3.82 (1.15)	3.13 (1.31)	1.99 (1.22)
	Teacher	3.53 (1.26)	3.78 (1.18)	3.21 (1.38)	1.65 (1.38)
Outlining	Student	3.27 (1.29)	3.42 (1.28)	2.82 (1.40)	1.84 (1.19)
	Teacher	2.96 (1.42)	2.76 (1.37)	2.41 (1.28)	1.34 (0.86)
Writing	Student	2.44 (1.47)	2.91 (1.42)	2.25 (1.43)	1.62 (1.12)
	Teacher	2.13 (1.48)	3.15 (1.30)	1.68 (1.04)	1.22 (0.79)
Revision	Student	2.80 (1.41)	3.54 (1.34)	2.62 (1.47)	1.92 (1.24)
	Teacher	2.53 (1.46)	3.67 (1.22)	2.10 (1.19)	1.44 (1.07)
Feedback	Student	2.98 (1.38)	3.23 (1.37)	2.80 (1.39)	2.01 (1.23)
	Teacher	3.23 (1.27)	3.71 (1.11)	3.03 (1.28)	1.82 (1.28)
Evaluation	Student	2.73 (1.38)	3.11 (1.43)	2.63 (1.42)	1.89 (1.31)
	Teacher	3.07 (1.29)	3.54 (1.14)	2.85 (1.39)	1.17 (1.17)
α	Student	0.87	0.78	0.85	0.91
	Teacher	0.84	0.79	0.83	0.90

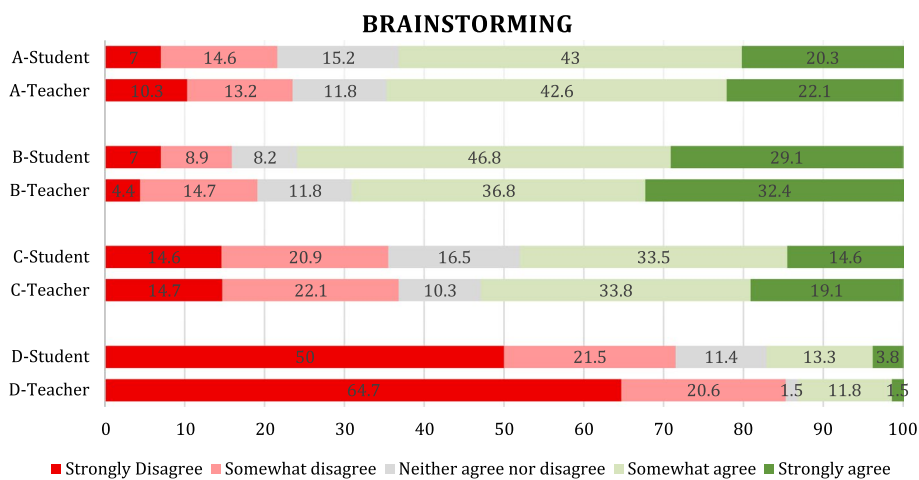


Fig. 1 Response frequencies for brainstorming task

disclosing the use of GenAI. Although teachers and students were predominantly in alignment in their disagreement that this use was acceptable, fewer students disagreed, and about 11% were uncertain. A Mann–Whitney U test comparing means between groups found a *p*-value of 0.032 ($U=4491, r=-0.143$) for division D on the brainstorming task.

Outlining

The outlining task resulted in more differences between student and teacher perceptions (Fig. 2). Means for division A of GenAI use (acceptable if the student is already a competent outliner) were not significantly different between groups, but for divisions B, C, and D significant differences were detected with students being more accepting of these uses. For using the GenAI output as a model, students were more accepting ($U=3949.5,$

