EFL Students' Perceptions of Language Learning and Assessment: Pre and in Generative AI (GenAI) Era[1]

Huan Zhao
Massey University
Auckland, New Zealand

Abstract

Anecdotal evidence suggests that the use of GenAI has both advantages and disadvantages in students' learning and assessment processes. However, despite the increasing volume, few studies have been conducted to explore English as a Foreign Language (EFL) students' perceptions towards language learning and assessments pre and in GenAI era. Therefore, employing a qualitative research approach, mainly through interviews and thematic analysis, the study aims to discover in-depth information about features and patterns of tertiary EFL students' perceptions and perception changes, regarding the impact of GenAI on language learning and the fairness of assessments in China. Findings of the study reveal that students do not embrace the idea of banning the use of GenAI in their learning and assessment, while they believe this could lead to not only advantages but also negative outcomes. Their perception towards assessment has shifted after the emergence of GenAI, compared with pre-GenAI era. To mitigate the negative consequences of GenAI in learning and assessment, participants presented some key suggestions. Tertiary education policy makers could draw on the findings of the study to create well-informed guidelines and strategies to effectively and ethically implement the use of GenAI tools, ultimately improving the teaching and learning experiences in Chinese and worldwide tertiary education sectors.

Keywords: GenAI, EFL learning; language assessment; tertiary education.

Huan Zhao

Resumen

Percepciones de estudiantes de EFL sobre el aprendizaje y la evaluación de lenguas: Antes y durante la era de la IA Generativa (GenAI)

La experiencia de los docentes sugiere que el uso de GenAI tiene ventajas y desventajas en los procesos de aprendizaje y evaluación de los estudiantes. Sin embargo, se han realizado pocos estudios para explorar las percepciones de los estudiantes de inglés como lengua extranjera (EFL) sobre el aprendizaje y la evaluación de idiomas antes y durante la era GenAI. Por lo tanto, empleando un enfoque de investigación cualitativa, principalmente a través de entrevistas y análisis temático, el estudio tiene como objetivo descubrir información en profundidad de las características y los patrones de las percepciones de los estudiantes de EFL y los cambios de percepción, en relación con el impacto de GenAI en el aprendizaje de idiomas y la imparcialidad de las evaluaciones. Los resultados del estudio revelan que los estudiantes no aceptan la idea de prohibir el uso de GenAI en su aprendizaje y evaluación, aunque creen que esto podría acarrear no sólo ventajas sino también resultados negativos. Su percepción de la evaluación ha cambiado tras la aparición de la GenAI, en comparación con la era pre-GenAI. Para mitigar las consecuencias negativas de la GenAI en el aprendizaje y la evaluación, los participantes presentaron algunas sugerencias clave. Los responsables políticos de la educación terciaria podrían basarse en las conclusiones del estudio para elaborar directrices y estrategias bien informadas que permitan aplicar de forma eficaz y ética el uso de las herramientas GenAI en los sectores de la educación terciaria de todo el mundo.

Palabras clave: GenAI; aprendizaje de EFL; evaluación lingüística; educación terciaria.

Résumé

Perceptions de l'apprentissage et de l'évaluation des langues chez des étudiants d'anglais : Avant et pendant l'ère de l'IA générative (GenAI)

Des preuves anecdotiques suggèrent que l'utilisation de la GenAI présente à la fois des avantages et des inconvénients dans les processus d'apprentissage et d'évaluation des étudiants. Cependant, peu d'études ont été menées pour explorer les perceptions des étudiants d'anglais langue étrangère (EFL) concernant l'apprentissage et l'évaluation des langues avant et pendant l'ère de la GenAI. Par conséquent, en utilisant une approche de recherche qualitative, principalement par le biais d'entretiens et d'analyses thématiques, l'étude vise à découvrir des informations approfondies sur les caractéristiques et les modèles des perceptions et des changements de perception des étudiants en anglais langue étrangère, concernant l'impact de la GenAI sur l'apprentissage des langues et l'équité des évaluations. Les résultats de l'étude révèlent que les étudiants n'adhèrent pas à l'idée d'interdire l'utilisation de la GenAI dans leur apprentissage et leur évaluation, tout en pensant que cela pourrait entraîner non seulement des avantages, mais aussi des résultats négatifs. Leur perception de l'évaluation a changé après l'émergence de la GenAI, par rapport à l'ère pré-GenAI. Pour atténuer les conséquences négatives de la GenAI sur l'apprentissage et l'évaluation, les participants ont présenté quelques suggestions clés.

