International Journal of TESOL Studies (2024)
Vol. 6 (3) 95-112 https://doi.org/10.58304/ijts.20240307

Article

Leveraging a Diagnostic Exam for Cultivating AI Literacy: Postgraduate Student Reflections on ChatGPT

Leslie Sherwood
Texas Tech University, USA

Kara Mac Donald*
Defense Language Institute, USA

Received: 13 August, 2024/Accepted: 2 October, 2024/Published: 18 October, 2024

Abstract
This article describes an action research project that uses a mixed-methods approach to explore postgraduate second language (L2) English Language Learner (ELL) students' views of artificial intelligence (AI) tools in terms of how such technology affects their thinking processes through an AI literacy intervention. The researchers coded the intervention output, thirty first-day writing diagnostic exams of students in a high-intermediate L2 English writing and communications course at a U.S. research university. The output was based on a reading and prompt discussing potential learning implications with the growing ubiquity of generative AI. Findings from students' perspectives are discussed through an AI literacy analytical framework as well as possible pedagogical implications for how English for Academic Purposes (EAP) instructors might support their students in navigating AI tools while in the EAP classroom.

Keywords
Artificial intelligence literacy, writing diagnostic exam, ChatGPT, English for Academic Purposes

1 Introduction

In late November 2022, OpenAI launched ChatGPT (Generative Pre-trained Transformer), a chatbot that permits a user to input a prompt, with the application producing an answer in a conversational or essay format. Educators have been considering the implications of using AI, with immediate discussions (Bowman, 2022; Herman, 2022; Roose, 2023) in reputable news publications and teacher practitioner publications **(**McMurtie, 2022; Yorio, 2023; Frances & Zimotti, 2023). One prominent topic has been the fury over the potential unethical use of AI, namely plagiarism, and the trepidation of its impact on the students' writing instruction. Some schools have blocked the app on school networks as a deterrent (Johnson, 2023). Additionally, one undergraduate computer science student was prompted to develop an app, GPTZero, for teachers to detect if ChatGPT was used to produce an essay **(**Bowman, 2023).

*Corresponding author. Email: kmacd@rocketmail.com

Since the release of the initial version of ChatGPT, there have been subsequent improved free versions (i.e., ChatGPT-3.5) and the enhanced paid ChatGPT-4. As a result, the AI chatbot landscape is rapidly changing and offering better functionality in under two years.

On the cusp of ChatGPT's presence in the educational sector, it is valuable to inquire how students view the technology as well as the implications for the higher education academic L2 writing classroom. Discussions in popular news and academic sources have been expanding viewpoints of teachers and school leadership among other stakeholders, but the views of postgraduate students second language (L2) English language learners (ELLs) are limited (Koltovskaia, et al. 2024). To address this, the authors examined master's degree and doctoral students' perceptions of ChatGPT through a writing assessment prompt, analyzed based on the students' points of view regarding the benefits and drawbacks of its use. Also, to hedge against possible misuse of generative AI, the instructor prompted students to reflect on the degree to which such technology may interfere with their thinking processes.

This present study summarizes existing scholarship on this topic over two time periods: the literature available at the time of the study and what has been produced since. Both periods jointly will contextualize the discussion for both generative AI advances now and in the future.

2 Literature Review

The advancement of generative AI has posed concerns for instructors in the months following ChatGPT's broad release (Huang, 2023; Weissman, 2023). In the postsecondary space, instructors who used writing for learning and assessment wrote in various media outlets about the threat that such technology would pose to students' learning, given instructors' current pedagogical approaches (Huang, 2023). In the year and a half since ChatGPT has become available, scholarship has emerged, proposing conceptual frameworks for ensuring academic honesty. Kohnke et al.'s (2023) review of dozens of American universities' guidelines for treating generative AI resulted in the team proposing generative artificial intelligence assessment literacy. The idea is that instructors would critically revise their curricula given the affordances available to students with generative AI. Indeed, the work of ensuring students' ethical use of technology continues.

2.1 AI literacy

AI literacy is defined in a myriad of ways. Some definitions include technical skills, e.g., programming, while others do not. Ng et al.'s (2021) literature review of thirty articles distilled the idea of AI literacy into the following three competencies: know and understand, use and apply, and evaluate and create, mapping these competencies to Bloom's Taxonomy. While noteworthy at the time, Ng et al.'s (2021) framework did not include generative AI, given the fact that such technology was not as publicly accessible as it is now, with the unveiling of ChatGPT in November 2022. For a generative AI context in L2 instruction, Tseng & Warschauer (2023)'s AI literacy framework is worth discussing. Recently, Warschauer, et al. (2023)'s AI framework discussion maps Ng et al.'s (2021)'s framework onto their initial framework proposal (Tseng & Warschauer, 2023). In their text, Warschauer, et al. (2023) argue that based on the proficiency level of students, with foundational writing skills needed before work with AI literacy, and the purpose of the learning activities, AI literacy should be cultivated in L2 writing instruction. Their framework outlines five components. First, the authors state that understanding how AI works, along with its affordances and limitations, is needed, just as Ng et al. (2021) note in their "know and understand" part. Second, like Ng et al. (2021)'s second competence of "use and apply," Warschauer, et al. (2023)'s framework includes "access and navigate" to ensure students know how to use the technology for their specific purposes. Third, Warschauer, et al. (2023)'s framework addresses prompting, or the iterative practice with refining prompts until the desired output is achieved. To this end, Huang (2023) offers a method based on experimental experience in EFL writing classroom. Specifically,

the author argues that for maximum effectiveness, prompts need to be direct and rubric-based, as well as requesting for feedback, not rewriting. Next, the authors' framework includes "corroboration," which allows for the learner to seek and confirm additional evidence for the output received. The framework concludes with "incorporation" of the output, which includes alignment with institutional policy regarding AI use and meeting required transparency measures.

Teng (2024a) offers a systematic review of the use of ChatGPT for EFL writing. The review supports the notion that AI literacy is necessary for students' effective use, including the preservation of learning. Given the above scholarship on the topic, we define AI literacy for this study as the critical use of AI tools, which means the human writer thoroughly analyzes the AI output before deciding how and if to incorporate it into their own writing product.