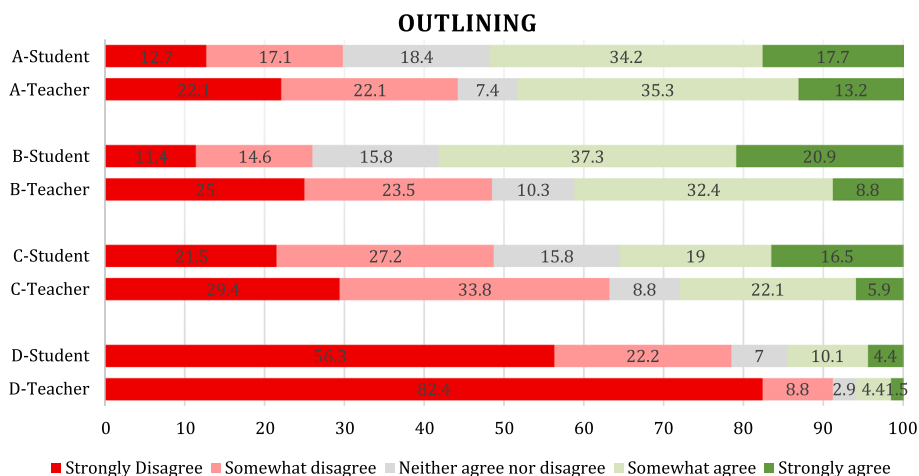


Fig. 2 Response frequencies for outlining task

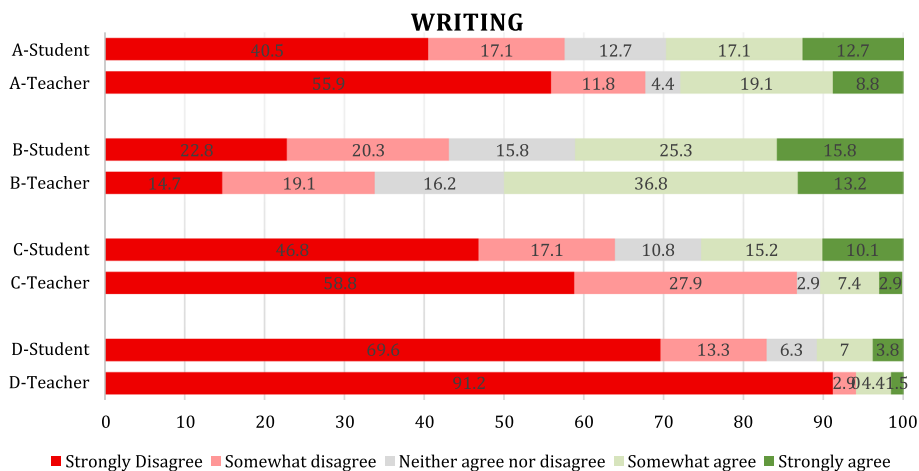


Fig. 3 Response frequencies for writing task

$r = -0.217, p = 0.001$). For submitting GenAI output with disclosure ($U = 4488.5, r = -0.134, p < 0.05$) and without disclosure ($U = 3973.5, r = -2.42, p < 0.001$) students were also more accepting than teachers.

Writing

Teachers and students had comparable perceptions on acceptable use of GenAI for writing an essay (Fig. 3) if the student was already competent in writing an essay (Division A pair) or if a student wanted to use a GenAI draft of an essay to model their own writing (Division B pair). Interestingly, both students and teachers predominantly viewed the use of GenAI for writing essays, even when the student is a competent writer, as inappropriate, but both students and teachers mostly agreed it was ok for students to use a GenAI generated essay as a model. Significant disagreement was found between teachers and students regarding submitting a GenAI-written essay with disclosure ($U = 4311.5,$

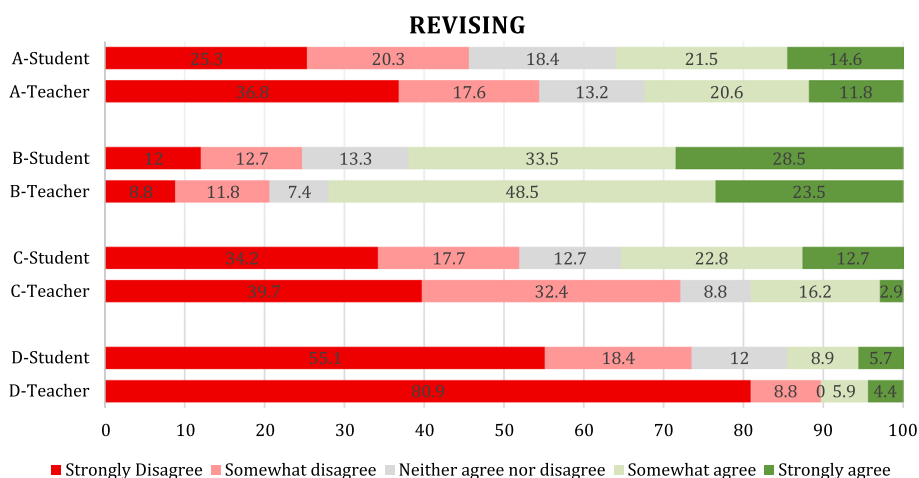


Fig. 4 Response frequencies for revision task

$r = -0.169, p = 0.011$) and without disclosure ($U = 4245.5, r = -0.223, p < 0.001$), with teachers more heavily disagreeing with this behavior.

Revision

Similar to writing, having GenAI revise an essay (Fig. 4) showed mixed perceptions between teachers and students with no differences detected for divisions A and B, but significant differences found in divisions C and D. Students were more accepting of using GenAI to revise their writing both with disclosure ($U = 4393, r = -0.15, p < 0.05$) and without disclosure ($U = 4030, r = -0.23, p < 0.001$). However, like the writing task, the majority of both groups saw this as inappropriate use of AI.

Feedback

Feedback (Fig. 5) was framed as a teacher use of GenAI as an AWE tool. Non-significant differences in mean response frequencies were found along divisions A, C, and D, however, teachers agreed significantly more than students that division B was acceptable (using GenAI generated feedback as a model) ($U = 6397, r = -0.157, p < 0.05$), although both groups were generally accepting of this behavior. Heavy disagreement was reported by both groups in using GenAI for providing writing feedback without disclosing the use of GenAI.

Evaluating

Using GenAI for evaluation of student writing was also framed as a teacher task (Fig. 6). Similar to the feedback task, significant differences on perceptions of acceptable use were only detected in division B (using the AI-generated evaluation as a model), again with teachers being more accepting of this use ($U = 6247, r = -0.133, p < 0.05$). Again, both teachers and students felt it was inappropriate to use GenAI for this purpose without disclosing the use of GenAI.