Les décideurs de l'enseignement supérieur pourraient s'inspirer des résultats de l'étude pour élaborer des lignes directrices et des stratégies bien informées afin de mettre en œuvre de manière efficace et éthique l'utilisation des outils de la GenAI dans les secteurs de l'enseignement supérieur à l'échelle mondiale.

Mots-clés : GenAI ; apprentissage de l'anglais langue étrangère ; évaluation des langues ; enseignement supérieur.

ABOUT THE AUTHOR

Huan Zhao, PhD is Learning Advisor at the Centre for Learner Success, Massey University, New Zealand. She was a lecturer at Fudan University, Shanghai, China, prior to her embarking on her PhD study in New Zealand. She has since graduated with the degree of Doctor of Philosophy in Education (Applied Linguistics and TESOL) from The University of Auckland. Her research interests include language teacher cognition, language teacher education, teacher professional development, and the learning and teaching of English as a foreign language (EFL), especially EFL writing.

Email: h.zhao2@massey.ac.nz

HOW TO CITE THIS ARTICLE

INTRODUCTION

GenAI and Tertiary Education

The release of GenAI tools has had a huge impact on tertiary teaching and learning. Since the launch of GenAI tools, especially ChatGPT in 2022, there has been a surge of peer-reviewed articles about the impact of students' use of GenAI (Adiguzel et al., 2023; Crawford et al., 2023; Warschauer et al., 2023; Zhou, Wu & Qu, 2024). A coined term, "GenAI assessment literacy", introduced by Moorehouse et al. (2023, p. 9), has appeared in many journal articles. Regarding utilising GenAI as a learning tool, studies reported that students have mostly recognised the potential of GenAI tools (Baidoo-Anu et al., 2024). Students are generally satisfied with the online learning and online assessment in different contexts, such as a study in Vietnam (Tran et al., 2021), which found that students rated satisfaction with applying GenAI in their learning as 4.08 out of 5. Students reported benefits of utilising GenAI tools when completing written assignments include increased and enhanced linguistic capability, while the concerns consisted of risks of lowering academic integrity, plagiarism (Moorhouse et al., 2023), and over-reliance on AI (Zhou, Zhang & Chan, 2024). Most studies reported that students prefer using GenAI tools to assist their linguistic capability (Attewell, 2024). This indicates that EFL students are a cohort that is more likely to use GenAI in their assessment preparation. The curriculum design of the English major reassures that the majority of the assignments are written assignments or include written assignments in their exams, especially for the first-year EFL students. Therefore, the first-year EFL student cohort, majoring in English, was chosen as participants for this study.

While the emergence of GenAI has a huge impact on students' learning, it has also affected the assessment practices of higher education institutions worldwide (Moorhouse et al., 2023). Also, the nature of the traditional assessment system means that it is extremely vulnerable in the GenAI era (Overono & Ditta, 2023). This means that the prevalence and easy access to GenAI tools have brought challenges to the tertiary assessment system. Studies discovered that students raised concerns about the potential for increased cheating (Zhou, Zhang, & Chan, 2024). It is also worth mentioning that students mentioned the need for a clear guideline for integrating AI and GenAI into their assignment writing (Zhou, Zhang, & Chan, 2024). Recent literature has called for a reform of assessment by designing changes for the age of AI (Lodge et al., 2023; TEQSA, 2023; Vallis, 2024). Meanwhile, there seems to be a lack of assessment guidelines consistent with the GenAI era (Overono & Ditta, 2023). Before the emergence of GenAI, students' concerns regarding assessments were primarily focused on human biases and inconsistencies in grading (Overono & Ditta, 2023). After the emergence of GenAI, while some of these concerns persist, new issues related to the transparency and accountability of GenAI systems have emerged. One major issue is the potential for academic dishonesty, as students might use GenAI tools to complete assignments without truly understanding the material (Dong, 2024).

