



## Navigating English Writing Proficiency Tests in the Era of Artificial Intelligence

**Benni Ichsanda Rahman Hz**

Universitas Islam Negeri Sumatera Utara Medan

[benni.ichsanda@uinsu.ac.id](mailto:benni.ichsanda@uinsu.ac.id)

**Rebecca Evelyn Laiya**

Universitas Nias Raya

[evelaiya@gmail.com](mailto:evelaiya@gmail.com)

**Martiman Suazisiwa Sarumaha**

Universitas Nias Raya

[martiman@uniraya.ac.id](mailto:martiman@uniraya.ac.id)

**Supiyandi**

Universitas Pembangunan Panca Budi

[supiyandi.mkom@gmail.com](mailto:supiyandi.mkom@gmail.com)

**Corresponding Email:** [benni.ichsanda@uinsu.ac.id](mailto:benni.ichsanda@uinsu.ac.id)

### Abstract

The establishment and growth of artificial intelligence (AI) technology has sparked tremendous interest in the field of writing skills, as AI writing tools have the potential to assist humans in creating and refining text. The availability of these tools may lead to a reliance on AI in writing, which raises concerns about the potential degradation of writing skills. To that extent, will the writing skill session in English proficiency tests (e.g: TOEFL, IELTS, PTE Academic, TOEIC) be displaced? Utilizing a phenomenological case study, the researchers interviewed four experienced test administrators in North Sumatera, Indonesia, to explore their perspective regarding to the impact of AI writing tools on the applications of writing tests in English proficiency tests. Insightful perspectives are elaborated by the participants, implicating both the future of writing skills in general, the future of English writing skills in proficiency tests, and the future of critical thinking of human in the age of AI. However, in case this study is conducted in Indonesia and only involves four test administrators, a wider scale of research is needed to further investigate the potential impact of AI writing tools on writing skills across various contexts and languages.

**Keywords:** *AI writing tools, critical thinking, proficiency test, writing*

### Introduction

Presently, language instructors have a plethora of intriguing technology options to enhance language learning, to the point where it can become overwhelming (Kessler, 2018). The emergence and advancement of artificial intelligence (AI) technologies have been arousing a remarkable level of interest among educational technologists in terms of how to effectively integrate these innovations into the realm of teaching and learning (Adams & Chuah, 2022; Shi et al., 2022), and have been

attracting growing interest within communities that use English as a Foreign Language (EFL) (Zhao, 2022), especially in writing skill. With the rise of AI technologies, a new generation of writing tools has emerged to assist users in the process of writing (Gayed et al., 2022; Godwin-Jones, 2022; McKee & Porter, 2020; Puntoni et al., 2021; Tang et al., 2018). These tools incorporate AI algorithms to provide users with advanced writing support and guidance.

Automated writing evaluation (e.g. Criterion, ETS e-rater, WriteToLearn, MY Access!), automated writing generator (e.g. Textio, Quillbot, Microsoft 365 Pilot, Articoolo, Jasper), automated essay scoring (Turnitin, Intellimetric, PEG Writing, LightSIDE), automated written corrective feedback (Grammarly, ProWritingAid, WhiteSmoke, LanguageTool), writing paraphrases (Quillbot, Spinbot, Prepostseo) and chatbot (ChatGPT, Siri, Google Assistant, Microsoft's Cortana) are the instances of AI Writing Tools (AIWTs) that increasingly being used as assistances for traditional writing methods. Or event, especially, ChatGPT can produce several tasks at one application, such as generating whole text, completing paragraph, correcting paragraph, paraphrasing, and changing the writing style. With the help of these AI-powered writing tools, which are now available on mobile devices and web-based, students can improve their writing skills in ways that may be difficult to achieve through traditional training methods. This is a promising development in the field of education (Nazari et al., 2021), and also causing a shift in the way second language (L2) writing is approached (Barrot, 2021).

AIWTs are beneficial in some aspects (Alharbi, 2023; Fitria, 2021; Gayed et al., 2022; Li & Li, 2018). The AIs enhance the precision of writing in a second language (Li & Li, 2018). AIWTs can help students become more engaged in their writing process by providing them with real-time feedback on their work (Alharbi, 2023), in which the tools can equip students with more personalized feedback, which can be tailored to the individual's needs. For example, AIWTs can suggest alternative words and phrases to improve the clarity and fluency of a student's writing. These tools can also provide grammar and punctuation corrections, and can accommodate students with corrections on spelling, and syntax, as well as suggestions on how to improve the flow and structure of their piece. These machines have the ability to complete tasks rapidly, which can greatly speed up progress and potentially result in ground-breaking discoveries in various fields (van Dis et al., 2023). The tools also provide instructors and other educators with data-driven insights into student performance, which can help

reduce the time needed for editing and proofreading (Fitria, 2021). Thus, these tools have the potential to enhance student learning outcomes (Gayed et al., 2022).