2.2 ELL students' perception of generative AI

Since ChatGPT was publicly released in 2022, a growing number of studies have contributed to reporting ELL students' views toward using the tool for learning. In terms of student views, research has shown that students are conflicted, reporting both positive and negative views of ChatGPT as a learning aid (Al-Alami, 2024; Chan and Hu, 2023; Tica & Krsmanovic, 2024; Xiao & Zhi, 2023; Wang, 2024; Yan, 2023).

Of the positive views that students held, some reported that the tool offered a quick and efficient method for obtaining feedback (Al-Alami, 2024; Chan and Hu, 2023; Tica & Krsmanovic, 2024; Wang, 2024; Yan, 2023). The immediacy of feedback generated by ChatGPT was noted in a variety of research contexts: ESP (Tica & Krsmanovic, 2024), ESL (Wang, 2024), and EFL (Al-Alami, 2024; Chan and Hu, 2023; Yan, 2023). Another benefit students reported with the tool was its value in providing useful language feedback (Al-Alami, 2024; Tica & Krsmanovic, 2024) and in supporting the writing process more generally (Al-Alami, 2024; Tica & Krsmanovic, 2024; Wang, 2024). For language feedback, students reported help with language accuracy (Al-Alami, 2024) and more specifically, ChatGPT's utility as aiding vocabulary development (Al-Alami, 2024; Tica & Krsmanovic, 2024). As for the writing process, students noted a number of affordances, including obtaining ideas from the output (Al-Alami, 2024), the output providing generally useful information (Tica & Krsmanovic, 2024) and utility at various phases of the writing process (Wang, 2024). Students also commented that ChatGPT was user-friendly (Al-Alami, 2024; Tica & Krsmanovic, 2024) and was useful for generating ideas (Al-Alami, 2024; Chan and Hu, 2023; Xiao & Zhi, 2023). Coupled with these individual studies, Teng (2024a)'s systematic review of ChatGPT for EFL writing reveals that students are generally positive about its potential for learning.

On the other hand, students reported a number of drawbacks with using ChatGPT as a learning tool, including plagiarism concerns (Al-Alami, 2024; Chan and Hu, 2023; Xiao & Zhi, 2023; Yan, 2023). Such concerns seemed to range from a general discussion of ethics (Chan and Hu, 2023) to the potential negative implications with submitting AI-assisted work to the respective educational institutions (Al-Alami, 2024; Xiao & Zhi, 2023; Yan, 2023). Students notably responded that the use of ChaGPT could interfere with their learning processes (Al-Alami, 2024; Chan and Hu, 2023; Yan, 2023). Students in Yan's (2023) study expressed concern that the tool could become a shortcut for learning, perhaps due to overreliance (Chan & Hu, 2023). Al-Alami's (2024) study showed a number of negative effects on the cultivation of creativity, writing skills, thinking skills, knowledge acquisition, and general learning practices. While some students applauded ChatGPT's language accuracy and general adherence to academic writing conventions(Tica & Krsmanovic, 2024; Yan, 2023),some students expressed concern over inaccuracy of some of its output (Chan and Hu, 2023; Tica & Krsmanovic, 2024; Wang, 2024; Xiao & Zhi, 2023). Wang (2024) reported that some students seemed alarmed by ChatGPT's "potential to mislead people" (p. 12). Students also expressed the potential loss of personalization when using ChatGPT as a collaborator during the writing process, whether this was a perceived lack of emotion (Al-

Alami, 2024) or an issue of unique voice (Wang, 2024). What is clear from the above discussion is ELL students' nuanced view of ChatGPT's utility as a learning aid in the writing process.

2.3 Writing process

Writing is a complex process that Hayes and Berninger (2014) describe as consisting of four source components: 1) Attention and the ability to concentrate and attend to a task, 2) the use of long-term memory, including the storage and recall of information over a long period of time, 3) the use of working memory, i.e., the part of short-term memory that is associated with immediate conscious perceptual and linguistic processing, and 4) reading, including reviewing the text and revisiting source texts. Thus, the iterative process of crafting a written piece requires a focused thought process, including reflecting on what was written while drawing on working and long-term memory, thereby making it a dynamic and complex process in producing a manuscript. The writing process at its heart goes far beyond the principal stages of composing and expressing ideas, organizing them, and inferring a conclusion for the ideas shared. Additional factors guide the actions and decisions that writers choose in the process of crafting a piece (Flower & Hayes, 1981). The authors' texts are informed by their audience and the context of the text, making the writing process an internal dialogue with potential readers as the text is written. Notably, Al-Alami's 2024 study reported students commenting that ChatGPT could disrupt the writing process by interfering with one's concentration. Given the complexity of the writing process as discussed above, at least for beginning ChatGPT users, ELL students may find it disruptive perhaps before integrating the tool into their workflow.

Writing is also a human process in that there is impact on both the author and the reader. For the author, there are physical (Murray, 2002) and psychological benefits to writing (Pennebaker, 1997; Watson et al., 2019). However, the health benefits of writing are not derived from the writing tools (i.e., pen or keyboard), but from the thoughts and emotions in the writer's mind (Murray, 2002). Writing clarifies thinking, allows the writer to articulate and explain themself to others, and/or allows for collaboration with others to build meaning that could not be done alone. How AI might shape the writer's physical and psychological benefits gained through writing, whether positive or negative, remains unclear. For the reader, the influence the author's words can have on them is boundless, as the author has the potential to touch the reader's core. The dialogue of communication between the reader and the author's crafted words is in itself a very human process. When AI enters as a collaborator with the writer, the impact on the reader likely changes, regardless of the reader knowing of the AI-assistance.

This complex, iterative process has been called post-process writing. In this process, Kalan (2014) describes activity where writing processes are acknowledged as less linear and predictable than the process writing approach; writing is not an individual, but a social act, involving other stakeholders in the process. Additionally, written texts are produced in a mess of cultural, social, and power dynamics (Kalan, 2014). Indeed, it is within this social context that writing takes shape and is produced. However, the atomization of writing through AI reduces the quality of the human experience of writing (Johnson, 2023) and can cause the authors' writing skills to atrophy and the readers' experiences to be less dynamic.

