Dissonance in generative AI use among student writers: How should curriculum managers respond?

Franklin G. Talaue1*

1English Department, Binus University, Jakarta Indonesia

Abstract. This qualitative study investigated the case of ChatGPT use by five students in an English for Academic Purposes (EAP) class comprised of 28 students in total. The results of the analysis revealed dissonance in the way the subjects of the study had intended to use ChatGPT in their written assessments and how they used it, six weeks into the writing program. The results of the study revealed that their intention to use ChatGPT may have been within the remit of fair use, but their written output as evaluated by an AI detection tool revealed that the students might have violated academic integrity. This study has clear implications for the role of educators to guide students on how to use AI in their written assessments in such a way that they uphold academic integrity. Educational institutions, therefore, should take a pedagogical stance on AI text generators, which means that proactive measures should be taken to embrace AI as an assistive tool for writing, but fair and guided use of the technology must be in place.

1 Introduction

Shantal (pseudonym), at the beginning of the semester, was asked what she might do with ChatGPT as a tool for her to use in an academic writing class. Surprisingly, she had not heard of ChatGPT although when briefed about it as being another form of Artificial Intelligence she quickly picked up from there and browsed what the fuss was all about.

Almost like her, half of the class enrolled in academic writing did not hear about it, but the other half were familiar with it, but had not used it before. Their unfamiliarity with generative AI was not surprising as they were just about to start the semester in February, a little over two months when ChatGPT was launched on 30 November 2022 [1]. Fast forward to five months later, ChatGPT now has reached over one billion users. It took the platform a mere 5 days to reach 1 million users as compared to Facebook which reached the same figure after 10 months in 2004, while Instagram had the same milestone after two months back in 2010 [2].

ChatGPT and its updated version GPT-4, which was launched in March as paid subscription [3] are seen as a game changer for many sectors, but this paper will focus on how this latest AI advancement has disrupted education, prompting schools and universities worldwide to address how it could be used or not used in education. Some schools have blocked access to ChatGPT asserting that it is not a tool for critical thinking and problem solving [4], while others cite its use as a downright breach of academic dishonesty. However, an increasing number of academics seem to agree that the use of GPT and its newer models like GPT 3.5 and GPT-4 should not be treated as an enemy of education, but rather as an ally [5].

This paper investigated how students in a writing class have used ChatGPT in their assessments and found dissonance between how they thought it could be used and how they used it in their written assessments. This paper supports the notion that generative AI should have a place in education, but there should be acceptable ways by which schools and university administrations should allow their use in classrooms.

The research study analyzed the cases of five students in an Academic English class who used AI to write their summative tasks either partly or entirely as detected by Turnitin. Before their summative tasks, they had been asked in writing how they could use ChatGPT in their subsequent work. They were part of an Academic English class with a total of 28 students, and so it is noteworthy that only around 20% of the students had used AI in writing, at least according to Turnitin which recently added an AI detection feature. It must be noted, however, that non-detection does not mean that students had not used AI at all, including forms or types other than generative chatbots like chatGPT.

Given the now popular belief among scholars that ChatGPT should not be banned in the classroom but rather be used as a learning tool [6], this study proposes ways to do so, following the findings of this study which analyzed the beliefs of students about ChatGPT and how they used it in their actual written assessments. As such, this paper will specifically report students’ beliefs about and actual use of ChatGPT in class, propose guidelines for using it in the writing classroom, and recommend sanctions for violations of the terms of use of ChatGPT in written assessments.

* Corresponding author: fgtalaue@binus.edu
2 Generative pre-trained model

Natural language processing (NLP) is a major category of the broader field of Artificial Intelligence which also includes computer vision, fuzzy logic, expert systems, robotics, machine learning, and neural/deep learning [7]. NLP accomplishes tasks such as language understanding and generation as aided by a machine that manipulates human language or data that resembles such a language, both written and spoken.

The Generative Pre-trained Model (GPT), which is under the generative NLP umbrella, is based on an architecture that makes tasks such as answering a question and text completion possible [8], language generation and understanding [9], including translation and summary generation [10].

ChatGPT is a type of GPT based on the GPT-3 architecture that features prompting to generate conversation-like responses – meaningfully strung words that have been culled from a large language model [11]. OpenAI launched ChatGPT late last year and has since then become a game changer in the online world. With a conversational feature, the model can simulate a question-and-answer situation and so the model responds according to how the user types in question prompts. It can generate contextual information based on these prompts and engage in a series of related questions. ChatGPT data was publicly available texts and licensed data until 2021.