Although AIWTs can offer advantages, such as enhancing writing skills and consistency, they also have drawbacks that must be considered (Alessandro et al., 2021; Biermann et al., 2022; Chaves & Gerosa, 2021; Floridi & Chiriatti, 2020; Lund & Wang, 2023; McKee & Porter, 2020). One of these deficiencies is that writers may become too reliant on pre-written phrases and templates, leading to a decrease in originality and creativity (Biermann et al., 2022; Floridi & Chiriatti, 2020). Students are also reliant to these tools as AI tools have become a familiar aid for students in their research, and some of them have resorted to using these tools to cheat in their assessments, homework, essays, and even their master theses (Dwivedi et al., 2023). In addition, the standardized writing style that AIWTs often produce can be viewed as a drawback. Automated correction algorithms and language models used in AIWTs can introduce bias (Chaves & Gerosa, 2021; Lund & Wang, 2023), especially if they are not trained on diverse datasets (Alessandro et al., 2021). The AIs can only remember and combine information that humans have previously generated, and it lacks the ability to demonstrate human qualities of scientific thinking, like inventiveness and abstract thinking (Dwivedi et al., 2023; van Dis et al., 2023). This contradicts the act of writing in which it is a process that involves multiple steps, and the final product is often influenced by the restrictions and conventions of its assigned genre (Harmer, 2004). Consequently, the accuracy of the suggestions provided by these tools may be limited due to a lack of diversity in the data used to train them, resulting in unintended bias. Eventually, these AI tools cannot understand the context of culture (McKee & Porter, 2020). As the consequence, they lead to bias information on its generating results. Even one of the AIWTs, ChatGPT, has been warning this limitation in its website: "may occasionally produce harmful instructions or biased content". Whereas, writing is a process of providing information to the audience where the audience understands the cultural context given (Scardamalia, 1987).

However, by the massive utilization of the tools, there is a possibility of a disruption, in way that many individuals will use these AIWTs to generate their writing rather than using their own critical thinking to develop the ideas. This could result in a lack of creativity and originality in writing, as well as a failure to fully engage with and understand the material being written about. Include in the world of language proficiency test, the emergence of AI writing tools also appears to be influencing the

way in which English proficiency tests are conducted. English proficiency tests (e.g: TOEFL, IELTS, PTE Academic, TOEIC) typically evaluate the reading, writing, listening, and speaking abilities of individuals seeking to demonstrate their proficiency in the English language. For many years, English proficiency tests have played a crucial role in the education system, influencing the development and achievement of language learning and proficiency (Frost et al., 2020; Owen et al., 2021), and the increasing trend of globalization has led to a need for a universal language of communication, which has resulted in the development of language proficiency tests to assess language skills across international borders (Anamaria-Mirabela & Monica-Ariana, 2013; Bachman & Palmer, 2022). Both in global scale: for those who are non-native speakers, and in Indonesia context: as English is a foreign language, English proficiency tests have become a general way for individuals to demonstrate their language skills. These tests are used for a variety of purposes, including academic admissions, employment opportunities, and immigration requirements. The tests are standardized, meaning that they are administered and scored consistently to ensure accuracy and fairness. As English continues to be the global language of business, education, and communication, the demand for English proficiency tests is likely to continue to increase.

For that extent, in particular, writing skills play a critical role in English language learning, especially in the context of English as a foreign language (Brown, 2000; Celce-Murcia & Olshtain, 2000). Similarly to speaking skill that require linguistic and cognitive process (Hirvela & Belcher, 2016) and lexical density (Malvern et al., 2004; Yu, 2010), the act of writing comprises cognitive processes that involve thinking (Brown, 2001: 336). It is an essential process that aids in the development of analytical skills, as it requires a systematic approach to thinking and problem solving (Jasper, 2005). It is also a cognitive activity that involves the creation and manipulation of mental images, as it forces the writer to confront their ideas and consider them in a more critical and coherent manner (Odell, 1980). Include in language proficiency test, notably writing skills are of utmost importance in English proficiency tests. It is evidenced by the way the tests enable individuals to communicate effectively and accurately (Ariana, 2010). Writing skills are essential for test takers to express themselves efficiently and effectively in a variety of different contexts, including writing letters, emails, essays, and other long-form assignments. Likewise, according to Escamilla et al. (2018), individuals who demonstrate writing proficiency on English

language tests often experience an increased level of confidence in their writing abilities, which can lead to higher education pursuits and a more successful career. In addition, Barkaoui (2014) found out that students who were better at writing tended to score higher on English proficiency tests, demonstrating that a student's writing ability is an essential component of their overall language proficiency. Additionally, writing skills have been found to play a vital role in determining how well a student does on a language test, as writing tests require students to express their ideas in a clear and organized manner (Hosseini et al., 2013). Students who are able to organize their ideas and write coherently tend to score higher than those who are unable to do so, indicating the importance of writing skills in English proficiency tests.

As new types of AI continue to develop, a growing number of challenges and difficulties arise. The appearance of AIWTs has led to concerns about their potential impact on the writing skills, integrity, and competences of students, teachers, scholars, academics, professionals in field of writing, and publics. As these tools are designed to assist writers by suggesting phrases, sentences, or even entire paragraphs, and in some cases, they can even generate entire articles or reports, it is worried that these tools may make individuals overly reliant on them, and therefore may not develop their writing skills as effectively (Biermann et al., 2022). Though writing is a complex skill that requires a combination of cognitive and linguistic abilities, and the more people rely on AI tools, the less they will practice and develop these skills. AI-generated writing is often characterized by an impersonal, robotic style of language that is not as sophisticated as human-generated writing, and is unable to capture the nuances of human-generated language and is more rigid in terms of its structure and syntax (He, 2019). Another concern is that the use of AIWTs could lead to a decline in creativity (Maiden et al., 2023; Ng et al., 2022). Writing is the act of artistic expression, and it requires ingenuity to produce engaging and compelling content. AIWTs may be able to generate content quickly and efficiently, but they are limited by their programming and lack the imagination and creativity of human writers (van Dis et al., 2023), in this way will negatively affect the writing competences and ingenuity in the long run.