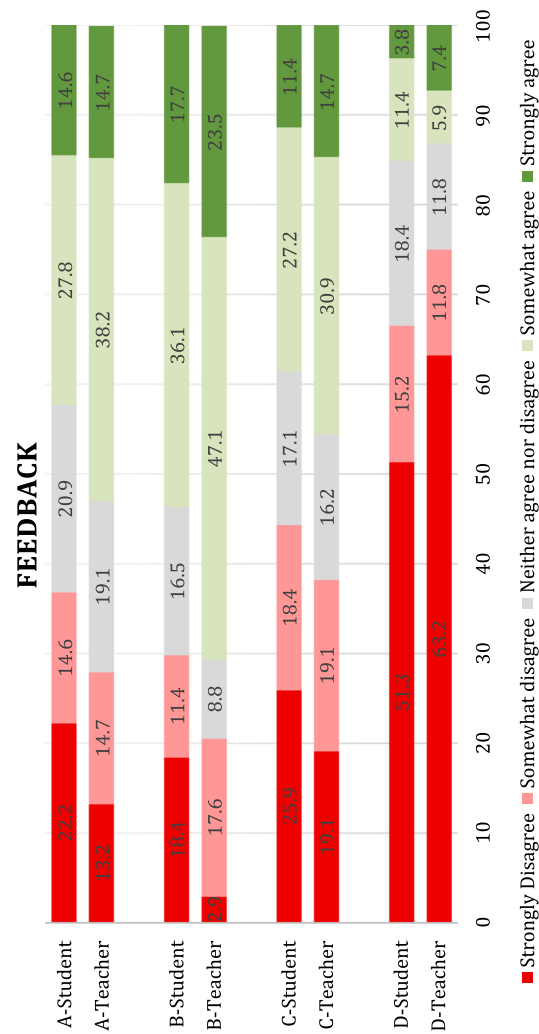


Fig. 5 Response frequencies for feedback task

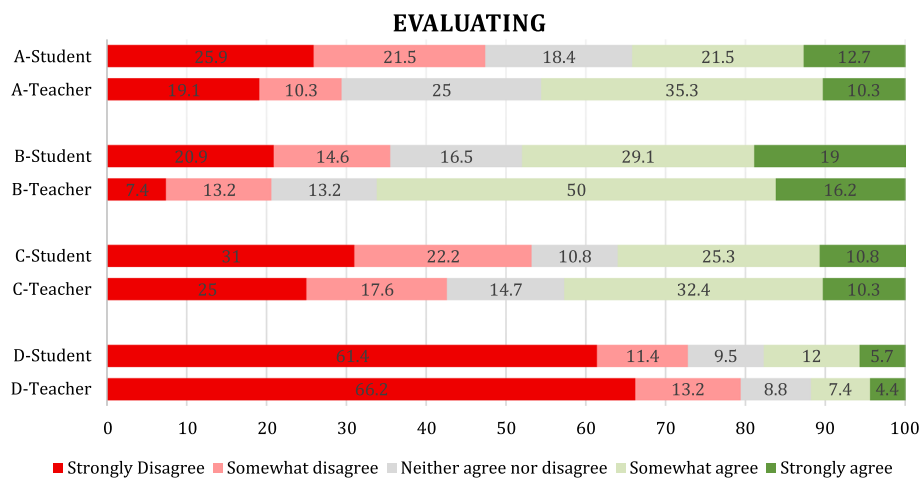


Fig. 6 Response frequencies for evaluation task

Table 6 Principal component analysis

	1	2	3
Writing D	0.846		
Outlining D	0.868		
Revision D	0.796		
Brainstorming D	0.749		
Writing C	0.747		
Evaluate D	0.695		
Feedback D	0.633		
Revision C	0.622		
Outline C	0.502		
Outlining A		0.747	
Outlining B		0.733	
Brainstorming A		0.677	
Brainstorming B		0.666	
Revision B		0.652	
Revision A		0.640	
Writing B		0.550	
Writing A		0.545	
Brainstorming C		0.322	
Evaluate B			0.766
Evaluate C			0.762
Feedback C			0.752
Evaluate A			0.747
Feedback B			0.742
Feedback A			0.725

Principal component analysis

In order to better understand how teachers and students perceive acceptable use of GenAI in the writing process a principal component analysis was conducted (Table 6). Dimension reduction was achieved with a varimax rotation specifying 3 factors, identified from components with an eigenvalue of 2 or greater on a scree plot, which

accounted for 60% of the variance. The KMI measure of sampling adequacy value was 0.862 and Bartlett's Test of Sphericity was significant ($p < 0.001$). For variables with loadings in more than one factor, the smaller loading was suppressed.

Factor 1 included all of division D as well as writing, outlining, and revising for division C. These use examples all had the lowest mean agreement of acceptability (refer to Table 5) so we can label this factor as highly unacceptable use. Factor 2, on the other hand, primarily contained divisions of use A and B for outlining, revision, brainstorming, and writing. These use examples had relatively high means and were seen as generally permissible uses of GenAI, despite writing A having a majority disagreement response frequency. Factor 3 contained evaluation and feedback for all A, B and C divisions of use. The ratings for these use examples were mostly supportive, especially along the A and B divisions of use.

