Evaluating the Impact of an AI Critical Friend on Student Research Proposals, Critical AI Literacy, and Transparent Collaborative Assessment Practices

Tim Gander, David Parsons
academyEX

This paper presents a work-in-progress report on an innovative approach to integrating generative AI capabilities into the assessment process, focusing on research proposal development. Specifically, the study explores the impact of incorporating an AI-powered "critical friend" that provides targeted feedback and guidance to students as they craft their research proposals. The aim is to evaluate how this AI-powered support system affects students' critical AI literacy, the quality and transparency of the collaborative assessment process, and, ultimately, the strength of the final research proposals. Through a mixed-methods approach combining surveys, focus groups, and analysis of student work and interactions with generative AI, the researchers seek to uncover key insights on the benefits, challenges, and ethical considerations of leveraging AI to enhance the research proposal development experience. The findings from this study will contribute to the growing body of literature on the role of AI in academic assessments and provide practical recommendations for educators and institutions looking to responsibly integrate these emerging technologies.

Keywords: AI, academic writing, research proposals, critical literacy, collaborative assessment

Introduction

Generative AI is rapidly being integrated into higher education, offering new avenues to support student learning. Beyond simply generating text, these tools are being explored for their potential to personalise and enhance the learning experience. Rojas et al. (2023) highlight how AI can tailor educational content to individual needs, while Bowles and Kruger (2023) emphasise its potential role in preparing students for a future workforce increasingly reliant on AI technologies. Ipek et al. (2023) provide a systematic review of educational applications of ChatGPT, highlighting its potential to transform teaching and learning processes through enhanced feedback mechanisms, personalised learning experiences, and improved student engagement. However, as with any new technology, careful consideration must be given to its ethical implications and effective implementation within existing pedagogical frameworks and assessment practices. Introducing AI-powered tools like ChatGPT may disrupt traditional notions of academic integrity and require educators to rethink assessment strategies (Fontenelle-Tereshchuk, 2024). Recommendations from TEQSA (Lodge et al., 2023) emphasise the need for a paradigm shift in assessment practices within AI-driven educational contexts. Their recommendations highlight the importance of moving beyond traditional assessment methods that focus solely on rote memorisation or potentially AI-generated content. Instead, they advocate for incorporating assessments that evaluate students' critical engagement with AI, their ability to discern biases, and their ethical awareness when utilising these powerful tools. Research on the unique intersection of AI-powered feedback and critical peer review to support transparent assessment in postgraduate education is limited. The objective of this research is to test the efficacy of employing an AI "critical friend" to support master's level students in a practice-based research course, with a specific focus on enhancing the quality of research proposals, fostering critical AI literacy, and promoting transparent and collaborative assessment practices.

Context

The study is set within a practice-based research course at a New Zealand tertiary organisation, where master's-level students are tasked with developing and presenting a research proposal as a core component of their curriculum. Traditionally, students have relied on multiple rounds of feedback from their instructors to refine their proposals, often resulting in siloed, linear, and enduring assessment processes focusing on the

outcome rather than the learning process of collaboratively developing a proposal. Students have reported integrating AI-generated text into research proposals, yielding mixed results. However, generative AI has evidently hindered some proposals and there is inequity between students who understand how to use the tools well and those who do not. This has prompted the researchers to investigate the potential benefits and challenges of proactively integrating an AI-powered "critical friend" into the assessment of the research proposal. Working closely with the students and the course instructors, the researchers have developed a GPT using OpenAI GPT builder and the GPT-4-o language model. This GPT is designed to engage with students throughout the research proposal development process, providing feedback on elements ranging from the aims and scope to the research methods and data analysis. What sets this study apart is its comprehensive examination of the intersection between AI feedback and peer review in a postgraduate education context. By leveraging an AI "critical friend," this research aims to provide a unique blend of automated and human feedback, enhancing the learning process through iterative, reflective practice.