Further, there has been a substantial body of literature on the correlation between writing skill and its surrounding aspects and language proficiency tests (e.g. Daria & Natalia, 2020; Kim, 2021; Muluk et al., 2022), such as how proficiency test impact students' attitude, and how to develop writing skills in proficiency test based on certain strategies. However, the study on whether AIWTs will replace writing skills in

English proficiency tests is a relatively new area of research. This field of investigation holds significant importance in the context of current technological advancements. As more and more AIWTs become available, there is a growing concern that writing skills may become obsolete, leading to a decline in the quality of writing. Therefore, this study is aimed to investigate whether writing skill is still required in English proficiency tests or not since the existence of AIWTs. This study focuses on exploring the perceptions of administrators regarding the existence of writing tools, which will help to identify the potential impact of these tools on writing skills. The findings of this study could be useful for policymakers, educators, and test administrators, providing them with insights into the future of writing. It may also contribute to the ongoing debate on the role of technology in education and language proficiency enhancement.

Thus, this study is developed based on the research questions: The research questions on this study are: to what extent do AIWTs affect the writing skills of humans? Will AIWTs impact humans' critical thinking? And will writing skills in English Proficiency Test be displaced?

## **Research method**

### **Design**

This study aims to explore the perceptions of test administrators regarding the potential impact of AIWTs on writing skills in English proficiency tests, so a qualitative case study is the most appropriate method. The study used a phenomenological approach to understand the lived experiences of administrators in this regard. A case study is a type of research that facilitates the investigation of a phenomenon in its contextual setting, utilizing diverse sources of information to uncover and comprehend various dimensions of the phenomenon (Baxter & Jack, 2008). This method will allow for a deep and comprehensive understanding of the perceptions of test administrators towards AIWTs, and how they perceive it to affect the English writing proficiency tests. The diverse sources of information used in the case study will provide valuable insights into the phenomenon and will allow for an in-depth analysis of the contextual factors that may influence these perceptions.

**Participants and Techniques of Data Collection**

A purposive sampling is utilized in selecting the participants: four qualifying and credential test administrators who have notable experience in designing and administering English proficiency tests in various conceivable educational institutions for various purposes in North Sumatera province, Indonesia. The participants were selected based on their knowledge and expertise in the development and administration of English proficiency tests as well as their experience and involvement in English language education as all the participants are also higher education teachers. All the participants also have been teaching EFL at credible universities in Indonesia for years. Data were collected through semi-structured interviews, which were audio-recorded and transcribed for analysis. In addition, a focus group discussion was also conducted to achieve more comprehensive diverse point of view from the participants. The interview questions have been developed based on the research questions and focused on the participants' perceptions and experiences regarding the availability of AIWTs and its effect on English proficiency tests.

**Data Analysis**

The data collected were analysed using thematic analysis, a widely used technique for identifying, analysing, and reporting patterns within qualitative data. The analysis process involved multiple stages, including familiarization with the data, generating initial codes, searching for themes, reviewing and defining themes, and producing the report (Braun & Clarke, 2006). The themes were developed based on the patterns and connections found in the data and have been used to answer the research questions and identify potential implications for the use of AIWTs in English proficiency tests.

**Trustworthiness**

Trustworthiness of the findings was ensured through a rigorous process of peer debriefing, member checking, and triangulation of data sources. Triangulation involves the use of multiple sources of data, methods, and researchers to enhance the credibility, dependability, and transferability of the study (Denzin, 2017). In this study, the use of multiple sources of data (interviews) from a diverse individual of participants, multiple researchers in the data analysis stage, and the incorporation of multiple perspectives in the development of the themes ensured the validity and

reliability of the findings. The results were also validated by providing a rich description of the participants' experiences and perceptions regarding the existence of AIWTs in English proficiency tests. This triangulation approach enhances the trustworthiness and robustness of the study, making it a valuable contribution to the existing literature on the use of AIWTs and writing skill language assessment.

## Findings and Discussion

### Findings

Four test managers of the two most commonly used English proficiency tests in Indonesia: IELTS and TOEFL, participated in the study. The demographic information of the participants is presented below.

Name	Gender	Age	Credential	Education Degree and Field	Main Profession	Teaching Experience
Interviewee 1	F	43	Test Administrator and Manager	Ph.D in English Applied Linguistics	Lecturer	18 years
Interviewee 2	F	32	Test Administrator and Manager	MA in English Education	Test Manager	7 years
Interviewee 3	M	32	Test Administrator and Manager	Ph.D in English Education	Lecturer	7 years
Interviewee 4	F	34	Test Administrator and Manager	M.Sc in Applied Linguistics	Lecturer	9 years

### The Effect of AIWTs to writing skill

This question is addressed in order to explore the insights and identify areas of concern from the Interviewees related the potential impact of AIWTs on writing skills. The interviewee 1 believed that the advent of AI writing generators will have an impact on human writing skills. As these digital tools serve to aid individuals in constructing their written works by providing suggestions for grammatical correctness and paraphrasing, among other features, Interviewee 1 posits that there may be an eventual decline in writing abilities if humans become overly dependent on AI tools.