This analysis supports the validity of the instrument. Factor 1 contains the divisions of use that involve GenAI doing all the writing of a writing assignment, which was rated as the most unacceptable use of GenAI. Factor 2 contained the most acceptable uses of GenAI across four student-oriented steps of the writing process, which included utilizing GenAI for tasks that the user is already proficient in, and to generate ideas or model answers. The factor 3 use examples were categorized as being teacher-oriented, and their inclusion in one factor here demonstrates that participants conceptualized acceptable teacher use of GenAI differently.

Survey results

In addition to measuring perceptions on the use of GenAI in the writing process, we also included several survey items to measure other aspects of AI use in educational contexts (Table 7).

We asked questions to better understand perceptions about the utility of AI in education (items 1 through 3). There was a tendency to agree that AI would be useful to students and teachers in education, although some trepidation can be insinuated from relatively less agreement (and neutrality for students) to item 3, about AI having a positive impact on education. Some disagreement is evident between teachers and students for item 2, about the utility of AI for teachers, with a significantly higher percentage of teachers agreeing on AI's utility compared to students ($U = 7098, p < 0.001$).

Concern for student use of AI (items 4 and 5) was fairly high for both teachers and students, with slightly less concern for teacher use. Students showed significantly more concern regarding teacher use of AI than did teachers ($U = 4288.5, p < 0.05$).

We asked two questions to get a sense of participants' general feelings toward new technology and innovation (items 6 and 7). Teachers reported significantly more openness to the use of new technologies and innovative tools and methods in their teaching than did students for their learning ($U = 6441.5, p < 0.05$; $U = 3577.5, p < 0.001$).

Item 8 inquired about participant familiarity with AI. Mean familiarity scores were not high, but students reported slightly more familiarity with AI than did teachers, however the difference was non-significant.

We asked three yes/no questions (items 9 through 11; Table 8) about AI policy and preparedness. The student group (who all attend the same university) reported mixed

Table 7 Survey item total response percentages (rounded) and mean comparisons between students (n = 158) and teachers (n = 68)

#	Item	S/T	SD	D	N	A	SA	M (SD)	U	P
1	AI will be a useful tool for students	S	10	11	19	43	17	3.47 (1.18)	5914.5	0.201
		T	3	12	13	56	16	3.71 (0.98)		
2	AI will be a useful tool for teachers	S	11	11	23	41	13	3.32 (1.18)	7098.0	<0.001
		T	0	6	10	63	21	3.99 (0.74)		
3	I think AI will have a positive impact on education	S	13	21	28	30	7	2.96 (1.15)	6084.5	0.101
		T	4	16	37	37	6	3.24 (0.95)		
4	I am concerned with how students will use AI	S	5	4	8	30	53	4.23 (1.08)	5522.0	0.713
		T	3	3	7	32	54	4.32 (0.95)		
5	I am concerned with how teachers will use AI	S	4	10	17	30	39	3.89 (1.16)	4288.5	0.012
		T	4	16	24	35	21	3.51 (1.13)		
6	I enjoy using new technologies for teaching [to help me learn]*	S	2	10	21	45	23	3.76 (0.98)	6441.5	0.012
		T	2	3	21	34	41	4.10 (0.93)		
7	I prefer to stick with tools and methods [of learning]* I am familiar with	S	3	17	23	36	21	3.58 (1.08)	3577.5	<0.001
		T	10	27	29	31	3	2.90 (1.05)		
8	I am familiar with how AI may be used for education	S	9	18	13	47	13	3.36 (1.18)	4728.5	0.133
		T	6	31	18	35	10	3.13 (1.15)		

S student, T teacher, SD strongly disagree, D disagree, N neither agree nor disagree, A agree, SA strongly agree

*Bracketed text was displayed to the student group only

Table 8 Yes/no item response percentages

#	Item	Student	Teacher
9	My school or university has a policy in place regarding the use of AI	YES = 52% NO = 47%	YES = 5.9% NO = 94.1%
10	My school or university has provided me with training about the use of AI	YES = 5.7% NO = 94.3%	YES = 4.4% NO = 95.6%
11	I have educated my students about the appropriate use of AI	N/A	YES = 10.3% NO = 89.7%

answers when asked if their university had a policy on AI use. Teachers (who affiliated with various institutions), by comparison, expressed much more certainty that their institutions did not have an AI policy. Around 95% of both groups reported that they had received no training on the use of AI, and about 90% of teachers have not taught their students about the appropriate use of AI.

Discussion

This study sought to better understand how student and teachers perceive of the use of GenAI in the writing process within a framework of acceptability. The goal of this research is to contribute to the burgeoning discussion on how GenAI can be integrated into educational contexts successfully (see Godwin-Jones, 2022; Yeo, 2023). The prevailing narrative in our results demonstrated that for all the steps of the writing process, students and teachers generally agreed that using GenAI to brainstorm ideas or model answers, or as a form of cognitive offloading for tasks that the user is already competent in, is acceptable. Conversely, using GenAI to complete writing task assignments, with or without disclosing the use of GenAI, is unacceptable.

Both students and teachers perceived GenAI use to be more acceptable in the early stages of the writing process (i.e., brainstorming and outlining) than in later stages. These results suggest that use of GenAI for writing purposes is viewed as more acceptable when it is fulfilling a supportive role focused on idea generation and organization rather than when leveraged as an automatic writing completion tool. As to differences in perspectives, students tended to disagree less than teachers that using GenAI without disclosure was appropriate, and teachers tended to disagree less than students about using GenAI to model feedback or for evaluation of student writing.