Some universities have proposed tools or approaches to help mitigating limitations of GenAI in the field of assessment, such as the un-grading approach (Overono & Ditta, 2023), the assessment validity framework at Unitech in New Zealand (Kabbar & Barmada, 2024) and the two-lane approach in University of Sydney, Australia (Liu & Bridgeman, 2023). Warschauer et al. (2023) proposed a five-part pedagogical framework so that GenAI could be better integrated into supporting EFL learners' writing development. Given the urgent need to adapt to the new era, universities must work hard to redesign the assessment system. A recent study has found that the number of universities that have developed and modified their assessment guidelines to address GenAI use is unsatisfactory (Moorhouse et al., 2023).

Scholars have realised that students' voices are crucial to inform responsible GenAI integrating in the tertiary education sector (Attewell, 2024), especially in the Chinese educational context, which is extremely examination-driven (Berry, 2011; Zhou, Zhang, & Chan, 2024). Meanwhile, although the trend of research of GenAI in higher education has shifted from the US to China (Crompton & Burke, 2023), there is a lack of in-depth studies on Chinese EFL students' perceptions towards learning and assessment in the GenAI era. The study aims to discover the dynamic nature of students' perception change of language learning and assessments after the rise of GenAI. This study explores how EFL students in Chinese universities perceive language learning and assessment before and after the rise of GenAI.

METHODOLOGY

Research Design

Based on the exploratory nature of the study, the qualitative approach was chosen for its ability to facilitate an in-depth examination of participants' experiences and perceptions (Creswell et al., 2007), which are essential for understanding the nuanced changes in perceptions towards learning and assessments. Considering the objective of this study, semi-structured interviews were employed to uncover the deep understanding of the differences for three reasons. Firstly, interview data provides more in-depth information than quantitative approaches. This method allows for flexibility in exploring participants' thoughts and experiences while ensuring that key topics are covered. Secondly, interview data are more valid because they are collected through direct contact with participants (Denscombe, 2017). Thirdly, the interview approach allows a dialogic discussion, enabling the researcher to discover any hidden perceptions of the students (Creswell et al., 2007).

Data Collection

Convenience and purposive sampling methods were used to recruit participants for this study. Given that there are a large number of universities in China, and they are from different tiers, two universities of different tiers were selected, intending to avoid bias in research. One is a "Double First-Class Initiative" university, while the other

one is a non- "Double First-Class Initiative"[2] university. Two teachers from the English departments of the two universities, which the researcher has personal contact with, were contacted by the researcher to recruit two students from each university to participate in the study. As requested by the researcher, the two teachers sent recruiting advertisements to their classes, consisting of 39 students altogether. Four participants expressed their interests. All participants signed the research consent forms and received the participant information sheet. The participants are referred to as Participant 1, 2, 3 and 4 in this study. The interviews were semi-structured interviews, which took 20-30 minutes. Semi-structured interviews can yield detailed and rich responses (Creswell et al., 2007) because open-ended questions are less structured. Also, you can build a better connection with participants and ask follow-up questions based on the answers, thereby eliciting richer, more nuanced data. The interviews were conducted in Chinese as requested by the interviewees. The use of Chinese helped to enhance the depth of the interview data. It also shortened the distance between the researcher and the participants. The interview outline includes questions about students' perceptions of language learning and language assessment fairness, both before and after the introduction of GenAI tools.

This is the semi-structured Interview outline:

1. Have you used GenAI with your assignment writing?
2. What are the benefits and concerns related to using GenAI in learning and assessment?
3. What is your perception of language assessment before GenAI emergence?
4. How about after the emergence of GenAI? What are the key differences?
5. How could the university help you to better integrate GenAI into your English learning?
6. How could the assessment procedures or policies adapted to suit the needs of GenAI era?

Data Analysis

Thematic analysis was used to analyse the interview data. This method involves identifying, analysing, and reporting patterns (themes) within the data. Given the relatively small number of participants, the interview data was manually analysed. The process includes the following steps. Firstly, translating, transcribing interviews and reading through the data to familiarise with the content. Secondly, manually generating initial codes from the data that capture major features relevant to the research aim. Thirdly, refining and defining the codes, ensuring each code tells a coherent story about the data. Lastly, writing up the findings, using quotes from participants to illustrate key codes.

---

[2] In China, top universities are classified as "Double First-Class Initiative" universities (Ministry of Education of China, n.d.).