Some scholars have pushed back against such a perceived threat that AI could have on the writing process, especially as it pertains to student learning. In discussing the impact of large language models (LLMs), such as ChatGPT on writing studies, Graham (2023) remarks that "...AI provides the opportunity to add multiple dimensions of recursion where prompt-engineering, output curation, fact-checking, and revision become an orthogonal dimension to traditional writing and learning processes" (p. 166). In what Graham calls a multidimensional recursive AI-assisted writing process, the LLM supplements the iterative potential the writer employs with the self and/or peer reflection they already have. Thus, the LLM would augment the current post-process writing practices that individuals currently engage in.

2.4 Background & research questions

At the time of the study, just two months after ChatGPT's broad release, what international postgraduate students thought about ChatGPT and generative AI more generally was not known. Instructors discussed in popular media and professional organizations the (mostly) negative effects such technology would have on the students' learning process, but the students' perspectives have been limited (Johnston et al., 2024) This study sought to begin the conversation of ChatGPT in terms of the dangers the uncritical use of it could pose for the learning process, from the student's perspective, i.e., commencing the conversation with an AI literacy approach, and 2) to further invite student perspectives as a starting point for revising the curriculum in real time to address the potential threat ChatGPT posed to the writing classroom specifically.

Despite the aforementioned investigations into student perspectives, there remains little understanding of postgraduate international students in an EAP classroom and their view of AI in the literature. This research seeks to add to the growing understanding of student perspectives, so that these can be taken into consideration by instructors and institutions accordingly.

The purpose of this study was to critically explore the students' views of AI-assisted writing tools, generative AI specifically, on the learners' writing process. This information would then subsequently inform the treatment of the said technology in the language learning classroom. The research question is *To what extent do postgraduate ELL students view generative AI tools as interfering with their thinking processes?* with the sub-question*: What additional effects of generative AI tools on their thinking processes do students perceive?*

3 Methods

3.1 Context

The study focused on analyzing the student-participants' diagnostic exams written in a U.S. research university setting. The exams were written in January 2023, two months after ChatGPT was broadly released to the public. The instructor had read a variety of ideas about how to – or whether to – address this technology in the writing classroom. Many ideas were offered, but very little student viewpoints were reported. The lack of awareness of student perceptions on the topic was coupled with the instructor's experience with similar student demographics who were immersed in AI in their fields. Thus, the instructor was curious as to the thoughts and opinions of the students, especially in the context of how to hedge against academic misconduct from using AI and how to utilize it in the EAP curriculum.

3.2 Study design

This study used a mixed-methods design common in applied linguistics action research (Burns, 2015). Along with investigating problems, Hammond (2020) asserts that the value of action research is elucidating opportunities that might also exist, especially opportunities concerning new technology. Such opposing interests motivated this research, as the leading questions assumed both problem and opportunity and the impacts of AI on the students' thoughts and communication. This study also followed the principal goals of action research, with the main aim to inform the instructor-practitioner's practice in their particular classroom and only secondarily to extrapolate the results exterior to the classroom (Efron & Ravid, 2020).

Since the results are intended to inform instructor-practitioners' practice, the authors utilized some numerical data in the analysis to move beyond vague qualifications such as 'most', 'some', and 'all most all', drawing on the benefits of quasi-statistics (Becker, 1970, 1990), Maxwell (2010) discusses the common practice of presentation of numerical data, citing Erickson (2007), Hammersley (1992), and

Miles and Huberman (1984),whose work understands numbers as a key indicator of presenting valid and generalizable data.

3.3 Participants

Thirty students were included in this study. These participants were one of the authors' students, enrolled at the graduate and PhD levels in two sections of the same class: A high-intermediate level multi-skills L2 English course designed for international graduate students. A cluster random sampling method was used since all students enrolled in the two classes composed the participant pool (Mackey & Gass, 2016). While eighteen students were enrolled in one class and seventeen in the second, two students from the first class and three from the class were absent on the day of the intervention, thus were excluded from this study. The students came from a variety of disciplines, including architecture and engineering. The students were placed into the course through the university's English language proficiency examination. Successful completion of the course resulted in fulfillment of the university English language requirement. Demographic variables of interest, including academic discipline, degree program, and first language (L1) are shown in Table 1.

Table 1
Demographic Variables

Variable	Number (Percentage)
Academic discipline	
Engineering	15 (50%)
Hard sciences	7 (23%)
Art	2 (7%)
Architecture	3 (10%)
Other	3 (10%)
Degree program	
Doctorate	3 (10%)
Master's	18 (60%)
Professional master's	9 (30%)
L1	
Chinese	16 (53%)
French	4 (13%)
Hindi	2 (7%)
Korean	2 (7%)
Other	6 (20%)

3.4 Intervention

To hedge against the possible misuse of generative AI in the EAP classroom, the instructor used a first-day diagnostic exam consisting of a short article from *Nature* (Stokel-Walker, 2022), which served as both the intervention and the data collecting tool. The diagnostic exam prompted the students to respond to a question concerning the extent to which they agreed with the idea that AI applications could result in an "outsourcing of one's thinking." A secondary question, which probed them to consider additional implications of such tools, elicited further perspectives. The instructor chose to analyze the students'

diagnostic exams, leveraging this teaching artifact as a consequence of teaching, not a research-restricted information gathering process. Indeed, as Burns (2015) notes, "the research should show it builds upon, rather than detracts from, practitioners' major responsibilities for teaching" (p. 194). The students were asked to support their ideas with evidence, had approximately one hour to write, and all but one student completed the writing using a computer. The students were informed that the diagnostic was graded based on completion, and that their views would be of interest to the instructor.

3.5 Data collection

Data elements included student information sheets and anonymized diagnostic exams. Descriptive statistics for understanding student response types holistically were calculated using Microsoft Excel. The researchers acknowledged their subjectivity in the data analysis process to achieve an understanding informed by existing literature in the field. Institutional Review Board (IRB) approval was granted by the teaching institution: 23-000419.