Assessment reform

With the integration of AI in education, assessment practices are changing. With a lack of guidance in New Zealand, the Australian TEQSA Assessment reform for the age of Artificial Intelligence (Lodge et al., 2023) guidelines highlight the need for approaches that integrate AI in a thoughtful, transparent and ethical way. Lodge et al. (2023) write that in the age of AI, assessments should emphasise digital mastery and critical reflection, encouraging students to move beyond simply using AI tools, to understanding how to leverage them effectively and ethically. This includes critically evaluating AI output and considering potential biases and limitations. Furthermore, assessments should focus on both the process and application of using AI tools, evaluating a student's ability to formulate effective prompts, curate data, and iterate based on AI-generated feedback. Ultimately, assessments should challenge students to apply their AI-enhanced skills to solve real-world problems, encouraging them to think critically about the practical implications and potential impact of AI in their field.

These assessments fall into what Liu and Bridgeman (2023) describe as a 'lane two' assessment or assessment as learning. Some examples include a marketing course that could task students with creating AI-generated content and then presenting the ethical implications and legal boundaries of using such technology. Similarly, an engineering program might challenge students to design a robot using AI, requiring them to document the limitations and risks associated with their AI-driven design choices.

Lodge et al. (2023) advocate for a multi-method assessment strategy that moves beyond traditional exams and incorporates projects where students critically evaluate AI-generated content or design AI-powered solutions while considering ethical implications. This approach aligns with the concept of "authentic assessment," as discussed by Tai et al. (2022), where tasks reflect real-world contexts and require students to apply their knowledge and skills in meaningful ways. There is a need for a shift towards assessments that encourage higher-order thinking, collaboration, and a deeper understanding of AI's capabilities and limitations. AI can be integrated into assessment practices to promote collaboration and transparency, encouraging students to engage in peer feedback and co-creation of assessment tasks (Zellner, 2023). These approaches seek to empower students to critically engage with AI, understand its capabilities and limitations, and consider the ethical implications of its use - all within the context of authentic assessment tasks that align with real-world applications.

Challenges with generative AI and assessment

The rapid advancement of large language models has raised concerns about potential misuse in academic settings, such as students' reliance on AI-generated text to fulfil writing assignments (Cotton et al., 2023; Yan et al., 2023). This raises fundamental questions about academic integrity, the authenticity of student work, and the ability of educators to effectively evaluate and assess learning outcomes. In addition, the opaque nature of many AI systems poses a significant challenge, as students may struggle to understand the reasoning behind the feedback or suggestions provided by these tools, limiting their ability to critically engage with the technology and develop essential academic skills Zirar (2023).

Chatbots have been implemented in a range of contexts (Farazouli et al., 2023; İpek et al.,2023), but the literature suggests there are limitations around the contextual awareness and tailored feedback that can be provided by generative AI systems, which may impact the effectiveness and clarity of the writing support (Cotton et al., 2023). In a recent study by Banihashem et al. (2024) it was found that there were significant differences in the feedback provided by generative AI compared to peer feedback. Generative AI was more descriptive and focused on the structure of the writing but lacked the nuanced, contextual analysis that human reviewers were able to provide, suggesting that there is benefit from both approaches. Consequently, more nuanced approaches to integrating AI into the assessment process, focusing on enhancing student learning through a combination of AI-supported approaches and promoting critical engagement with the technology, are needed.

Designing a solution

This study aims to evaluate the effectiveness of interactions between learners and a GPT-based AI tool in three key areas: improving the quality of student research proposals, fostering critical AI literacy, and promoting transparent and collaborative AI-assisted assessment practices. By examining these aspects, the research seeks to understand the potential of AI tools to transform educational practices and support students in developing essential skills for the digital era.

Based on the guidelines presented by Lodge et al. (2023), we designed a multi-method assessment approach that integrates AI while maintaining academic integrity and emphasising critical thinking. Our solution shifts the focus from traditional summative project evaluations to more authentic assessments. It requires students to critically evaluate AI-generated feedback with their peers, considering ethical implications and limitations. This approach ensures transparency and academic integrity by incorporating student peer feedback and co-creating next steps and goals, fostering a collaborative learning environment.

Intervention:

1 **Initial Proposal Development**:
- Students draft initial proposals during scaffolded lectures over 4 weeks, covering essential components such as aims, scope, research methods, and data analysis.

2 **AI-Powered Feedback**:
- Students use the custom GPT tool to receive targeted feedback on their proposals, focusing on clarity, coherence, and relevance.