3 Method

This study is an action research that investigated the use of AI text generators in an English for Academic Purposes class at one private university in Indonesia. Specifically, it employed the case study approach in which 5 cases were picked from the total class population. Those 5 cases were purposively selected to study aspects of dissonance between student beliefs on and actual practices in the use of AI in academic writing.

In the first week of the semester, students were asked to write how they could use ChatGPT to write their class assessments, which mostly centered on producing referenced essays.

In the sixth week, the class of 28 students submitted their essays to Turnitin. Five cases of student writing were found to have used AI in their writing based on the AI reports generated by the software. Historically, Turnitin only had a similarity report feature, but as a response to the growing use of AI in writing, the software now also has an AI report feature which puts a percentage report regarding AI use.

The 5 instances of student writing which returned AI-use percentages based on Turnitin reports became the subjects of this case study. The researcher then compared the subjects’ declaration of intention at the beginning of the semester and their actual writing.

The study employed a thematic approach in its analysis of the subjects’ intention regarding the use of AI to deduce their beliefs and attitudes toward the use of AI text generators and at the same time relied on Turnitin data to check how they have used AI in their writing.

This comparison of cognition and actual use of AI puts to the fore dissonance in generative AI use among students enrolled in an academic writing class, which falls under the remit of English for Academic Purposes.

4 Findings and discussion

4.1 Student beliefs and attitudes toward ChatGPT

All the subjects of the study believe that ChatGPT will make their lives easier as writing students. Shantal stressed that she can have error-free writing as ChatGPT can make recommendations about how she can phrase her thoughts. She praises the platform as “incredibly helpful with writing tasks” as she could ask for recommendations to fix her writing in English.

Claire, on the other hand, sounded very positive that she could “create a research paper outline, write research abstracts, translate text from one language to another, and rewrite text passages” using ChatGPT.

Via believes that with ChatGPT she could improve her essays by adding new information, insights, and analysis, as well as providing evidence all with the help of generative AI.

Rey mentioned that with AI, he could improve his writing by asking ChatGPT to develop his ideas and fix his grammatical mistakes.

John, on the other hand, believes that he could improve his writing by asking ChatGPT to elaborate concepts, supply new details or information, and provide more perspectives to develop his thesis.

Overall, these beliefs about the AI platform are commonly held among users of the platform, but the subjects wrote very little to nothing about how they could ethically use it in their writing.

However, all 5 subjects’ opinions about the use of ChatGPT fairly aligned with academics’ opinions of how AI could be used to enhance student learning. Some of the world’s leading academic institutions are of the opinion that generative AI could assist both students and educators with teaching and learning.

Intuitively, however, academia is calling for the ethical use of the AI text generators to preserve academic integrity. While adapting to AI seems to be the order of the day, universities are calling for fair use of the platform and a review of current assessment practices especially about written assessments.

4.2 How the subjects have used AI in their writing

The subjects of the study, along with the rest of the class, had been asked how they could use text-based generative AI. All 5 subjects have been detected to have used it in their writing through a report generated by Turnitin, which has recently launched an AI-detection feature. Turnitin claims that their percentage report for possible AI use has a 98% confidence level [12].
This paper reports the AI use results for two assessment tasks. Out of the 56 submissions, for an English for Academic Purposes class, 7 counts of AI use were recorded, ranging from 29% to 100% percentage of use. Table 1 below provides information on the AI use of the 5 student subjects.

Table 1. AI use of the subjects of the study.

<table>
<thead>
<tr>
<th>Student (pseudonym)</th>
<th>Assessment 1</th>
<th>Assessment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claire</td>
<td>29</td>
<td>100</td>
</tr>
<tr>
<td>Via</td>
<td>31</td>
<td>-</td>
</tr>
<tr>
<td>Rey</td>
<td>91</td>
<td>-</td>
</tr>
<tr>
<td>John</td>
<td>31</td>
<td>46</td>
</tr>
<tr>
<td>Shantal</td>
<td>30</td>
<td>41</td>
</tr>
</tbody>
</table>

As can be seen from Table 1, Clarissa, John, and Shantal all had two counts of AI use, while Via and Rey used AI only once to write two expository essays. A 100 percent score for AI use means that Claire might have simply cut and pasted the text generated by an AI platform. This goes against her idea that she would use it for generating ideas and creating an outline. She might have been tempted to conveniently ask the AI chatbot and treat the generated text as a product of her writing, not as a tool for the early stages of her writing development.