The approach of obtaining students' perceptions could be understood as weakening the validity of the research in the following ways: i) the data collection was a timed exam, possibly limiting students' expression and comfortability to share their views, especially for ELLs, ii) there was only one data source. Regarding the first, participants were Master's, Professional Masters's and Doctoral students at a research university, requiring language proficiency at the level of an IELTS (International English Language Testing System) minimum score of 7.0 on a 0 to 9-band system, which documents' individuals enrolled in these programs as possessing full proficiency in English for specialized work and professional settings (IELTS, 2024). With this command of English, an hour to produce a page (i.e., 500 words approximately) offered ample time for participants to take time to organize their thoughts, produce a draft and review if they wished.

Regarding the second, the data collection principal aim was to inform the instructor-practitioner's practice in a particular classroom and only secondarily to extrapolate the results exterior to the classroom (Efron & Ravid, 2020). Acknowledging the secondary goal to share the outcomes with other instructor-practitioners, the study utilized a procedure that included simultaneous work by more than one researcher to code and interpret the data. Working collaboratively with a second researcher to inductively code the student exams, negotiating meaning in multiple coding sessions, increased the trustworthiness of the results (Burns, 2015).

3.6 Data coding and analysis

To prepare the data, the instructor-researcher anonymized each student exam. To enhance trustworthiness (Burns, 2015), the study utilized the "investigator triangulation method", a procedure that includes simultaneous work by more than one researcher. Specifically, the instructor-researcher worked collaboratively with the second researcher to inductively code the student exams, negotiating meaning in multiple coding sessions.

The researchers systematically coded the essays using thematic analysis, a method that went beyond the instructor-researcher's reading of the exam to identify emerging views and opinions. Similar coding categories were identified based on a negotiated discussion. Multiple themes identified from each student paper, were marked once per paper, regardless of elaborative depth or repetition. The coding categories were informed by the research, where the researchers identified statements describing generative AI tools as interfering with their thinking processes, as well as any other positive or negative views. The codes identified, and which ones were merged, are presented in Appendix A.

As part of coding the diagnostic exams, a catalog of the entire data set was generated. Then, an approach to interpreting that data as identified, based on the goal of the action research; to inform instructor-practitioner classroom practice. The core question in analyzing the data was to identify the

salient/common themes among the coded data. There were two phases in the coding process. An initial phase, where all content was coded with a unique themed code. Next, in a process of identifying what was being expressed, codes that were similar were merged (Appendix A). Lastly, the authors interpreted the possible implication of the most salient coding categories.

4 Results

The results of the study will be discussed in this section.

RQ:

To what extent do postgraduate ELL students view generative AI as interfering with their thinking process?

What additional effects of generative AI tools on their thinking processes do students perceive?

The response type, positive or negative, and description of the response is noted in Table 2. After the second phase of data coding, the number of positive and negative response tokens were tallied in each category. To obtain. Figure that would be easily interpreted, a percentage for each tallied category was found by dividing the value by the total value and then multiplying the results by 100 (i.e., (value/total value) x 100%).

Table 2

Responses Coded in Diagnostic Exam

Positive response	**Number (Percentage)**
Task efficiency/versatility	16 (29.09%)
Enhances thinking	9 (16.36%)
Increasing insights	8 (14.55%)
Fostering creative output for society	7 (12.73%)
Leveraging the apps' potential for education and elsewhere	6 (10.91%)
Necessitates update of pedagogy	5 (9.09%)
Need for openness to technological evolution	4 (7.27%)
Total positive responses	55 (100%)
Negative response	**Number (Percentage)**
Writing/creating as a human process	17 (19.77%)
Displaces thinking and learning	15 (17.44%)
Academic integrity of individual	11 (12.79%)
Reliance on applications	6 (6.98%)
Loss of creative/productive output for society	6 (6.98%)
Reduction of critical thinking	5 (5.81%)
Erroneous information	5 (5.81%)
Loss of creativity output for individual	4 (4.65%)
Communication is inherently human	4 (4.65%)
Moral access and impact	3 (3.49%)
Interference with the writing process and authorship	3 (3.49%)
Interferes with thinking	2 (2.33%)
Total negative responses	81(100%)

In general, the students reported more negative than positive views in terms of how AI impacts their thinking processes. Of all the views expressed by the students in their diagnostics, 60.99% of them were negative, while just over 39% were positive.

The most commonly reported negative concerns were, in order of most to least frequency: Writing and creating as a human process, concerns over displacing thinking and learning, and issues of academic integrity. The most common positive responses were, in descending order: Task efficiency/versatility, the enhancement of thinking, and AI's perceived ability to increase insights.

4.1 Selected data samples

Each of the above categories is worthy of a lengthy discussion. However, key data is presented that has been categorized in two groupings based on the research questions to highlight participants' views on the role and impact of AI tools. The two themes are: Views of Generative AI as Interfering with the Thinking Process and Views of Generative AI as Facilitating Thinking Processes . Within each category, the data is grouped into additional subcategories with example excerpts from participants as illustrations.

Theme 1: Views of Generative AI as interfering with the thinking process

Academic integrity

One significant concern, which is reflected in the literature popular media discussions, is the influence that generative AI could have on students' adherence to academic principles and may use it to skirt research and writing processes based on the ease the tools offer.

Participants shared particular concern on how it could not only threaten academic integrity, but also reduce the human thinking process.

> *I believe that most students who count on ChatGPT to write will merely use its function to automatically generate essays which from my perspective, they might not even have a chance to think about the topics of essays. (*Participant 23)

Essentially, the student is saying that ChatGPT will be leveraged to skate steps and responsibilities in the learning and writing process. In doing so, students fail to adhere to academic requirements that the work presented in their own. The student's quote also touches on how student will bypass the essential stages of writing where the author engages with the topic, critically analyzes and reflects on the significance of the issues, problem and/or outcome, which also undermines academic integrity.

Interference in the learning process

The following comments further address participants' views of how generative AI tools could inhibit thinking and learning regarding the following coding category: Writing and Creating as a Human Process (N-RHP).

> *The underlying logic is that algorithms cannot invent new things from nothing which is we human's strength and the most important or distinctive part of essay and paper as well, and they just collect existing data and do the predictions or imitations instead, so they can never touch the essence of an outstanding essay.* (Participant 13)

The student is arguing that computer calculations come from existing data and therefore, do not produce new and innovative views on topics and/or issues, which is the hallmark of the human brain and experience.