"100% writing skill will be impacted" – Interviewee 1.

In addition, Interviewee 4 expressed a more specific perspective, highlighting that in specific situations, the use of AIWTs could lead to a decline in human creativity.



"I think it depends on the level of proficiency of the person. If the person is a beginner or doesn't care too much about quality, then using AI writing tools may suffice. However, if the person is at a higher level and requires more substantial content, then relying solely on AI writing tools may not be sufficient. For example, the output of GPT chat may be grammatically correct, but lacking in substance. If the person is a non-native English speaker who struggles with academic writing in English, they may turn to AI writing tools for help in generating content, and be satisfied with the results. Unfortunately, there may be many people in this situation." Interviewee 4

Interviewee 4 emphasizes that the impact of AIWTs on writing skills depends on the level of proficiency of the user. For beginners or those who prioritize quantity over quality, AIWTs may be sufficient. However, for those at a higher level who require substantial content, relying solely on AIWTs may not be adequate. On this matter, Interviewee 4 stressed other important remark:

"If students are just learning to write in English and are constantly overwhelmed with a lot of assignments, it's understandable that many of them would turn to GPT for help. However, if we're talking about professional researchers, they understand the value of research and the purpose of writing, and may not want to solely rely on GPT chat for their work. Nonetheless, GPT chat can be used as an assistant for tasks like summarizing. In that sense, it can be a useful tool." – Interviewee 4

### **Will AIWTs, specifically writing generator, impact human's writing skill?**

Related to the questions whether AI writing generator (in this context: ChatGPT) impact human's writing skill, Interviewee 2 explained a comprehensive perspective. He tried to compare with previous popular tools that assist human brain works: a calculator.

"I am sure, for sure, the effect there is just what I think about it. Those who are involved in the world of English proficiency tests are not too worried about that because I think, we used to, well, back to few centuries ago, let us use the calculators as the imagery, it is just more basic, which for sure will help us in counting. But now the ability to count does not disappear because no matter what, the ability to count is related to our daily needs." –Interviewee 2

Interviewee 2 emphasizes that the impact of AI writing generators on human writing skills is comparable to the impact of calculators on human ability to perform mathematical calculations. While calculators have made basic math more accessible and easier to perform, they have not eliminated the need for human understanding and skill in mathematics. Similarly, AI writing generators can assist writers in producing text, but they cannot replace the need for human creativity, critical thinking, and writing skills.

**Will AIWTs impact academicians' critical thinking?**

One intriguing aspect that emerged from the interview is the impact of AIWTs on critical thinking among academicians. Interviewee 4 presented a particularly interesting perspective on this topic.

"Talking about critical thinking is quite complicated because it is not clear whether we can learn critical thinking from AI or not. It is essential to understand what critical thinking is and how to demonstrate it. Using AI to write raises questions about whether it helps us develop critical thinking or not. For example, if someone asks for help from the GPT chat to improve their writing, whether it positively impacts their critical thinking or not depends on whether they evaluate and analyze the suggestions made by the AI. If the person takes the AI's suggestions and uses them to improve their writing, then they are demonstrating critical thinking; yet it is critical thinking literally. It is difficult to determine whether someone's critical thinking will be negatively or positively impacted by using AI to write, as it depends on the individual's willingness to evaluate and analyze the suggestions made by the AI. Therefore, it is important to review the procedure and the form of interaction with AI to ensure that critical thinking is taking place. On the contrary, simply asking the AI to complete an assignment without evaluating or analyzing its suggestions does not demonstrate critical thinking. " –Interviewee 4

Interviewee 4 emphasizes that whether someone write by utilizing AI for the assistance, in which the writer takes control what she writes, it is still critical thinking. However, if someone writes the article just by asking the AI, it is then not critical in which will eliminate critical thinking.

Likewise, Interviewee 4 argued similar limitation of AI writing generators. She emphasized that the tools cannot comprehend the nuance being discussed on the text.

"It is important to consider that there are various levels of proficiency in writing, as previously mentioned. There are nuances in writing that require attention to meaning, implications, and subtleties which GPT chat cannot provide due to its literal interpretation. While there may be some academic writings that GPT chat can produce, there are many levels and nuances of writing that require human interaction, evaluation, and learning. Therefore, it is important to assess the individual's writing needs and level of proficiency before relying solely on AI writing tools. It is uncertain whether GPT chat can produce a paper of Q1 quality which discuss specify context. Ultimately, the impact of AI writing tools on writing skills will depend on the specific context and level of the user." -Interviewee 4

**Will Writing Skills in English Proficiency Test Disappear?**

This ultimate question is asked to seek whether the emergence of AIWTs will eventually lead to the demise of writing skills. Shedding light on this matter, Interviewee 3 puts forth his perspective:

"I think it still required, and it still survives. Writing skill in language proficiency tests is still necessary and relevant in many context. Such as in the IELTS writing test. For instance, in Writing Task 1, test takers are required to describe tables or charts, and it may be possible to replace this task with AI-generated descriptions in everyday life. However, in Writing Task 2, test takers are asked to present personal arguments, and it would feel strange if these arguments were generated by AI. Writing skills are still relevant and cannot be fully replaced, especially when it comes to personal arguments. While it may be possible to replace them with AI, it would feel unnatural." –Interviewee 3

The interviewee acknowledges that AI-generated descriptions could replace tasks such as describing tables or charts in everyday life. However, in tasks that require personal arguments, such as in Writing Task 2 of the IELTS test, the interviewee believes that AI-generated arguments would feel strange and unnatural. Thus, the significance of writing skills remains irrefutable, and their complete replacement seems implausible, particularly in the realm of personal opinions and arguments.