Ensuring Validity and Reliability

To ensure the validity and reliability of the study, several strategies were employed. The first strategy was member check: Sharing findings with participants to verify the accuracy of the interpretations. In this research, the findings were sent back to the participants once transcribed. With the participants' consent, the research was then carried out to the next page. The second strategy was peer review. The researcher requested two expert colleagues to review the research process and findings to provide feedback to improve validity and reliability.

Ethical Considerations

Participants were informed about the study's purpose, their right to withdraw at any time, and the measures taken to ensure confidentiality and anonymity. Informed consents were obtained from all participants before data collection began.

RESULTS AND DISCUSSION

Benefits and Concerns of Using GenAI in Language Learning

Overall, participants reported that we should not ban the use of GenAI but embrace it into our learning and teaching (Crompton & Burke, 2023; Warschauer et al., 2023). However, participants reported that the use of GenAI tools had both positive and negative impacts (Thorp, 2023; Zhou et al., 2024). On the positive side, GenAI tools provided valuable support for learning, helping students to improve their skills and gain confidence. On the negative side, there were concerns about the potential for over-reliance on GenAI, which might undermine the development of independent learning skills.

Benefits

All participants reported they have used GenAI tools in English writing for various purposes, including drafting the structure of written assessments (all participants), practicing language skills (participants 1 and 3), and seeking feedback on their drafts (participant 4). This is consistent with the research findings of Chan and Hu (2023), that most students embrace the use of GenAI into their learning. Participants appreciated the immediate feedback provided by GenAI tools, which helped them identify and correct errors in their writing (Attewell, 2024; Jiang et al., 2023). They also valued the personalized learning experience, as the GenAI could adapt to their individual needs and provide tailored suggestions or feedback in an accessible way (Jiang et al., 2023). Additionally, the convenience of having an available tool for practice, especially practicing "grammatical knowledge" (participant 3), was seen as a significant advantage. Participant 1 mentioned, "As a non-native English speaker, I

always use Grammarly to check my draft before submitting them because my grammar was not very good." Participant 3 stated:

> It feels like my confidence has boosted after using AI tools like Grammarly. Although I found it struggling at the beginning of using GenAI, now I have found it has taught me a lot, as I think the frequency of committing the same grammatical mistakes has been largely reduced. (participant 3)

This is consistent with the findings of Overono and Ditta (2023), who proposed that the emergence of ChatGPT and the continued evolution of GenAI pose an exciting challenge for higher education—to refocus students on work that is personally meaningful and accomplishable only by humans.

Concerns

On the other hand, participants also expressed concerns about the use of GenAI, especially regarding the reliability and accuracy of GenAI-generated feedback. Some noted that the GenAI occasionally provided incorrect or irrelevant suggestions, which could be either at a superficial level (Dong, 2024) or even misleading. There were also worries about the potential for over-reliance on GenAI tools, which might hinder the development of originality, independent critical thinking and problem-solving skills. Participant 2 stated, "I am afraid I do not think it is safe to share my draft with GenAI tools, such as Grammarly, to check, as I think it might be shared with someone else so that my originality might be challenged". This aligns with Baidoo-Anu et al.'s (2024) study, which found that students are concerned about a lack of originality in assignments. Participant 4 stated, "My teacher is very strict. They said they are going to use GenAI detector to check my work, so I do not use it to complete my work, but I cannot guarantee that my classmates do the same as me." This aligns with studies which have found that concerns of students include integrity policy violations (Baidoo-Anu et al., 2024; Moorhouse et al., 2023).

Participants also complained it was difficult for them to decide which GenAI tool to use, given that there are so many tools available currently. Participant 4 stated:

> After the emergence of so many GenAI tools, there are many similar tools online, such as Wenxin Yiyan. I feel it is difficult for me to choose which tool is reliable to use. It is a kind of waste of time. This is one reason that I am not a huge fan of using GenAI tools to complete my assignment. (participant 4)

The overreliance on GenAI could compromise students' efforts and ability to "develop writing competency" (participant 4), as well as critical thinking ability. Additionally, the potential for academic dishonesty, where students might use GenAI to complete assignments without genuine understanding, poses a challenge to academic integrity.