It must be noted that 20 percent of the class had resorted to AI, and so the rest of the class even though they echoed what the five students had mentioned about how they could be assisted with AI in writing did not use the platform at all. This implies the students’ low level of comfort with the technology. It could be that most of the class were apprehensive at least or likely fearful of what the consequences could be if they had been caught using it regardless of their intentions.

There are no policies at the research site as to AI chatbot use, but intuitively, unethical use of AI falls within the remit of an academic integrity breach which is governed by its Campus Life Code of Conduct issued by the University Rector through Decree No. 1426/SK/PTKK-UBN/VII/2016.

Plagiarism is covered in Section 2 Article 9 of the Campus Life Code of Conduct. According to the decree, any student who is caught violating the academic code of ethics such as plagiarizing, cheating during exams, and engaging in other similar deceitful activities will be expelled from the university [13].

A general clause on cheating in exams and other assessments in which plagiarism is covered will provide stakeholders a better framework to deal with acts of student academic dishonesty. With the rise of generative AI, this proposition becomes even more imperative.

### 4.3. AI text generators and their implications for assessments

This case study on AI has clear implications on assessment practices in writing classes at the university level. Overall, writing programs must guide their students as to what is acceptable when it comes to AI use in academic writing. As ChatGPT becomes more widespread at an exponential speed, it is predicted that more and more students would resort to using AI chatbots, and so universities and schools have started to act quickly on the matter.

Based on the results of the study, the dissonance between how a significant percentage of one class of English for Academic Purposes had intended to use ChatGPT and how they used it reveals that intentionality is a problematic premise for the current gray areas of how and when AI should be brought to the AI classroom.

Unethical use of AI text generators for assessments, whether willingly or unwittingly, therefore, should be defined by writing programs and all other university departments and school authorities. Overall, the use of AI in assessments can consider aspects of when to allow the use of AI considering the types of assessments in place, to what extent it can be used, and how assessments can be variegated and task prompts be made to allow for more nuanced assessment practices amid calls for a proactive and regulated use of generative AI [14].

### 4.4. When should AI be used in writing classes?

AI can be a double-edged sword for various reasons. Some assert that is not as intelligent as once thought, and so exercising critical thinking in using it is imperative. The fallibility of AI technology, therefore, means that students may not be allowed to use it during summative assessment exercises. For example, AI technology has been criticized for fabricated references, as well as “biased and toxic content,” [15] among others.

However, writing programs should specify whether a particular piece of written assessment may use AI. For example, formative assessment exercises may indicate AI use in the task prompt, while summative assessments may only be done in the classroom, with AI chatbots made inaccessible to students.

During formative assessment exercises, AI can be used for idea generation during the initial stages of writing. This entails prompting ChatGPT, for example, with a topic list and asking it to scan for references about the chosen topic. However, lecturers must be able to secure a ChatGPT audit for their students, which shows how AI has been ethically used in writing. In other words, students should be able to declare how they have queried AI chatbots in their writing.

Writing programs must also specify when AI chatbots may not be used. This should be indicated, for example, in major examination task prompts in which a clear stipulation on whether access to AI is totally banned. In general, school and university administrations and writing programs may work with a framework to understand the current AI situation in relation to education, decide which AI can be integrated into the school system, and monitor performance and equity.

In summary, students should be encouraged to use AI during formative assessment exercises whether in class or outside of class to assist with idea generation, providing more perspectives, and exploring credible
sources on the internet. This approach emphasizes the idea of embracing AI while managing its use [16].

4.5 How should assessments be set amid the presence of AI?

Personalizing assessments in the writing class may be one way to put to the fore students’ experiential knowledge around the topic that they write. One way of personalizing assessments is to bring a local context to the writing task. This way, students can have a better sense of ownership of the ideas that they bring to the fore.

In addition, students can also be asked to evaluate AI-generated essays. This practice can be a test for the critical evaluation of essays [17]. Doing this means that lecturers will do their query of ChatGPT and have students work in class for a discussion of how an essay could be further substantiated, personalized, and improved in terms of source credibility. This practice is underpinned by the now common knowledge that AI chatbots may not provide reliable information.