Likewise, another prominence was incorporated by Interviewee 2.

"And even if you can excel in one skill, there will an inquiry why a significant gap between the tested skills occurred. For instance, you may perform very well in writing, but your speaking test you are poor. It is challenging to achieve a high score in writing if your listening, reading, and speaking skills are around 5 or 6.5. As an examiner, I have observed a noticeable gap between speaking, listening, and reading skills. The required score for speaking and writing tasks is usually very high, around 7, 7.5, or even 8. Therefore, if there is a significant gap between a student's speaking and writing scores, there may be a suspicion of their intellectual abilities. Both speaking and writing demonstrate a student's intelligence, so a high gap between the two may indicate a problem." Interviewee 2

## **Discussion**

### **The Effect of AIWTs to writing skill**

There is possibility of humans relying heavily on AIWTs for assistance rather than harnessing their own writing skills. Interviewee 1 stressed that an over-reliance on AI writing generators could result in a deterioration of creativity and originality in written works, with individuals becoming too dependent on the prompts and suggestions provided by the tools. This highlight is similar to Dwivedi et al's., (2023) notion that students have become accustomed to utilizing AIWTs to assist them in their assignments, and in some cases, to cheat on their assignments, homework, essays, and even their theses. Human dependency on these tools could potentially result in a regression of writing skills is the point of their emphasis.

These perspective highlights the potential risks associated with the widespread use of AIWTs in writing skill. This issue really threats the writing skill of human. The prevalence of students relying on AI tools for their academic work, including using them to cheat, raises concerns about the development of their writing skills. While these tools can be useful for research purposes, they can also potentially impede the development of critical thinking and analytical skills that are crucial for effective writing. Moreover, if students resort to cheating, they not only undermine the educational system but also deprive themselves of the opportunity to improve their writing skills through practice and feedback. As such, it is important for educators and institutions to address this issue and encourage students to approach their academic work with integrity and prioritize the development of their writing skills. Furthermore, the over-reliance on AI tools could result in a decline in the quality of academic work produced by students. These tools are often designed to provide quick solutions and shortcuts to complex problems, which could lead to a lack of originality and depth in students' work. In contrast, human writing requires thoughtful analysis and synthesis of information, which is vital for producing high-quality academic work. Additionally, the use of AI tools to cheat in academic work not only undermines the integrity of the educational system but also deprives students of the opportunity to develop their writing skills through practice and feedback. Cheating hinders the development of important skills such as critical thinking, research, and analysis, which are critical for effective writing.

### **Will AIWTs, specifically writing generator, impact human's writing skill?**

Interviewee 2 suggests that the role of AI writing generators should be viewed as a helpful tool rather than a replacement for human writing skills. Human can use the tools to write, but still require their brain to control the substance of the text. Human can utilize the AIs to assist in some ways, AIWTs provide real-time feedback to students, including suggestions for alternative words and phrases, grammar and punctuation corrections, and feedback on flow and structure, yet they also provide data-driven insights to instructors, reducing the time needed for editing and proofreading (Alharbi, 2023; Fitria, 2021).

### **The Gap between Human vs AI Capacity**

Interviewee 1 added another insights that there is limitations of these AIWTs. A significant limitation of these tools is their inability to distinguish the contextual

phenomena of a given topic, including its pragmatic dimensions. This, in turn, can limit the accuracy and appropriateness of their suggestions and recommendations.

"There must be, the human brain must definitely be able to differentiate. For example, what is there called context? When the machine is used, in my opinion will not understanding a very pragmatic condition. So, I don't think a machine will ever be able to read situations that are very pragmatic in nature. That's on one side, there that certainly cannot be 100% controlled by the machine."- Interviewee 1

Context refers to the circumstances that surround a particular situation, event or statement. It includes factors such as the time, place, people involved and their motivations, and socio-culture which can greatly influence the meaning and interpretation of information. While machines can be programmed to recognize patterns and make predictions based on data, they may struggle to understand the nuances of context that humans are able to perceive. AIWTs' systems will require help with understanding rhetorical context, including immediate communication and broader cultural context involving ethics (McKee & Porter, 2020). They will need assistance with higher-order cognitive skills such as synthesis, decision-making, uncertainty, and ambiguity (McKee & Porter, 2020).