The survey results (Tables 7, 8) further illuminated the findings on acceptable use of AI in educational contexts. Students and teacher both agreed that artificial intelligence would be a useful tool for teachers and students, but teachers tended to have a more positive outlook on teacher use of AI than did students. Yet both groups responded more cautiously when asked if AI would have a positive impact on education, and both groups reported concern about how AI might be used by teachers and students. The apparent trepidation regarding AI in education seems to be countered by the perceived utility of the tool. These are apprehensions that can be addressed by establishing clear policies on the use of AI and by educating both teachers and students on acceptable use.

Given the positive impact that university and classroom honor codes have on academic integrity by delimiting inappropriate practices (Ely et al., 2013; Konheim-Kalkstein et al., 2008), it is alarming that 94.1% of teachers reported their university as not having a policy in place regarding the use of AI and that 89.7% of teachers acknowledged

they had never educated their students on acceptable use of AI. Clear university policies and statements on ethical use of GenAI are needed, such as the framework proposed by Chan (2023).

Additionally, teachers showed more openness to innovation, but 95.6% of teachers reported receiving no training on the use of AI from their institution. Many may be hesitant to embrace GenAI tools, such as ChatGPT, due to concerns related to cost, privacy, and legality (Kumar, 2023), in addition to a naivety as to how these tools can be appropriately used for educational purposes. Complicating this issue is the question of whether any prescriptions regarding the use of GenAI by students should also apply to teachers. Teacher training on these issues is urgently required.

The findings of this study underscore the importance of students and teachers being transparent in their use of GenAI tools. Using AI without disclosure, by both students and teachers, was considered the least acceptable practice. Yeo (2023) suggests that educators “accept and befriend [GenAI] by showing learners how to use AI authoring ethically and gainfully to achieve their learning intentions and goals” (p. 10). If educators are transparent and clearly state and model acceptable uses of GenAI then students may follow. While there is a nascent body of literature offering practical suggestions and advice for using GenAI in education by educators, learners, and researchers (Ingle & Pack, 2023; Pack & Maloney, 2023b; Bonner et al., 2023), an evidence-centered framework for leveraging GenAI in writing and in higher education in general needs further attention.

Conclusion

By the time this paper is published the GenAI tools and examples of use in this study will likely be anachronistic to a wider variety of available programs that users might interact with; a common limitation in emerging technology research. Even as this paper was being prepared, advances in prompt engineering have shown how students or educators might be able to use GenAI tools in ever more sophisticated ways. Regardless, the seemingly rudimentary prompts exemplified in this study represent how a teacher or student without much experience or training in prompt engineering would likely interact with a GenAI tool like ChatGPT. For this reason, the measured perceptions herein will likely be useful when considering approaches to GenAI integration in education.

Limitations of the study include the non-random sample which inhibits generalizing findings to a wider population, and the heterogeneous population of teachers and students that render between-group comparisons tenuous. Also, the cross-sectional design of the current study does not take into account the dynamic nature of student and teacher perceptions. It is reasonable to assume that both student and teacher perceptions of the ethics of using GenAI in writing education will change as individuals become more aware of and competent in the technology, and as the technology develops and changes over time.

While this research provides initial insights into student and teacher perspectives on using GenAI in academic writing, future studies could further examine potential differences in perceptions of students and educators in varying cultural and

educational backgrounds, or take a longitudinal approach to see if opinions change as the technology becomes more prevalent. In addition, investigating stakeholder perceptions of the use of GenAI through more advanced prompts (such as having GenAI act as a teacher in a writing conference to guide a student through the writing process) would likely prove insightful.

Appendix

#1- Brainstorming

Student input:

Please brainstorm some ideas for an essay on the topic of urban challenges and global warming.

AI output:

1. Discuss the impact of increased urbanization on global warming and climate change.
2. Explore the role of city governments in mitigating the effects of global warming, such as implementing renewable energy sources and reducing greenhouse gas emissions.
3. Examine the challenges faced by cities in developing countries in addressing global warming and climate change, such as limited access to resources and funding.
4. Analyze the impact of global warming on urban ecosystems, such as heat islands and the loss of green space.
5. Discuss the ways in which urban residents can play a role in addressing global warming and reducing their carbon footprint.
6. Evaluate the effectiveness of existing solutions to urban challenges related to global warming, such as green buildings and public transportation.
7. Explore the potential long-term consequences of global warming for urban areas, such as sea level rise and extreme weather events.
8. Consider the potential benefits of addressing global warming in urban areas, such as improved air and water quality, and increased public health.

#2- Outline

Student input:

Write an outline of a 5-paragraph essay on the topic of urban challenges and global warming.