There also have been calls to rethink how schools and universities assess students. A multimodal form of assessment is one such practice that can be heightened in the age of AI. Implementing a multimodal assessment in writing may incorporate oral presentations and survey modes in which authentic data dictates the content of any piece of assessment [18]. Debates, podcasting, talk shows, and other similar activities can also be done in class to assess students in the four skill areas of speaking, reading, listening, and writing.

Before starting the semester, writing instructors should have some form of diagnostic writing activity for their students to create a writing profile for their students. This activity must be done in the classroom without any access to the computer or at least an internet connection. One-to-two paragraph sample of students’ writing can provide a wealth of information about students’ natural writing abilities which may be used later to compare their written assessments done outside of the classroom.

Also, requiring students a trail of how they have developed their writing, including AI use, can be put in place. This means that they may submit screen grabs of how they have queried, for example, ChatGPT. This practice may instill in students the value of transparency when they write as they use source ideas.

However, this paper calls for more assessments done in the classroom with very limited access to the internet, especially during summative exams. Currently, the assessment practices for academic writing at the research site favor out-of-classroom written assessments, and this may have resulted in the use of available AI tools without much regard for ethical use. This situation could have been prevented had students been made more aware of how to engage AI chatbots in their assessments.

Overall, lecturers must not take policing roles in student use of AI text generators but instead embrace practices that ethically integrate AI into university writing programs. Acts of disclosure within the remit of transparent, ethical, and fair use will preserve writing integrity in the age of AI.

4.6 Sanctions against the use of ChatGPT

With or without the intention to cheat, students must be made aware of the consequences of unfair or unethical use of AI in summative assessments. This study outlines sanctions that could be applied to student assessments found to have employed AI chatbots unfairly. Overall, this study asserts that unfair AI use is tantamount to plagiarism, after scrutiny of how those 5 subjects in one English for Academic Purposes class used AI in their writing.

Taking the cue from color codes provided by Turnitin for similarity reports, this research suggests percentages of score deductions in initial assessment scores depending on the reported percentages of AI use.

Lecturers may deduct points ranging from -5 to -100 depending on the extent of AI use. Like the Turnitin green zone of around 20% similarity report, this research suggests that AI use of up to 20% may be regarded as fair use. From there, universities may deduct corresponding points for every AI percentage score interval of 10 points. Students who may have fully used AI chatbots may be given a score of 0. This means that full AI use could be considered plagiarism which merits a failing mark. It must be noted however that this recommendation is arbitrary.

5. Conclusions and recommendations

This study investigated the cases of 5 students’ intended and actual use of ChatGPT in their written assessments for an Academic English class. Following a qualitative approach, the analysis has established the dissonance between cognition and practice in the use of ChatGPT. The 5 student cases accounted for 20 percent of the class, all of whom had been asked to describe how they would use the AI platform. Everyone’s intention to use ChatGPT indicated fair use except for the actual practice of the 5 subjects whose Turnitin AI reports returned what could be described as unethical or unfair use of the AI chatbot. The AI reports of the students revealed that the text generated by ChatGPT accounted for anywhere between 29 to 100 percent.

The results of the study indicate that intention and actual practice may differ, but the reasons for the difference have not been determined. This limitation may pave the way for other researchers to investigate this aspect by way of in-depth interviews with a bigger sample of respondents.

The study also recommends that future studies investgate thresholds for what is acceptable or unacceptable in AI use, if the percentages set by AI detection tools were to be used as a good measure for ethical or unethical use of AI in written assessments.

In addition, future investigations of AI use among students may include in-depth interviews with the subjects and include a more thorough analysis of their written output to include how they query AI before, during, and after writing.
The study ultimately has clear implications for university writing programs to establish rules of engagement in AI use. Such rules should not emphasize policing practices but rather embrace a pedagogical stance on the use of AI in writing programs and content courses across all other departments. Universities may approach the guided use of AI from one department to another or may come up with blanket policies and/or guidelines.

University writing programs should also consider multimodal assessments to create robust regimes that shield academic integrity from possible breaches amid the fast developments in generative AI technology. Lastly, universities should constantly monitor developments in AI and not see it as a threat to written assessments but rather as an assistive tool for student learning.

References