Language Assessment Fairness Before the Emergence of GenAI

Fairness in assessments is often linked to the transparency of the assessment criteria and the perceived impartiality of the grading process. In this study, participants reported that their perceptions of language assessments before the emergence of GenAI tools were primarily centred around the subjective nature of grading. All participants felt that different instructors had "varying standards" (participant 1), which led to inconsistencies in grading. Students often questioned the fairness of assessments, citing examples where they believed personal biases of instructors influenced their grades. Some students noted that their performance could vary significantly depending on the instructor or the specific assessment task. They felt that the lack of standardized grading rubrics contributed to this inconsistency. Participant 3 stated, "In our faculty, there are two Intensive English teachers. My superior told me that one teacher tended to mark higher than the other. They told me to try to get into that teacher's class if possible." They also highlighted the lack of transparency in the grading criteria, which made it difficult for them to understand how their performance was evaluated. Participant 2 mentioned, "For one of my papers, the marking rubric is very general, which made me confused when completing the assignment."

Language Assessment Fairness After the Emergence of GenAI

Research findings show that students' perceptions of language assessments after the emergence of GenAI tools began to shift. While some of the pre-existing concerns remained, new issues related to the use of GenAI when preparing for written assessments emerged. Students continued to question the fairness of assessments, particularly in the era when GenAI tools were used to help complete assignments. All participants have raised this concern. For instance:

> I feel the assessment is not as fair as before, after the emergence of GenAI. For example, I know some of my classmates use GenAI to help with their writing. However, I have not used GenAI at all in completing my work. I felt frustrated when I found their scores were higher than mine. (participant 1)

Same with the finding of one study that students complained that the AI detectors currently are not capable of detecting AI-generated output (Chan & Hu, 2023), all participants in this study reported that guidelines should adapt to the technological change "completely and urgently" (participant 2 and 3) to improve the fairness of language assignment assessment.

Ways to Mitigate Limitations of GenAI for EFL students

GenAI Literacy Training

Despite all participants reporting that we should embrace GenAI tools into our learning and teaching, they reported that they processed low GenAI literacy. As mentioned by participant 3 that there are "so many writing generating tools, and I am very confused about which one is the best one to use". Participant 3 also mentioned, "I am afraid I would waste a lot of time on scrutinising the GenAI tools. Therefore, I did not want to start. Instead, I would prefer spending time on completing the task myself." This is consistent with findings of Attewell (2024), which argue for a need to cultivate GenAI literacy both for learners and trainers. However, findings of the study also support those of Warschauer et al. (2023) and Dong (2024), that we need to "have clear guidance on how to choose the correct GenAI tools" (participant 2) and we need to be "trained to use them effectively and ethically" (participant 2).

Assessment Guideline or Policy Change

All participants reported a need to change the assessment guidelines or policies. Participants 2 and 3 mentioned that the change needs to happen "urgently". Participant 2 argued, "The assignment guideline needs to be clearer about the involvement of GenAI tools. They need to specify to what extent GenAI could be used in assignment writing. Currently, it is very unclear, and needs to be changed urgently." This is consistent with many studies proposing the need to develop evidence-based guidelines and policies for GenAI integration (Chan & Lee, 2023).

> I think the design of assessments needs to be changed. I prefer the type of assessment where students sit together and do the assessment without access to outside sources such as the internet. In this way, it seems like I prefer the old school test, not written take-home assignments. (participant 3)

This is in line with studies' findings that the number of universities that have already developed policy changes correspondingly to deal with the emergence of GenAI is unsatisfactory, and also, most of them are of low quality as they were produced in haste (Moorhouse et al., 2023). In the foreseeable future, as GenAI may potentially be widely used in formal academic settings, institutions should also develop policies and provide formal guidance on the use of GenAI (Crompton & Burke, 2023).

Participant 3 recommended the use of peer assessment (Zhao, 2018), mentioning:

> During my high school time, our English teacher tended to assign us to work in pairs to mark each other's writing. I felt this is a good way as sometimes we

could provide very effective feedback to classmates as we are learning the same thing, and we are of the same age. (participant 3)

This echoes the partial findings in Zhao (2018) that tertiary education tutors have limited information on peer assessment. Therefore, training is needed to enhance tutors' awareness about peer assessment (Zhao, 2018).

CONCLUSION, IMPLICATIONS AND LIMITATIONS

The study's findings underscore the evolving landscape of students' perceptions towards language learning and language assessments in the context of GenAI tools. The introduction of these tools has brought both opportunities and challenges, reshaping students' perceptions of language learning and assessments.