3 **Collaborative Group Sessions**:
- Students compare their original proposals with AI feedback in group sessions, discussing the ethical implications and potential biases, fostering critical AI literacy and reflective thinking.

4 **Final Proposal Submission**:
- Students submit a transcript and video of their group discussion, along with the revised proposal, emphasising transparency and collaborative assessment.

Research Questions

1) How effective is the GPT-based AI tool in providing constructive feedback on student research proposals?
2) In what ways does interacting with the GPT-based AI tool improve students' AI literacy?
3) How does the use of the GPT-based AI tool influence students' critical reflection on their research ideas?
4) In what ways does the GPT-based AI tool engage students in transparent and collaborative assessment practices?

Design Framework

This study employs a mixed-methods design, integrating both qualitative and quantitative data to provide a comprehensive evaluation of the impact of a GPT-based AI tool on student research proposals, AI literacy, and collaborative assessment practices. Data collection will occur in several phases:

Pre-Intervention: Strzelecki (2023) explores student acceptance and use of ChatGPT in higher education using relevant components from the Unified Theory of Acceptance and Use of Technology (UTAUT) scale (Venkatesh et al., 2003). This scale will be implemented through quantitative methods using questionnaires (similar to those used by Strzelecki) to measure student perceptions of GPT's impact on learning, perceived usefulness, ease of use, and intention to continue using the tool. In addition, questions regarding AI literacy using the framework from MacCallum et al. (2023) and levels of critical reflection while using the tools are included.

Intervention: Students will use the GPT-based AI tool to receive feedback on their project proposals. All interactions with the GPT tool will be recorded to facilitate detailed analysis of the feedback process. Collaborative meetings will be organised where students discuss the feedback provided by the GPT tool. These sessions will be recorded and transcribed for qualitative analysis. During focus groups, UTAUT themes will be discussed along with AI literacy and critical reflection.

Post-Intervention: Students will revise their proposals based on the feedback received from the GPT tool. The revised proposals will be collected to assess improvements and the effectiveness of the AI feedback. Survey data will also be collected on students' experiences with the GPT tool, their engagement in the collaborative process, and any changes in their AI literacy and critical reflection skills. This survey will also include UTUAT-related items to reassess and compare the effectiveness of the GPT tool with their previous experiences with AI tools.

Data Analysis

Pre- and post-intervention survey data will be analysed using statistical methods to measure changes in AI literacy, critical reflection, and engagement. Content analysis will be performed on transcripts of interactions with the GPT tool to evaluate the effectiveness and quality of the feedback provided. These interactions will be analysed using the Interaction Analysis Model (IAM) (Gunawardena et al., 1997) and will meet RQ 2 (AI literacy) and RQ3 (critical reflection). Thematic analysis will also be applied to focus group transcripts to identify common themes and patterns, particularly focusing on the depth of critical discussion and AI literacy (RQs 1-4). The quality of initial and revised research proposals will be compared using rubric-based evaluation RQ1 (feedback effectiveness), RQ4 (collaborative assessment). This analysis will assess the effectiveness of the GPT feedback and the collaborative review process in improving research proposals, in comparison to traditional teaching methods. By integrating the UTAUT framework into both surveys and focus groups, and by comparing AI literacy and critical reflection alongside student interactions with the GPT tool, this study systematically examines the factors influencing students' acceptance and use of the GPT-based AI tool.

Significance and expected outcomes

The innovative approach described in this concise paper aims to meet these recommendations by providing students with opportunities to critically engage with a GPT-based AI tool, fostering their understanding of AI's capabilities and limitations, and promoting transparent, collaborative assessment practices that enhance their critical AI literacy. This study expects to contribute to the growing body of research on AI in education, with the expectation of observing several significant outcomes. First, drawing upon the work of Song et al. (2024) and Yan et al. (2023), this research anticipates fostering enhanced AI literacy among student participants, deepening their understanding of AI's capabilities, limitations, and ethical dimensions. Second, aligning with the findings of Zirar (2023) and Yulia and Susilowati (2020), this study aims to cultivate critical thinking skills by encouraging students to engage thoughtfully and analytically with AI tools. Third, based on the research by Banihashem et al. (2024), this study expects to see an improvement in the quality of student work through the combined use of AI and peer feedback. Finally, echoing the insights of Bowles and Kruger (2023), this study aims to equip students with authentic and future-ready skills, preparing them to navigate and thrive in a world increasingly shaped by AI technologies.