Both humans and machines will need to learn how to communicate with each other, and humans must be aware when they are speaking to machines to ensure effective communication. Whereas, writing is a cognitive activity that involves organizing and structuring information in a meaningful way, requiring careful consideration of audience, purpose, and context (Scardamalia, 1987). For this reason, human intervention may still be necessary to ensure that AI writing systems are able to provide accurate and meaningful responses. Ultimately, the key to effective communication between humans and machines will be the development of a shared language and understanding of how to interact with one another. This will require ongoing collaboration and learning on the part of both humans and machines, working to bridge the gap between artificial intelligence and human, to avoid the bias ((Alessandro et al., 2021; Chaves & Gerosa, 2021; Lund & Wang, 2023).

### **Will AIWTs impact academicians' critical thinking?**

Interviewee 4 notes that whether AI positively or negatively impacts critical thinking depends on the individual's willingness to evaluate and analyse the suggestions made by the AI. Using AI to complete an assignment without any critical evaluation does not demonstrate critical thinking. As, writing is the process

generating text by involving thinking, analytical skill, and critical and coherent manner (Brown, 2001:336; Jasper, 2005; Odell, 1980).

### **The Effect of AI writing tools to proficiency test**

The emphasize of impact of AIWTs to language proficiency test is considered as threatening for the test creator by Interviewee 4, rather than to the skill itself.

"Well, that's very interesting. I think there's a potential threat to the global language assessment industry, such as those offered by Cambridge and IDP. However, I also think it's a good thing because it makes them aware that the world is evolving and humans are creative, always seeking help from anywhere that can assist them. As a result, they have to be innovative and think about how to adapt their tests to keep up with the times. Rather than seeing AI writing tools as a disruption that eliminates writing skills, they should take advantage of this technology to further improve the accuracy of writing skills themselves." – Interviewee 4

Interviewee 4 highlights the potential threat that AIWTs pose to the global language assessment industry. The interviewee recognizes that these AIWTs are becoming increasingly popular and accessible to people, which could potentially reduce the demand for language proficiency tests. However, she sees this as an opportunity for test creators to adapt and innovate their tests to keep up with the changing times. The interviewee believes that rather than being seen as a disruption that eliminates writing skills, AIWTs should be used as a tool to further improve the accuracy of writing skills themselves. Overall, the interviewee's perspective is that AIWTs can serve as a wake-up call for the language assessment industry to be more innovative and adaptable to changing times, rather than just relying on current methods. By doing so, they can improve the quality of language assessment and ensure that language proficiency tests continue to be relevant and effective in the future.

### **Will Writing Skills in English Proficiency Test Disappear?**

Interviewee 2 believed that speaking and writing skills are linked domain. The Interviewee 2 pointed out that a significant gap between a student's writing and speaking scores may indicate a problem with their overall intellectual abilities. The importance of developing both skills simultaneously is explained by the Interviewee 2, in way that these two units complement each other and contribute to a well-rounded communication ability. These points evidenced Hirvela & Belcher's 2016) and Malvern et al's., (2004) notions, that showing high proficiency in writing should be

followed by proficiency in speaking, as this connection is a two-way street, with speaking and writing skills influencing each other, in ways both skills require similar linguistic and cognitive processes, such as organizing thoughts, selecting appropriate vocabulary and grammar, and applying rhetorical strategies. Moreover, showing proficiency in both skills demonstrate the lexical diversity of the test takers (Yu, 2010), as lexical diversity is an essential aspect of both writing and speaking, in which it allows individuals to effectively communicate their thoughts and ideas while also demonstrating their level of mastery of a language. Individuals with a higher level of lexical diversity tend to be perceived as more intelligent and competent (Yu, 2010). In Contrary, a lack of lexical diversity can lead to redundancy and repetition, which can be tedious for the audience to listen to or read.

“..what being written by the test takers on writing skill test reflect their selves, their way of thinking, and their point of view” -Interviewee 3

As, personal argument requires human imagination. Even though it may be possible to replace personal arguments with AI, the result would not have the same human touch as writing produced by humans. Notably, AI can only generate article that the data have been created previously by human or systems (Dwivedi et al., 2023; van Dis et al., 2023). AI operates based on pre-existing data and algorithms, which means that it cannot create anything new that has not been programmed into it. Therefore, AI tools can only produce writing based on the patterns and structures it has been trained on. They cannot think outside the box, come up with unique ideas, or create content that reflects personal experiences or emotions. They cannot replace the value of human writing skills when it comes to producing high-quality and compelling content. Ultimately, it is the human touch that can bring authenticity, creativity, and originality to writing, and these aspects are highly valued in many fields and industries. Nevertheless, writing skills are crucial in creating engaging and compelling content that resonates with the audience. This is something that AI may not be able to fully replicate as it lacks the emotional intelligence and creativity of human writers.

## **Conclusion**

Evidently, insight from all respondents conclude the impact of AIWTs on human writing skills depends on how they are used. If they are used for falsification, AIWTs

may result in dependence on tools rather than utilizing one's own ideas and notions. However, if they are used as tools to assist humans in the writing process, the presence of AIWTs should be considered a potential asset. Otherwise, in the aspect of critical thinking, using AIWTs should not be relied upon entirely as a writing tool. However, it can be viewed as a helpful tool in simplifying the writing process. Utilizing AIWTs to improve text by organizing and refining sentences or paragraphs is literally a nature of critical thinking. Technically, AIWTs still have some limitations, such as difficulty in understanding the cultural context of the text. Therefore, it is crucial for humans to play a role in refining their writing, instead of solely depending on AIWTs, recognizing that AIWTs are not as ingenious as humans in developing the intentions and beliefs.