AI output:

- I. Introduction
- II. Briefly introduce the topic of global warming and its impact on urban areas.
- III. Thesis statement: Despite the challenges, addressing global warming in urban areas is essential for the health and sustainability of our cities.
- IV. Body Paragraph1: The impact of urbanization on global warming and climate change
- V. Discuss the increased greenhouse gas emissions resulting from urbanization
- VI. Explore the impact of heat islands on global warming

- VII. Mention the role of cities in contributing to global warming and climate change
- VIII. Body Paragraph 2: The role of city governments in addressing global warming
- IX. Discuss the ways in which city governments can mitigate the effects of global warming, such as implementing renewable energy sources and reducing greenhouse gas emissions
- X. Evaluate the effectiveness of existing solutions, such as green buildings and public transportation
- XI. Body Paragraph 3: The challenges faced by cities in developing countries in addressing global warming
- XII. Explore the limited access to resources and funding faced by cities in developing countries
- XIII. Discuss the potential consequences of global warming for these cities, such as sea level rise and extreme weather events
- XIV. Conclusion
- XV. Summarize the main points made in the essay
- XVI. Restate the thesis
- XVII. Call to action: urge readers to consider the importance of addressing global warming in urban areas and take action to reduce their carbon footprint.

#3- Writing

Student input:

Write a 5-paragraph essay on the topic of urban challenges and global warming.

AI output:

Urban areas around the world are facing a range of challenges related to global warming and climate change. The increased greenhouse gas emissions resulting from urbanization contribute to global warming and exacerbate the impacts of climate change. Additionally, the heat generated by urban areas can create heat islands, which further contribute to global warming.

City governments have an important role to play in addressing global warming in urban areas. They can implement policies and programs aimed at reducing greenhouse gas emissions, such as promoting the use of renewable energy sources and increasing energy efficiency. They can also invest in infrastructure that supports sustainable transportation, such as public transit and bike lanes.

However, cities in developing countries often face challenges in addressing global warming and climate change. They may have limited access to resources and funding, making it difficult to implement solutions to mitigate the effects of global warming. These cities may also be more vulnerable to the impacts of global warming, such as sea level rise and extreme weather events.

Despite these challenges, addressing global warming in urban areas is essential for the health and sustainability of our cities. By taking action to reduce greenhouse gas emissions and adapt to the impacts of global warming, we can protect our urban environments and ensure that they remain livable for future generations.

In conclusion, global warming and climate change present significant challenges for urban areas. City governments have an important role to play in addressing these challenges, but they must also overcome obstacles such as limited resources and funding. By taking action to reduce greenhouse gas emissions and adapt to the impacts of global warming, we can ensure that our cities remain sustainable and livable for the future.

#4- Revision

Student input:

Rewrite the following paragraph and make it sound more academic:

Global warming is important for several reasons. First, global warming will affect many people. All countries and their people will be affected by global warming. Second, the effects of global warming will be costly. Third, many people will die because of global warming.

AI output:

The significance of global warming is multifaceted. Its impacts will be far-reaching, affecting individuals and communities across all nations. Additionally, the economic costs of global warming will be significant. Furthermore, it is anticipated that global warming will result in a significant loss of human life.

#5- Feedback

Teacher input:

Can you provide some suggestions on how to improve this student's writing:

Culture shock is a feeling of disorient that occur when someone is in a new culture. It can be caused by factors such as differences in language, culture, and ideas and culture shock include homesickness, difficulty eat to new foods, and feeling of isolation or depression and it is important to recognize that culture shock is normal part of adjusting to new culture. And it is possible to overcome it with time and effort. Some-way to cope with culture shock is connect with friends and family back home, and get out support from others who also have same feelings. By choose to learn about and adapt to a new culture, individuals can slowly learn enjoy and appreciate new culture.

AI output:

Here are some suggestions for improving the writing:

1. Use clear and concise language: Avoid using long or complex sentences, and try to use straightforward language.
2. Use transitions: Add transitional phrases or words (e.g., "for example," "however," "in addition") to help the reader follow the flow of the writing.
3. Use concrete examples: Provide specific examples to illustrate the points you are making. This will help the reader understand and relate to your writing more easily.
4. Use active voice: Write in the active voice (e.g., "She overcame culture shock by learning about the new culture") rather than the passive voice (e.g., "Culture shock

was overcome by her learning about the new culture"). The active voice makes the writing more engaging and easier to follow.

5. Use proper grammar and punctuation: Review your writing for errors in grammar and punctuation, and correct them as needed. This will help ensure that your writing is clear and easy to understand.

#6- Evaluate

Teacher input:

Evaluate the quality of ideas expressed in the following students' paragraph:

I believe arming teachers with guns will make the classroom safer. To begin with, the only way to stop a bad guy with a gun is a good guy with a gun. Teachers are good people and people trust them. Students and teachers can feel safer if the teacher has a gun. Second, having a gun is protected by the government. The reason why this is protected is because guns can protect people from dangerous people and governments. Therefore, teachers with guns will make the classroom safer.

AI output:

While the paragraph presents some ideas about arming teachers with guns, the reasoning and logic used to support these ideas are problematic and flawed. The first point made in the paragraph relies on the premise that the only way to stop a bad person with a gun is to have a good person with a gun, which is not necessarily accurate or supported by evidence. The second point asserts that having a gun is protected by the government, but does not provide any context or explanation for this claim. Additionally, the final statement that teachers with guns will make the classroom safer is not supported by any evidence or logical reasoning. Overall, the ideas expressed in this paragraph are not well-developed or thoughtfully presented and do not effectively convey a coherent or logical argument.

Acknowledgements

Not applicable.