Key Findings

All participants mentioned that we should learn to incorporate GenAI into our learning, instead of pushing it away from us. Most of them have positive attitudes towards integrating GenAI into their learning, but have reported a lack of GenAI literacy. Before the emergence of GenAI, students' concerns about assessment were primarily centred around the subjective nature of grading and the lack of transparency in assessment criteria. These issues persisted post-GenAI, but new concerns emerged, particularly regarding fairness and inconsistency of assessment procedures and guidelines.

Based on the findings of this study, several recommendations can be made to enhance the use of GenAI tools in language assessments within tertiary education. These recommendations aim to address the challenges identified and leverage the benefits of GenAI to improve the teaching and learning experience.

AI literacy training

Providing comprehensive training for educators on how to effectively integrate GenAI tools into their teaching and assessment practices is recommended. This training should cover the technical aspects of using GenAI tools, as well as strategies for interpreting AI-generated feedback. This includes providing clear explanations of how GenAI works and evaluates written work and training towards how effectively, efficiently and ethically it incorporates GenAI with their assignment writing. The training could cover topics such as academic integrity, critical thinking, and the limitations of GenAI.

Introducing a Hybrid Assessment Model

It is advisable to use GenAI tools to complement, not to replace, human judgment. For instance, GenAI can assist initial grading and provide feedback on grammar and syntax, while human evaluators can assess more nuanced aspects of language use,

such as context and creativity. To explore the effectiveness of this hybrid approach, piloting assessment models that combine both AI and human input is encouraged.

Ethical Guidelines

It is recommended to create clear ethical guidelines for the use of GenAI in education. These guidelines should address issues such as data privacy, academic integrity, and the responsible use of GenAI tools.

Policy Development

To guide the integration of GenAI in education, it is essential to develop policy frameworks at the institutional and national levels. These policies should ensure that GenAI tools are used in a way that promotes fairness in assessments.

Implications

This study contributes to the growing body of knowledge on the impact of GenAI in education, particularly in the context of English language learning and assessments. By providing insights into students' perceptions before and after the introduction of GenAI, the study offers valuable guidance for educators and policymakers aiming to enhance the fairness of assessments in the era of GenAI. Firstly, the study highlights the need for a balanced approach to integrating GenAI into tertiary learning and assessments for EFL students. Secondly, the findings of this study have several implications for educators and policymakers. Insights from this study can inform policy development, aiming to create well-informed guidelines as a response to the need for clear guidelines on the ethical use of GenAI in assessments to ensure fairness. Educators should be trained on how to effectively integrate GenAI tools into their teaching and assessment practices. Policymakers should consider developing policies that promote transparency and accountability in the use of GenAI in education.

Limitations and Recommendations for Future Research

The study acknowledges potential limitations, such as the reliance on self-reported data, which may be subject to bias (Creswell et al., 2007). Additionally, the findings may not be generalizable beyond the specific context of Chinese tertiary education. However, the in-depth qualitative approach provides valuable insights that can inform future research and practice.

Future research should continue to explore the long-term impact of GenAI on students' perceptions and learning outcomes. Investigating the effectiveness of hybrid assessment models and developing GenAI tools that can better evaluate aspects of language learning outcomes will be crucial. Additionally, examining the broader implications of GenAI in education, including its impact on teaching practices and student engagement, will provide valuable insights for the future.

REFERENCES

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology, 15*(3), ep429. https://doi.org/10.30935/cedtech/13152

Attewell, S. (2024). *Student perceptions of generative AI*. Jisc. https://beta.jisc.ac.uk/reports/student-perceptions-of-generative-ai

Baidoo-Anu, D., Asamoah, D., Amoako, I., & Mahama, I. (2024). Exploring student perspectives on generative artificial intelligence in higher education learning. *Discover Education, 3*(1), 98. https://doi.org/10.1007/s44217-024-00173-z

Berry, R. (2011). Assessment trends in Hong Kong: Seeking to establish formative assessment in an examination culture. *Assessment in Education: Principles, Policy & Practice, 18*(2), 199–211. https://doi.org/10.1080/0969594X.2010.527701

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*, 45. https://doi.org/10.1186/s41239-023-00411-8

Chan, C. K. Y., & Lee, K. K. W. (2023). The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and

learning than their Gen X and millennial generation teachers? *Smart Learning Environments, 10*(1), 60. https://doi.org/10.1186/s40561-023-00269-3