On the whole, AIWTs tools will not degrade writing skills in English proficiency test as the test is still required to demonstrate the language mastery of the test-takers overall. Ultimately, writing skills will continue to be relevant in English proficiency test context and other various contexts, and AI tools should be used wisely to enhance the writing process, rather than replace it entirely.

However, this study has some limitations, including the small sample size of only four participants, which may not be representative of a larger population. A larger number of participants is required in order to increase the generalizability of the findings. In addition, the study only focuses on the impact of AIWTs on writing skills in English proficiency tests, and it does not take into account the impact on writing skills in other contexts or languages. Therefore, further research is needed to investigate the potential impact of AIWTs on writing skills across various contexts and languages.

## References

- Adams, D., & Chuah, K.-M. (2022). Artificial Intelligence-Based Tools in Research Writing: Current Trends and Future Potentials. *Artificial Intelligence in Higher Education*, 169–184. <https://doi.org/10.1201/9781003184157-9>
- Alessandro, C., Lorenzo, B., Pierpaolo, L., Pinfield, S., & Bianchi, G. (2021). AI-assisted peer review. *Humanities & Social Sciences Communications*, 8(1). <https://doi.org/10.1057/s41599-020-00703-8>
- Alharbi, W. (2023). AI in the Foreign Language Classroom: A Pedagogical Overview of Automated Writing Assistance Tools. *Education Research International*, 2023. <https://doi.org/10.1155/2023/4253331>
- Anamaria-Mirabela, P., & Monica-Ariana, S. (2013). Benefits of english language learning-language proficiency certificates-a prerequisite for the business graduate. *Annals of the University of Oradea, Economic Science Series*, 22(2). <https://ideas.repec.org/a/ora/journl/v1y2013i2p167-176.html>
- Ariana, S. M. (2010). SOME THOUGHTS ON WRITING SKILLS. *Annals of the University of*



- Oradea, Economic Science Series, 19(1).  
[https://econpapers.repec.org/article/orajournal/v\\_3a1\\_3ay\\_3a2010\\_3ai\\_3a1\\_3ap\\_3a127-133.htm](https://econpapers.repec.org/article/orajournal/v_3a1_3ay_3a2010_3ai_3a1_3ap_3a127-133.htm)
- Bachman, L., & Palmer, A. (2022). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press.
- Barkaoui, K. (2014). Examining the impact of L2 proficiency and keyboarding skills on scores on TOEFL-iBT writing tasks. *Language Testing*, 31(2), 241–259. <https://doi.org/10.1177/0265532213509810>
- Barrot, J. S. (2021). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning*, 1–24. <https://doi.org/10.1080/09588221.2021.1936071>
- Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report*, 13(4), 544–559. <https://nsuworks.nova.edu/tqr/vol13/iss4/2/>
- Biermann, O. C., Ma, N. F., & Yoon, D. (2022). From tool to companion: Storywriters want AI writers to respect their personal values and writing strategies. *Designing Interactive Systems Conference*, 1209–1227. <https://doi.org/10.1145/3532106.3533506>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Brown, H. . (2001). *Teaching by Principles: An Interactive Approach to Language Pedagogy*. Addition Wesley Longman.
- Brown, H. D. (2000). *Principles of language learning and teaching (Vol. 4)*: Longman New York.
- Celce-Murcia, M., & Olshtain, E. (2000). *Discourse and context in language teaching: A guide for language teachers*. Cambridge University Press.
- Chaves, A. P., & Gerosa, M. A. (2021). How should my chatbot interact? A survey on social characteristics in human–chatbot interaction design. *International Journal of Human–Computer Interaction*, 37(8), 729–758. <https://doi.org/10.1080/10447318.2020.1841438>
- Daria, A., & Natalia, K. (2020). Using Debate to Develop Writing Skills for IELTS Writing Task 2 among STEM Students. *Journal of Language and Education*, 6(4 (24)), 30–43. <https://doi.org/10.17323/jle.2020.10424>
- Denzin, N. K. (2017). *The research act: A theoretical introduction to sociological methods*. Transaction publishers. <https://doi.org/10.4324/9781315134543>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., & Ahuja, M. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Escamilla, K., Butvilofsky, S., & Hopewell, S. (2018). What gets lost when English-only writing assessment is used to assess writing proficiency in Spanish-English emerging bilingual learners? *International Multilingual Research Journal*, 12(4), 221–236. <https://doi.org/10.1080/19313152.2016.1273740>
- Fitria, T. N. (2021). Grammarly as AI-powered English writing assistant: Students' alternative for writing English. *Metathesis: Journal of English Language, Literature, and Teaching*, 5(1), 65–78. <https://doi.org/10.31002/metathesis.v5i1.3519>
- Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences.