Funding

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

The authors declare that they do not have any competing interests.

Received: 25 July 2023 Accepted: 23 October 2023

Published online: 10 November 2023

References

- Carlson, M., Pack, A. & Escalante, J. (2023). Utilizing OpenAI's GPT-4 for written feedback. *TESOL Journal*. <https://doi.org/10.1002/tesj.759>
- Baker, R. S., & Hawn, A. (2021). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, 32, 1052–1092. <https://doi.org/10.1007/s40593-021-00285-9>
- Bland, J., & Altman, D. (1997). Statistics notes: Cronbach's alpha. *BMJ*, 314, 275. <https://doi.org/10.1136/bmj.314.7080.572>

- Bonner, E., Lege, R., & Frazier, E. (2023). Large language model-based artificial intelligence in the language classroom: Practical ideas for teaching. *Teaching English with Technology*, 23(1), 23–41. <https://doi.org/10.56297/BKAM1691/WIEO1749>
- Bridgeman, B., Trapani, C., & Attali, Y. (2012). Comparison of human and machine scoring of essays: Differences by gender, ethnicity, and country. *Applied Measurement in Education*, 25(1), 27–40. <https://doi.org/10.1080/08957347.2012.635502>
- Chan, C. Y. K. (2023). A comprehensive AI policy education framework for university teaching and learning. [preprint in arXiv]. <https://doi.org/10.48550/arXiv.2305.00280>
- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- Chomsky, N. (1991). Linguistics and cognitive science: Problems and mysteries. In Aka Kasher (Ed.), *The Chomskyan Turn*. Blackwell, pp. 26–53.
- CWPA, NCTE & NWP. (2011). *National Framework for success in postsecondary writing*. Council of Writing Program Administrators, the National Council of Teachers of English, and the National Writing Project. <http://wpacouncil.org/files/framework-for-success-postsecondarywriting.pdf>
- Dehouche, N. (2021). Plagiarism in the age of massive generative pre-trained transformers (GPT-3). *Ethics in Science and Environmental Politics*, 21, 17–23. <https://doi.org/10.3354/eseep00195>
- Ely, J. J., Henderson, L., & Wachsmann, Y. (2013). Testing the effectiveness of the university honor code. *Academy of Educational Leadership Journal*, 17(4), 95–104.
- Evering, L. C., & Moorman, G. (2012). Rethinking plagiarism in the digital age. *Journal of Adolescent & Adult Literacy*, 56(1), 35–44. <https://doi.org/10.1002/JAAL.00100>
- Fan, N. (2023). Exploring the effects of automated written corrective feedback on EFL students' writing quality: A mixed-methods study. *SAGE Open*. <https://doi.org/10.1177/21582440231181296>
- Farrokhnia, M., Banihashem, S. K., Norooz, O. I., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*. <https://doi.org/10.1080/14703297.2023.2195846>
- Fitria, T. N. (2021). "Grammarly" as AI-powered English writing assistant: Students' alternative for English writing. *Metathesis Journal of English Language Literature and Teaching*, 5(1), 65–78. <https://doi.org/10.31002/metathesis.v5i1.3519>
- Gardner, J., O'Leary, M., & Yuan, L. (2020). Artificial intelligence in educational assessment: 'Breakthrough? Or buncombe and ballyhoo?' *Journal of Computer Assisted Learning*, 37, 1207–1216. <https://doi.org/10.1111/jcal.12577>
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, 26(2), 5–24.
- Graham, M., De Sabbata, S., & Zook, M. A. (2015). Towards a study of information geographies: (im)mutable augmentations and a mapping of the geographies of information. *Geography and Environment*, 2(1), 88–105. <https://doi.org/10.1002/geo2.8>
- Graham, S. (2019). Changing how writing is taught. *Review of Research in Education*, 43(1), 277–303. <https://doi.org/10.3102/0091732X18821125>
- Graham, S., Kihara, S. A., & MacKay, M. (2020). The effects of writing on learning in science, social studies, and mathematics: A meta-analysis. *Review of Educational Research*, 90(2), 179–226. <https://doi.org/10.3102/0034654320914744>
- Hockley, N. (2018). Automated writing evaluation. *ELT Journal*, 73(1), 82–88. <https://doi.org/10.1093/elt/ccy044>
- Hu, K. (2023). *ChatGPT sets record for fastest-growing user base—Analyst note*. Reuters. <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>
- Huawei, S., & Aryadoust, V. (2023). A systematic review of automated writing evaluation systems. *Education and Information Technologies*, 28, 771–795. <https://doi.org/10.1007/s10639-022-11200-7>
- Ingle, S. J., & Pack, A. (2023). Leveraging AI tools to develop the writer rather than the writing. *Trends in Ecology & Evolution*, 38(9):785–787. <https://doi.org/10.1016/j.tree.2023.05.007>
- Jackson, M. C. (2021). Artificial intelligence and algorithmic bias: The issue with technology reflecting history and humans. *Journal of Business & Technology Law*, 16(2), 299–316.
- Kasneji, E., Sessler, K., Kuchemann, S., Bannert, M., Dementieva, D., Fischer, F., Grasser, U., Groh, G., Gunnemann, S., Hüllermeier, E., Krusche, S., Kutynoik, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneji, G. (2023). ChatGPT for good? On opportunity and challenges of large language models for education. *Learning and Individual Differences*. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kaplan, A., & Haenlein, M. (2018). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15–25. <https://doi.org/10.1016/j.bushor.2018.08.004>
- Konheim-Kalkstein, Y. L., Stellmack, M. A., & Shilkey, M. L. (2008). Comparison of honor code and non-honor code classrooms at non-honor code university. *Journal of College and Character*. <https://doi.org/10.2202/1940-1639.1115>
- Kumar, R. (2023). Faculty members' use of artificial intelligence to grade student papers: a case of implications. *International Journal for Educational Integrity*. <https://doi.org/10.1007/s40979-023-00130-7>
- Lampropoulos, G., Ferdig, R. E., & Kaplan-Rakowski, R. (2023). A social media data analysis of general and educational use of ChatGPT: Understanding emotional educators. *SSRN*. <https://doi.org/10.2139/ssrn.4468181>
- Lo, C. K. (2023). What is the impact of ChatGPT on Education? A rapid review of the literature. *Education Science*. <https://doi.org/10.3390/educsci13040410>
- National Council of Teachers of English. (2013). *NCTE position statement on machine scoring*. https://ncte.org/statement/machine_scoring/
- O'Neill, R., & Russell, A. (2019). Stop! Grammar time: University students' perceptions of the automated feedback program Grammarly. *Australasian Journal of Educational Technology*, 35(1), 42–56. <https://doi.org/10.14742/ajet.3795>
- OpenAI. (2023). Terms of use. <https://openai.com/policies/terms-of-use>. Accessed 4 Jul 2023.
- Pack, A. & Maloney, J. (2023a). Using generative artificial intelligence for language education research: Insights from using OpenAI's ChatGPT. *TESOL Quarterly*. <https://doi.org/10.1002/tesq.3253>