Comments directly addressing concerns of generative AI interfering with the learning process (i.e., coding category: N-ILP) are shared below.

> *As it becomes so easy and simple to get all the information you need in order to get a good grade, students might skip the most important part of any research, the part where you're wondering around, making mistakes or understand things wrong. I believe that those are very meaningful moments that brings creative and original thoughts.* (Participant 1)

The student states that people learn by engaging with their thoughts and experiencing the effects of their actions. In essence, trial and error learning can be a tedious process, but it is essential for deeper learning.

Of the three principal concerns around generative AI interfering with the thinking process, views related to compromising individuals' academic integrity were significant (i.e., 12.79%), but participants were more concerned on how it would inhibit writing and/or creating as a human process (i.e., 19.77%) and displaces thinking and learning (i.e., 17.44%). This may suggest that while the field and popular media were most concerned with academic integrity being the primary concern, the international postgraduate students as participants in the study were more focused on the negative impact on the development of their and others' learning and writing processes.

Theme 2: Views of Generative AI as facilitating thinking processes

Although participants expressed concerns of the negative impact of generative AI tools' impact on learning and thinking, there was a notable difference among the top three coding categories: Task Efficiency/Versatility (i.e., P- TV 29.0%), Enhances Thinking (i.e., P-ET 16.36%) and Increasing Insights (i.e., P-II 14.55%), compared to the data reflecting the primary three areas of concern.

Task efficiency and enhancing thinking and perspectives

Participants expressed how generative AI could serve as a mechanism to broaden individuals' understanding of their field and the academic writing process. Selected code category data for Enhancing Perspectives (i.e., P-EP) are shared below.

> *Because the results can be the inspirational source of thinking, which means it can lead to thinking further or generate different ideas.* (Participant 25)

The student is expressing that external input beyond one's surrounding is required to grow and obtain new understanding, and that AI tools can serve to bring such new perspectives to individuals to expand their understanding.

Additionally, the responses coded for Diverse Input (i.e., P-DI) are shared below.

> *They are available even at midnight, suitable for naive to learn, and can answer questions in different fields.* (Participant 24)

The student believes the ubiquitous access to knowledge resources permits even novice learners to have reliable support to proceed on a task when struggles of gaps in understanding are encountered.

Participants view the generative AI as a tool that is adaptable to their needs and always available, reflecting to some degree of their expectations in other parts of their lives due to technology.

Additional affordances

The following comments extracted from participants' diagnostic writing regarding the following coding categories: Update Instructional Pedagogy (i.e., P-UIP), Update Instructional Pedagogy View (i.e.,

P-IPSV), and Critical Review of Tech Policy (i.e., P-CRTP) in schools are shared below. The excerpts are grouped together as they serve to represent in part how participants understood how AI could positively inform the update of pedagogy to reflect current communication demands and practices, including the use of technology in a rapid evolutionary climate within education.

Update instructional pedagogy

This category reflects students' views of how generative AI can, or will, shift the way writing is taught and assessed. Its presence offers an opportunity to enhance and/or revise instructional practices to better meet students' needs.

> *However, with the advancing technology and the presence of AIs, I think there is a need to rethink how we should evaluate student essays in this perhaps changing academic landscape. AI is there, perhaps be more humane or situate more contextualized, student-specific prompts?* (Participant 9)

The student expresses that the former traditional ways of conducting writing assessment may no longer apply and that AI offers opportunities to adjust assessment practices to better meet students' individual needs and gaols.

Instructional pedagogy view

This category shows students' views of how generative AI can be used to adjust pedagogical practices beyond mastering the mechanics of writing and summarizing content taught, but to ask students to critically engage with the content they are writing about.

> *This new technology may change the Academic studies, and make more room for critical assignments, where the students require not only to sum up some information and imitate others in their fields, but also, and even mainly, to express their own point of view*. (Participant 1)

This student suggests that because AI can create a summary of existing information, writing instructors will need to shift from tasking student to merely report fact to interpreting and providing their view of the facts.

Critical review of technology policy

This category highlights students' understanding that new challenges presented by generative AI give rise to innovative approaches to solving these challenges and foster critical engagement with the academic landscape.

> *But it will bring bad things to like more cheating in the academy, but I think it is an opportunity to reinforce our checking software. Innovation brings innovation, so we will find a way to check the submissions and to put the light on the ones done with Chat GPT.* (Participant 8)

This student expresses that even though now there may appear to be new challenges that undermine writing and writing instruction, humans adapt to change and with access to technological solution the new challenges will be addressed.

These excerpts extensively reflect participants' view of how generative AI can enhance their thinking and academic work and its generation of possibilities for instruction, with noted consideration for evaluating academic work in its presence.

5 Discussion

The results of this study's reported L2 ELL postgraduate students' views of AI tools based on their responses from a diagnostic exam that served as the intervention. Approximately two-thirds of the participants' views were negative, including concerns about the impact AI may have on established academic processes (i.e., academic integrity, reading, and learning to publish) and human development (i.e., writing process, thinking, morals). One-third of the participants expressed views that were positive, including how AI tools provided advantages and efficiency, with potential for use in academia. Participants generally expressed both negative and positive views in their diagnostic exam responses, as reflected in Participant 1's two excerpts in the Results section.

5.1 AI literacy

In terms of AI literacy, it is important to consider the degree to which the students' responses aligned with the five elements outlined in Tseng & Warschauer's 2023 framework: *understand, access and navigate, prompt, corroborate,* and *incorporate.*

As for the first element, *understand*, the instructor had observed that this student population has had their disciplinary studies enmeshed in AI and machine learning for the past few years. Thus, it is reasonable to assume the students would understand what such technology entails. Further, the intervention reading provided an extended definition of ChatGPT, making it accessible to a broad audience and included discussion of some affordances and limitations for learning.. Thus, part of the diagnostic reading conveyed this knowledge to those unfamiliar with the topic and thus helped to meet this criterion in the framework. Additionally, student responses of the tool's affordances, such as increasing insights, fostering creative output for society, etc. and limitations. erroneous information, interfering with thinking, etc. suggest that the text, if not mentioning these items, prompted students to consider possible implications involved with using the tool.