- Minds and Machines*, 30, 681–694. <https://doi.org/10.1007/s11023-020-09548-1>
- Frost, K., Clothier, J., Huisman, A., & Wigglesworth, G. (2020). Responding to a TOEFL iBT integrated speaking task: Mapping task demands and test takers' use of stimulus content. *Language Testing*, 37(1), 133–155. <https://doi.org/10.1177/0265532219860750>
- Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, 3, 100055. <https://doi.org/10.1016/j.caeai.2022.100055>
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, 26(2). <https://doi.org/10.14705/rpnet.2022.61.1446>
- Harmer, J. (2004). *How to Teach Writing*. England: Pearson Education Limited.
- He, T. (2019). The sentimental fools and the fictitious authors: rethinking the copyright issues of AI-generated contents in China. *Asia Pacific Law Review*, 27(2), 218–238. <https://doi.org/10.1080/10192557.2019.1703520>
- Hirvela, A., & Belcher, D. (2016). 27 Reading/Writing and Speaking/Writing Connections: The Advantages of Multimodal Pedagogy. *Handbook of Second and Foreign Language Writing*, 587–612. <https://doi.org/10.1515/9781614511335-030>
- Hosseini, M., Taghizadeh, M. E., Abedin, M. J. Z., & Naseri, E. (2013). In the importance of EFL learners' writing skill: Is there any relation between writing skill and content score of English essay test. *International Letters of Social and Humanistic Sciences*, 6(1), 1–12. <https://doi.org/10.18052/www.scipress.com/ilshs.6.1>
- Jasper, M. A. (2005). Using reflective writing within research. *Journal of Research in Nursing*, 10(3), 247–260. <https://doi.org/10.1177/174498710501000303>
- Kessler, G. (2018). Technology and the future of language teaching. *Foreign Language Annals*, 51(1), 205–218. <https://doi.org/10.1111/flan.12318>
- Kim, S. (2021). Prepping for the TOEFL iBT Writing test, Gangnam style. *Assessing Writing*, 49, 100544. <https://doi.org/10.1016/j.asw.2021.100544>
- Li, J., & Li, M. (2018). Turnitin and Peer Review in ESL Academic Writing Classrooms. *Language Learning & Technology*, 22(1), 27–41. <https://eric.ed.gov/?id=EJ1168811>
- Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: how may AI and GPT impact academia and libraries? *Library Hi Tech News*. <https://doi.org/10.1108/lhtn-01-2023-0009>
- Maiden, N., Lockerbie, J., Zachos, K., Wolf, A., & Brown, A. (2023). Designing new digital tools to augment human creative thinking at work: An application in elite sports coaching. *Expert Systems*, 40(3), e13194. <https://doi.org/10.1111/exsy.13194>
- Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development*. Springer. <https://doi.org/10.1057/9780230511804>
- McKee, H. A., & Porter, J. E. (2020). Ethics for AI writing: The importance of rhetorical context. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 110–116. <https://doi.org/10.1145/3375627.3375811>
- Muluk, S., Zainuddin, Z., & Dahliana, S. (2022). Flipping an IELTS Writing Course: Investigating its impacts on students' performance and their attitudes. *Studies in English Language and Education*, 9(2), 591–612. <https://doi.org/10.24815/siele.v9i2.23314>
- Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial. *Heliyon*, 7(5), e07014. <https://doi.org/10.1016/j.heliyon.2021.e07014>

- Ng, D. T. K., Luo, W., Chan, H. M. Y., & Chu, S. K. W. (2022). Using digital story writing as a pedagogy to develop AI literacy among primary students. *Computers and Education: Artificial Intelligence*, 3, 100054. <https://doi.org/10.1016/j.caeai.2022.100054>
- Odell, L. (1980). The process of writing and the process of learning. *College Composition and Communication*, 31(1), 42–50. <https://doi.org/10.2307/356632>
- Owen, N., Shrestha, P. N., & Hultgren, A. K. (2021). Researching academic reading in two contrasting English as a medium of instruction contexts at a university level. *ETS Research Report Series*, 2021(1), 1–28. <https://doi.org/10.2307/356632>
- Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and artificial intelligence: An experiential perspective. *Journal of Marketing*, 85(1), 131–151. <https://doi.org/10.1177/0022242920953847>
- Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: L. Erlbaum Associates. [https://doi.org/10.1016/0091-2182\(87\)90137-6](https://doi.org/10.1016/0091-2182(87)90137-6)
- Shi, Z., Liu, F., Lai, C., & Jin, T. (2022). Enhancing the use of evidence in argumentative writing through collaborative processing of content-based automated writing evaluation feedback. *Language Learning & Technology*, 106–128. <https://eric.ed.gov/?id=EJ1347885>
- Tang, A., Tam, R., Cadrin-Chênevert, A., Guest, W., Chong, J., Barfett, J., Chepelev, L., Cairns, R., Mitchell, J. R., & Cicero, M. D. (2018). Canadian Association of Radiologists white paper on artificial intelligence in radiology. *Canadian Association of Radiologists Journal*, 69(2), 120–135. <https://doi.org/10.1016/j.carj.2018.02.002>
- van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: five priorities for research. *Nature*, 614(7947), 224–226. <https://doi.org/10.1038/d41586-023-00288-7>
- Yu, G. (2010). Lexical diversity in writing and speaking task performances. *Applied Linguistics*, 31(2), 236–259. <https://doi.org/10.1093/applin/amp024>
- Zhao, X. (2022). Leveraging artificial intelligence (AI) technology for English writing: Introducing wordtune as a digital writing assistant for EFL writers. *RELC Journal*, 0(0), 1–5. <https://doi.org/10.1177/00336882221094089>