- Pack, A., & Maloney, J. (2023b). Potential affordances of generative AI in language education: Demonstrations and an evaluative framework. *Teaching English with Technology*, 23(2), 4–24. doi: 10.56297/buka4060/vrro1747
- Pennycook, A. (1996). Borrowing others' words: Text, ownership, memory, and plagiarism. *TESOL Quarterly*, 30(2), 201–230.
- Sadeghi, R. (2019). The attitude of scholars has not changed towards plagiarism since the medieval period: Definition of plagiarism according to Shams-e-Qays, thirteenth-century Persian literary scientist. *Research Ethics*, 15(2), 1–3. <https://doi.org/10.1177/1747016116654065>
- Seow, A. (2002). The writing process and process writing. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 315–320). Cambridge University Press.
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching*. <https://doi.org/10.37074/jalt.2023.6.1.17>
- Sutherland-Smith, W. (2005). Pandora's box: Academic perceptions of student plagiarism in writing. *Journal of English for Academic Purposes*, 4, 83–95. <https://doi.org/10.1016/j.jeap.2004.07.007>
- Tatum, H. E. (2022). Honor codes and academic integrity: Three decades of research. *Journal of College and Character*, 23(1), 32–47. <https://doi.org/10.1080/2194587X.2021.2017977>
- Tseng, W., & Warschauer, M. (2023). AI-writing tools in education: If you can't beat them, join them. *Journal of China Computer-Assisted Language Learning*. <https://doi.org/10.1515/jccall-2023-0008>
- Urlaub, P., & Dessein, E. (2022). From disrupted classrooms to human-machine collaboration? The pocket calculator, Google Translate, and the future of language education. *L2 Journal*, 14(1), 45–59. <https://doi.org/10.5070/L214151790>
- Wang, Z., & Han, F. (2022). The effects of teacher feedback and automated feedback on cognitive and psychological aspects of foreign language writing: A mixed-methods research. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2022.909802>
- Weigle, S. C. (2013). English as a second language writing and automated essay evaluation. In M. D. Shermis & J. Burstein (Eds.), *The handbook of automated essay evaluation: Current applications and new directions*. Routledge.
- Yang, M. (2023). *New York City schools ban AI chatbot that writes essays and answers prompts*. The Guardian. <https://www.theguardian.com/us-news/2023/jan/06/new-york-city-schools-ban-ai-chatbot-chatgpt>
- Yeo, M. A. (2023). Academic integrity in the age of artificial intelligence (AI) authoring apps. *TESOL Journal*. <https://doi.org/10.1002/tesj.716>
- Yeo, S. (2007). First-year university science and engineering students' understanding of plagiarism. *Higher Education Research & Development*, 26(2), 199–216. <https://doi.org/10.1080/07294360701310813>
- Yeo, S., & Chien, R. (2007). Evaluation of a process and proforma for making consistent decisions about the seriousness of plagiarism incidents. *Quality in Higher Education*, 13(2), 187–204. <https://doi.org/10.1080/13538320701629202>
- Yu, H., & Guo, Y. (2023). Generative artificial intelligence empowers educational reform: Current status, issues, and prospects. *Frontiers in Education*, 8, 1183162. <https://doi.org/10.3389/feduc.2023.1183162>
- Zhang, Z. V. (2020). Engaging with automated writing evaluation (AWE) feedback on L2 writing: Student perceptions and revisions. *Assessing Writing*. <https://doi.org/10.1016/j.asw.2019.100439>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