In terms of the *access and navigate* and *prompt* elements, while the students were not invited to use AI during the diagnostic writing, this appeared later in subsequent curriculum redesign efforts. Thus, this paper's described intervention did not provide an opportunity for students to meet this criterion. Interventions as described by Huang (2023) would provide students with useful prompting criteria when seeking personalized feedback on their individual writing specifically. With *corroboration*, some students did mention erroneous information as an issue when using the tool. This suggests that again, the reading and/or writing process of this intervention allowed for some reflection and connection with this AI literacy element. Finally, on the topic of *incorporation*, we see some students' anticipation of aligning their use of the tool with expectations within the broader community of practice. Specifically, students' expressed concerns about academic integrity of the individual and interference with the writing process and authorship. These issues, expressed in their writing diagnostics, point to this final element of Tseng & Warschauer's 2023's AI literacy framework. For EAP instructors to prepare those postgraduate students interested in publication, *incorporation* is essential. This can be accomplished in part by crafting clear guidance as to how to use and report use of ChatGPT in the writing process and how to document such use to ensure transparency (Teng, 2024b).

5.2 ELL students' perceptions on the writing process

In terms of student views, it's important to consider the extent to which the student's views here are aligned or not with those reported elsewhere in the literature. Perhaps unsurprisingly, students in this study also reported ambivalent feelings towards the use of AI for their learning and communication habits. Like previous scholarship (Al-Alami, 2024; Chan and Hu, 2023; Tica & Krsmanovic, 2024; Wang, 2024; Yan, 2023), the students in this study expressed the efficiency of the tool for obtaining

feedback as a significant affordance. In fact, as the most commonly reported affordance in this study, the student's positions as postgraduate students may have something to do with this response frequency. While trying to balance a variety of academic, social, and professional tasks, students in this group may feel especially pressured in time and thus see this benefit as being the most useful immediately for their contexts.

Students in this study did also share the position of the tool's potential for increasing insights, a position shared elsewhere (Al-Alami, 2024; Chan and Hu, 2023; Xiao & Zhi, 2023). However, in our study, students also reported, at a higher rate, that the tool also posed threats to creativity for the individual and society. Thus, this study contributes to the idea that while students see potential with AI fostering creativity, they also see the tool as a threat to this important phenomenon.

The present study also adds additional variation of viewpoints on this topic. Students in this study shared the following affordances: enhancement of thinking, leveraging the application's potential for education and elsewhere, necessitating update of pedagogy, and the need for openness to technological evolution.

In terms of drawbacks, similar to elsewhere (Al-Alami, 2024; Chan and Hu, 2023; Xiao & Zhi, 2023; Yan, 2023), students reported concerns with plagiarism, specifically as it relates to academic integrity of the individual and the possible interference with the writing process and authorship. Taken together, these concerns, which consist of the third most commonly reported limitation, show that like teachers, students have such issues on their minds. It may be then that students share Tseng & Warschauer (2023)'s fifth element in their AI literacy framework: *incorporation*. Time spent on identifying institutional guidance on if and how to use the tool and how to make any permissive use transparent would be time well spent for all stakeholders involved.

The threat of interfering with one's learning was expressed by students in this study, as it was elsewhere (Al-Alami, 2024; Chan and Hu, 2023; Yan, 2023). Whether by displacing thinking and learning, relying on the applications, or reducing critical thinking, students seem to see the potential detriment this tool could have on the learning process. As well, potentially erroneous information was reported by students, similar to Chan and Hu, 2023; Tica & Krsmanovic, 2024; Wang, 2024; and Xiao & Zhi, 2023. Critical use, including corroboration, the fifth element in Tseng & Warschauer (2023)'s framework seems to resonate with such responses. Noteworthy, however, is that the participants in this study reported more negative than positive views of ChatGPT on their thinking and writing process than positive perspectives. This is in contrast to the findings in the recent systematic review on ChatGPT for EFL writing (Teng, 2024a), which found that students generally had positive perspectives toward the tool. This disconnect may be due to the timing of the study: the study reported here was conducted on day one of the two classes, while the review collated student views from throughout, including the end of the term. This suggests that student views might change based on how, when, and if ChatGPT is used explicitly in the classroom setting.

Beyond the current scholarship reviewed, this study contributes additional perspectives of graduate students. Of interest perhaps is some students' views that this tool might provide the tipping point to revising curriculum and pedagogy. Some students seem to sense that while Grammarly and Google Translate may not have been as significant of shifts as ChatGPT, perhaps along with the aforementioned tools, and the implications for a reform of communication assignment tasks. This view may be the canary in the coal mine–if the canary wasn't present before– for those involved in curriculum and instruction. Some students also expressed the idea of communication being inherently human. This concern is worth further research to understand what factors AI tools cannot replicate in communication.

5.3 Pedagogical implications

Based on students' reported views in this study, a multiplicity of pedagogical implications are apparent. First, given students' responses were varied, it appears that students are somewhat open to the use of

ChatGPT as a learning tool. However, the students' concerns with the tool must be concerned with designing pedagogy in response to this technological affordance. Given the serious concern students have with the view that writing is fundamentally a human process, critical thinking must be embedded into a ChatGPT assisted writing curriculum, such that students see the limitations of the tool. This can be achieved in part through cross-referencing the output from ChatGPT with an additional tool, such as a corpus or online writing lab, or a human, e.g., an instructor or writing center tutor. Additionally, reflective questions can be built into lessons that leverage the tool for learning purposes, which can provide space for students to thoroughly consider their experience. Such interventions with a third party and reflective work can also help to hedge against students' worry that the tool can displace thinking and learning. While instructors creatively work to ensure learning within a system that includes ChatGPT as an affordance, they ultimately may be reassured that students too worry about this, pointing to perspectives of learner motivation. Finally, the concern regarding academic integrity is one that instructors and institutions can consider responding to by proactively providing guidance and best practices to their students, informing their choices with using and making transparent any tool use in the development of communication artifacts. Ultimately, instructors may choose to seek their students' views and adjust pedagogy accordingly.