Crawford, J., Cowling, M., & Allen, K.-A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching & Learning Practice, 20*(3), 1–19. https://doi.org/10.53761/1.20.3.02

Creswell, J. W., Hanson, W. E., Plano, V. L. C., & Morales, A. (2007). Qualitative Research Designs: Selection and Implementation. *The Counseling Psychologist, 35*(2), 236–264. https://doi.org/10.1177/0011000006287390

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education, 20*(1), 22. https://doi.org/10.1186/s41239-023-00392-8

Denscombe, M. (2017). *The good research guide: For small-scale social research projects* (E-book). McGraw-Hill Education (UK).

Dong, L. (2024). ChatGPT in language writing education: Reflections and a research agenda for a ChatGPT feedback engagement framework. *Language Teaching Research Quarterly, 43*, 121–131. https://doi.org/10.32038/ltrq.2024.43.07

Jiang, Z., Xu, Z., Pan, Z., He, J., & Xie, K. (2023). Exploring the role of artificial intelligence in facilitating assessment of writing performance in second language learning. *Languages, 8*(4), Article 4. https://doi.org/10.3390/languages8040247

Kabbar, E., & Barmada, B. (2024, July 24). *Assessment validity in the era of generative AI tools*. SciSpace. https://doi.org/10.34074/proc.240105

Liu, D., & Bridgeman, A. (2023, July 12). *What to do about assessments if we can't out-design or out-run AI?* Teaching@Sydney. https://educational-innovation.sydney.edu.au/teaching@sydney/what-to-do-about-assessments-if-we-cant-out-design-or-out-run-ai/

Lodge, J., Howard, S., & Bearman, M. (2023). *Assessment reform for the age of artificial intelligence*. Tertiary Education Quality and Standards Agency. https://www.teqsa.gov.au/guides-resources/resources/corporate-publications/assessment-reform-age-artificial-intelligence

Ministry of Education of China. (n.d.). *China to further promote the Double First-Class Initiative*. Ministry of Education of the People's Republic of China. Retrieved March 17, 2025, from http://en.moe.gov.cn/news/press_releases/202203/t20220301_603547.html

Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open, 5*, 100151. https://doi.org/10.1016/j.caeo.2023.100151

Overono, A. L., & Ditta, A. S. (2023). The rise of artificial intelligence: A clarion call for higher education to redefine learning and reimagine assessment. *College Teaching*, 1–4. https://doi.org/10.1080/87567555.2023.2233653

TEQSA. (2023). *Assessment reform for the age of artificial intelligence.* https://www.teqsa.gov.au/sites/default/files/2023-09/assessment-reform-age-artificial-intelligence-discussion-paper.pdf

Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science, 379*(6630), 313. https://doi.org/10.1126/science.adg7879

Tran, T. T. T., Nguyen, P. L. A., Nguyen, N. H. T., & Tran, D. T. (2021). An empirical study on students' perception and satisfaction towards online assessment and testing in tertiary education. *Advances in Social Science, Education and Humanities Research, 621*, 207–216. https://doi.org/10.2991/assehr.k.211224.021

Vallis, C. (2024). Authentic assessment in higher education: The spectre of lost futures. Teaching in Higher Education. *Teaching in Higher Education, 30*(3), 744–751. https://doi.org/10.1080/13562517.2024.2362217

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. *Social Science Research Network*. https://doi.org/10.2139/ssrn.4404380

Zhao, H. (2018). Exploring tertiary English as a foreign language writing tutors' perceptions of the appropriateness of peer assessment for writing. *Assessment & Evaluation in Higher Education, 43*(7), 1133–1145. https://doi.org/10.1080/02602938.2018.1434610

Zhou, X., Zhang, J., & Chan, C. (2024). Unveiling students' experiences and perceptions of artificial intelligence usage in higher education. *Journal of University Teaching and Learning Practice, 21*(06). https://doi.org/10.53761/xzjprb23

Zhou, Y., Wu, X., & Qu, K. (2024). The Role of ChatGPT in English language learning: A hedonic motivation perspective on student adoption in Chinese universities. *Language Teaching Research Quarterly, 43*, 132–154. https://doi.org/10.32038/ltrq.2024.43.08