References

Banihashem, S K., Kerman, N T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: peer- generated or AI-generated feedback? *International Journal of Educational Technology in Higher Education*, 21(23). https://doi.org/10.1186/s41239-024-00455-4

ASCILITE 2024

Navigating the Terrain:

Emerging Frontiers in Learning Spaces, Pedagogies, and Technologies

Bowles, D C., & Kruger, J S. (2023). *Generating Employable, Intelligent Graduates in a World With Generative AI: Thoughts for Educators*. *Pedagogy in Health Promotion*,9(2), 75-77. https://doi.org/10.1177/23733799231175171

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Farazouli, A., Pargman, T C., Laksov, K B., & McGrath, C. (2023). *Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices*. Taylor & Francis, 1-13. https://doi.org/10.1080/02602938.2023.2241676

Fontenelle-Tereshchuk, D. Academic writing and ChatGPT: Students transitioning into college in the shadow of the COVID-19 pandemic. *Discov Educ* **3**, 6 (2024). https://doi.org/10.1007/s44217-023-00076-5

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a Global Online Debate and the Development of an Interaction Analysis Model for Examining Social Construction of Knowledge in Computer Conferencing. *Journal of Educational Computing Research*, *17*(4), 397-431. https://doi.org/10.2190/7MQV-X9UJ-C7Q3-NRAG

Ipek, Z. H., Gözüm, A. İ. C., Papadakis, S., & Kallogiannakis, M. (2023). Educational Applications of the ChatGPT AI System: A Systematic Review Research. *Educational Process International Journal*, *12*(3), 26–55. https://doi.org/10.22521/edupij.2023.123.2

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://educationalinnovation.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., & Associates (2023). *Assessment Reform for The Age of Artificial Intelligence*. Tertiary Education Quality and Standards Agency. https://www.teqsa.gov.au/sites/default/files/2023-09/assessment-reform-age-artificial-intelligence-discussion-paper.pdf

MacCallum, K., Parsons, D., & Mohaghegh, M. (2023). Identifying the components of foundational Artificial Intelligence (AI) literacy - Early results from a Delphi study. *Australasian Society for Computers in Learning in Tertiary Education*, 157-166. https://doi.org/10.14742/apubs.2023.672

Ruiz-Rojas, L I., Acosta-Vargas, P., De-Moreta-Llovet, J., & González, M. (2023). Empowering Education with Generative Artificial Intelligence Tools: Approach with an Instructional Design Matrix. *Multidisciplinary Digital Publishing Institute*, 15(15), 11524-11524. https://doi.org/10.3390/su151511524

Song, Y., Weisberg, L. R., Zhang, S., Tian, X., Boyer, K. E., & Israel, M. (2024). A framework for inclusive AI learning design for diverse learners. *Computers and Education: Artificial Intelligence*, *6*(November 2023), 100212. https://doi.org/10.1016/j.caeai.2024.100212

Strzelecki, A. (2023). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive Learning Environments*, 1–14. https://doi.org/10.1080/10494820.2023.2209881

Tai, J., Ajjawi, R., Bearman, M., Boud, D., Dawson, P., & Jorre de St Jorre, T. (2022). Assessment for inclusion: rethinking contemporary strategies in assessment design. *Higher Education Research & Development*, *42*(2), 483–497. https://doi.org/10.1080/07294360.2022.2057451Venkatesh.,

Morris, C F., Davis, S R., & Davis, S R. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425-425. https://doi.org/10.2307/30036540

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, 55, 90–112. https://doi.org/10.1111/bjet.13370

Zellner, A. (2023). Asking the Right Questions: The meaning of teaching and learning in the age of generative AI. *Irish Learning Technology Association*, 7(2), 52-60. https://doi.org/10.22554/ijtel.v7i2.132

ASCILITE 2024

Navigating the Terrain:

Emerging Frontiers in Learning Spaces, Pedagogies, and Technologies

Zirar, A. (2023). Exploring the impact of language models, such as ChatGPT, on student learning and assessment. *Review of Education*, 11, e3433. https://doi.org/10.1002/rev3.3433