6 Limitations

One of the main limitations of this study was the small sample size. Coupled with the study being action research, great care must be taken when extrapolating results to another context. Additional studies are needed, with the goals of fostering AI literacy and of trying to understand the students' complex views to inform the teaching with/amidst the evolution of technology generally and generative AI specifically. In terms of data analysis, it should be noted that a number of statistical tests were run, and the p-value was not adjusted. The authors chose a greater likelihood of Type 2 errors than Type 1. Thus, the results should be interpreted cautiously. Given society's ambivalence, the complexity of the students' views, and the dynamic nature of AI, continual research at various levels is warranted, and curricular projects and interventions should be adjusted based on findings.

7 Conclusion

Classroom practice becomes at some level a form of personal habit, and so, a teacher can often become comfortable with a routine instructional approach that they are unaware of until someone or something requires a deliberate reflection. Generative AI tools, like ChatGPT, have prompted stakeholders across the various sectors of education to do so. As part of the ongoing discussion on the implications of AI on learning, this article reported on a qualitative collaborative action research project, which both sought to develop postgraduate second language (L2) English Language Learner (ELL) students' AI literacy competencies (Tseng & Warschauer, 2023) and to explore their views of generative AI in terms of how this technology affects their thinking processes. Ultimately, the first day diagnostic served as an intervention to begin to develop identified elements crucial for AI literacy. Such an intervention, along with subsequent curricular revisions, might be useful for instructors as new AI tools arise, if they are interested in cultivating the students' literacy with such tools. While this study contributes to the field by reporting L2 postgraduate reflections concerning generative AI's effects on their thinking, future research on this population is necessary, given the limited studies on this topic. Further, the evolution of generative AI tools will certainly necessitate future research on how such developments impact the students' thinking, with implications for their academic and professional work.

Appendix A

Coding Descriptions

*Combined into first code presented

Code	Description
[N-AII]	Negative: academic integrity of individual
[N-HP]	Negative: Writing/creating as a human process
[N-LA]*	Negative: Loss of authorship
[N-WC]*	Negative-Interference with the writing process and authorship
[N-RCT]	Negative: Reduction of critical thinking
[N-RMR]*	Negative: Reducing memory retrieval
[N-IT]*	Negative-interferes with thinking
[P-DI]*	Positive: Diverse input
[P-EP]*	Positive: Expanding perspectives
[P-ET]	Positive: Enhance thinking
[N-ILP]*	Negative: Interferes with the learning process
[N-RVL]*	Negative: Reduces the value of learning
[N-RHT]*	Negative: Replace human thinking
[N-LCOI]	Negative: Loss of creativity output for individual
[N-LCOS]	Negative: Loss of creative/productive output for society
[P-FCOS]	Positive: Fostering creative output for society
[P-UPSV]*	Positive: Update of pedagogy, allowing for student voice
[P-UIP]*	Positive: Update instructional pedagogy
[P-CRTP]*	Positive: Critical review of tech policy in schools
[P-LAP]	Positive: Leveraging for the apps potential for education and elsewhere
[P-NOTE]	Positive: Need for openness to tech evolution
[N-RA]	Negative: Reliance on app
[P-TE]*	Positive: Task efficiency
[P-TV]*	Positive: Task versatility
[N-FN]	Negative: Erroneous info
[N-MIS]*	Negative: Moral impact on society
[N-PP]*	Negative: Privileging the privileged
[N-CIH]	Negative: Communication is inherently human

References

Al-Alami, S. E. (2024). EFL Learners' Attitudes Towards Utilizing ChatGPT for Acquiring Writing Skills in Higher Education: A Case Study of Computing Students. *Journal of Language Teaching and Research*, *15*(4), 1029–1038. https://doi.org/10.17507/jltr.1504.01

Becker, H. S. (1970). *Sociological work: Method and substance*. Transaction Books.

Becker, H. S. (1990). Generalizing from case studies. In E. Eisner & A. Peshkin (Eds.), *Qualitative inquiry in education: The continuing debate* (pp. 233-242). Teachers College Press.

Bowman, E. (2022, December 19). A new AI chatbot might do your homework for you. But it's still not an A+ student. *NPR*. https://www.npr.org/2022/12/19/1143912956/chatgpt-ai-chatbot-homework-academia

Bowman, E. (2023, January 9). A college student created an app that can tell whether AI wrote an essay. *NPR*. https://www.npr.org/2023/01/09/1147549845/gptzero-ai-chatgpt-edward-tian-plagiarism

Burns, A. (2015). Action research. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics* (pp. 187-204). Bloomsbury Academic.

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*(1), 43–18. https://doi.org/10.1186/s41239-023-00411-8

Efron, S. E., & Ravid, R. (2020). *Action research in education: A practical guide*. The Guilford.

Erickson, F. (2007, May). Specifying "usually" and "some": Using simple descriptive statistics in qualitative inquiry. Presented at the 2007 Congress of Qualitative Inquiry, Urbana, IL.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, *32*(4), 365–387. https://doi.org/10.2307/356600

Frances, C. & Zimotti, G. (2023, January 17). Robots vs. humans: Does ChatGPT pose a challenge to second language writing? *The FLTMag*. https://fltmag.com/chatgpt-second-language-writing/

Graham, S. S. (2023). *Post-process but not post-writing: Large language models and a future for composition pedagogy*. https://eric.ed.gov/?id=EJ1390327

Hammersley, M. (1992). *What's wrong with ethnography? Methodological explorations.* Routledge.

Hammond, M., & Wellington, J. (2020). *Education research: The basics*. Routledge.

Hayes, J.R. & Berninger, V.W. (2014). Cognitive processes in writing: A framework. In B. Arfe, J. Dockrell & V. Berninger (Eds.), *Writing development in children with hearing loss, dyslexia, or oral language problems: Implications for assessment and instruction* (pp.3-15). https://doi.org/10.1093/acprof:oso/9780199827282.003.0001

Herman, D. (2022, December 9). The end of high-school English, *The Atlantic*. https://www.theatlantic.com/technology/archive/2022/12/openai-chatgpt-writing-high-school-english-essay/672412/

Huang, J. (2023). Engineering ChatGPT Prompts for EFL Writing Classes. *International Journal of TESOL Studies*, *5*(4), 73-. https://doi.org/10.58304/ijts.20230405

Huang, K. (2023, January 16). Alarmed by A.I. chatbots, universities start revamping how they teach *The New York Times*. https://www.nytimes.com/2023/01/16/technology/chatgpt-artificial-intelligence-universities.html

IELTS. (2024). International English Language Testing System, Understanding Your IELTS Score, *British Council*. https://takeielts.britishcouncil.org/teach-ielts/test-information/ielts-scores-explained

Johnson, A. (2023, January 21). ChatGPT in schools: Here's where it's banned—And how it could potentially help students. *Forbes.* https://www.forbes.com/sites/ariannajohnson/2023/01/18/chatgpt-in-schools-heres-where-its-banned-and-how-it-could-potentially-help-students/?sh=1bcd51696e2c

Johnston, H., Wells, R.F., Shanks, E.M.. Boey, T., & Parsons, B.N. (2024). Student perspectives on the use of generative artificial intelligence technologies in higher education. *International Journal of Educational Integr*ity, *20*, 2. https://doi.org/10.1007/s40979-024-00149-4

Kalan, A. (2014). A Practice-Oriented Definition of Post-Process Second Language Writing Theory. *TESL Canada Journal*, *32*(1), 1. https://doi.org/10.18806/tesl.v32i1.1196

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for Language Teaching and Learning. RELC Journal, 54(2), 537-550. https://doi.org/10.1177/00336882231162868

Koltovskaia, S., Rahmati, P., & Saeli, H. (2024). Graduate students' use of ChatGPT for academic text revision: Behavioral, cognitive, and affective engagement. *Journal of Second Language Writing*, *65*, 101130-. https://doi.org/10.1016/j.jslw.2024.101130\

Mackey, A., & Gass, S.M. (2016). Second Language Research: Methodology and Design (2nd ed.). Routledge. https://doi.org/10.4324/9781315750606

Maxwell, J.A. (2010). *Using Numbers in Qualitative Research. Qualitative Inquiry 16*(6) 475 –482. https://doi.org/10.1177/1077800410364740

McMurtrie, B. (2022, December 13). AI and the future of undergraduate writing. *The Chronicle of Higher Education*. https://www.chronicle.com/article/ai-and-the-future-of-undergraduate-writing

Miles, M. B., & Huberman, A. M. (1984). *Qualitative data analysis: A sourcebook of new methods*. Sage.

Murray, B. (2002). Writing to heal. Monitor on Psychology, 33(6), 54. https://www.apa.org/monitor/jun02/writing

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, *2*, 100041. https://doi.org/10.1016/j.caeai.2021.100041

Pennebaker, J.W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science*, *8*(3), 162- 166. https://doi.org/10.1111/j.1467-9280.1997.tb00403.x

Roose. K. (2023, March 30). How should I use A.I. chatbots like ChatGPT? *The New York Times.* https://www.nytimes.com/2023/03/30/technology/ai-chatbot-chatgpt-uses-work-life.html

Stokel-Walker, C. (2022). AI bot ChatGPT writes smart essays—Should professors worry? *Nature (London)*. https://www.nature.com/articles/d41586-022-04397-7

Teng, M.F. (2023). Scientific writing, reviewing, and editing for open-access TESOL journals: The role of ChatGPT. *International Journal of TESOL Studies 5*(1), 87-91.

Teng, M.F. (2024a). A systematic review of ChatGPT for English as a foreign language writing: Opportunities, challenges, and recommendations. *International Journal of TESOL Studies 6*(3), 36-57. https://doi.org/10.58304/ijts.20240304

Teng, M. F. (2024b). "ChatGPT is the companion, not enemies": EFL learners' perceptions and experiences in using ChatGPT for feedback in writing. *Computers and Education. Artificial Intelligence*, *7*, 100270-. https://doi.org/10.1016/j.caeai.2024.100270

Tica, Lena & Krsmanovic, Ivana. (2024). *Overcoming the writer's block? Exploring students' motivation and perspectives on using ChatGPT as a writing assistance tool in ESP*. ELOPE English Language Overseas Perspectives and Enquiries. 21. 129-149. https://doi.org/10.4312/elope.21.1.129-149

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

Wang, C. (2024). Exploring students' generative AI-assisted writing processes: Perceptions and experiences from native and nonnative English speakers. *Technology, Knowledge and Learning*. https://doi.org/10.1007/s10758-024-09744-3

Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. Journal of Second Language Writing, 62, 101071. https://doi.org/10.1016/j.jslw.2023.101071

Watson, L.R., Fraser, M., & Ballas, P. (2019). Journaling for mental health. *Literacy Research and Instruction*, *49*(2), 194-208. doi:10.1080/19388070902947360.

Weissman, J. (2023, February 8). ChatGPT is a plague upon education. *Inside Higher Ed*. https://www.insidehighered.com/views/2023/02/09/chatgpt-plague-upon-education-opinion

Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages (Basel)*, *8*(3), 212-. https://doi.org/10.3390/languages8030212

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*. https://doi.org/10.1007/s10639-023-11742-4

Yorio, K. (2023, January 17). School librarians explore possibilities of ChatGPT. *School Library Journa*l. https://www.schoollibraryjournal.com/story/School-Librarians-Explore-Possibilities-of-ChatGPT

Leslie Sherwood is a doctoral student in Curriculum and Instruction (Language, Diversity, and Literacy Studies) at Texas Tech University. She has been teaching multilingual students in the U.S. American higher education setting for the past twelve years. Before that, she worked with migrant and immigrant English learners in the U.S. and Australia. Her research interests include assessment, digital literacies, and multimodal communication.

Kara Mac Donald is a Full Professor at the Defense Language Institute in Monterey, CA, USA. She conducts pre-service and in-service faculty training and offers academic support to students. Her background consists of over twenty years in English and foreign language teaching, teacher training, and faculty development across elementary, secondary and higher education. She earned a Masters in Applied Linguistics, TESOL and a Doctorate in Applied Linguistics.