Beyond conventional teaching towards networked learning: The role of generative AI chatbots in enhancing program evaluation skills

Soroush Sabbaghan, University of Calgary, Canada, ssabbagh@ucalgary.ca
Barbara Brown, University of Calgary, Canada, babrown@ucalgary.ca

Abstract

With Generative Artificial Intelligence (GenAI) adoption growing, education has seen the emergence of innovative technologies like chatbots. However, little research has examined the impacts of GenAI integration in specialized higher education contexts. This study explored graduate students' experiences using a GenAI chatbot, PEARL, within a graduate-level teacher education course focused on teaching students how to collaboratively conduct program evaluations using practice cases. Four students participated in our study and shared perceptions of interviewing personas with PEARL when evaluating the practice cases. Thematic analysis identified advantages like enhanced efficiency and accessibility, plus limitations regarding authenticity of artificial interactions. Findings emphasized the continued importance of human guidance and peer learning to enrich GenAI-enabled education aligning with principles of networked learning. Students highlighted the need for ethical considerations despite interacting with artificial entities, underscoring nuanced understanding. The significance of collaborative analysis and ongoing iterative improvements also emerged as themes integral to meaningful learning. Although GenAI presents transformational potential in instructional designs, findings support the use of blended approaches that strategically integrate its advantages with human activity and collaborative inquiry. The study makes contributions by elucidating domain-specific nuances of integrating GenAI into teaching in higher education. Practical implications encourage scaffolding GenAI curricula to promote authenticity and collaborative knowledge construction. Further research could examine variations across disciplines, technologies, and demographics. Overall, as GenAI shapes academia's evolution, reflective pedagogical examination will be key to evidence-guided integration. This exploratory study presents a preliminary yet important step, unveiling opportunities for networked learning and complexities of GenAI adoption in teaching program evaluation skills in education contexts.

Keywords
Artificial Intelligence; Higher Education; Generative AI Integration; Chatbots

Introduction

The application of Generative Artificial Intelligence (GenAI) in higher education has become increasingly pervasive in recent years, significantly transforming the manner in which educators and students interact with educational content and delivery mechanisms. Diverse applications of GenAI, including but not limited to adaptive learning systems, intelligent tutoring systems, and natural language processing tools like ChatGPT, have been adopted across various educational settings (Rudolph et al., 2023). The availability of these types of applications is insufficient for transforming education; change requires alignment with the principles of networked learning, fostering collaborative and connected learning experiences that bridge people, ideas, and resources, transforming traditional educational models (Gourlay et al., 2021; Networked Learning Editorial Collective (NLEC), 2021). Networked learning is defined as: "processes of collaborative, co-operative and collective inquiry, knowledge-creation and knowledgeable action, underpinned by trusting relationships, motivated by a sense of shared challenge and enabled by convivial technologies" (NLEC, 2021, p. 319).These powerful technologies have been successfully harnessed to augment teaching and learning experiences, stimulate student engagement, and ultimately elevate educational outcomes (Malik et al., 2023). An exemplary illustration of an impactful natural language processing tool is ChatGPT, which has shown marked potential in higher education domains. It efficiently assists in answering student queries, providing insightful feedback on assignments, and promoting collaborative learning experiences (Atlas, 2023). In essence, educators designing learning focused on principles of networked learning (NLEC, 2021) and using ChatGPT and similar technologies have the capacity to modernize

how students and educators engage with educational content, thereby making the learning process more universally accessible and immersive (Malik et al., 2023).

Data-driven persona development is known to be a field of study that is complex with inherent limitations (Salminen et al., 2021). Despite the limitations, the integration of GenAI-powered persona generating programs in higher education contexts introduces a plethora of possibilities. Not only do these systems augment learning experiences, but they also foster personalized learning processes. By incorporating personal experiences and contextual elements, designing learning with educational technologies can enhance student engagement, understanding, and knowledge retention, as explored in (Atlas, 2023) in the context of a graduate level educational setting.

A key area that calls for GenAI's attention within higher education context is the enhancement of research skills among graduate students. Graduate students often have limited experience with research activities and processes and encounter challenges during their data collection phase for inquiry-oriented projects. These challenges encompass ethical concerns surrounding human participants, limitations associated with participant accessibility, and subsequent restrictions on engaging with diverse participants thus limiting diversity in responses (see Cypress, 2018 for full review). These obstacles can compromise the effectiveness and efficiency of data collection, primarily in qualitative research. Moreover, time constraints for data collection from human participants often hamper the practice of research skill development in graduate school courses. A solution that addresses these issues while promoting students' data collection skills is thus of considerable value.

In response to this need, the Persona Emulating Adaptive Research and Learning bot (PEARL – LINK REMOVED), powered by GPT4 API, emerged as a promising tool to support novice researchers develop research skills. An integral component of PEARL is the inclusion of personas which are meticulously developed by the instructor, embedding them with memories and lived experiences. This unique feature not only reduces the risk of GenAI persona stereotypes, but also allows students to interact with these AI-generated entities to practice qualitative data collection, effectively counteracting the need for ethical approval to practice research-skill development and addressing problems related to participant accessibility for training or educational purposes. By providing a controlled and safe environment for the practice of interviewing methods, PEARL offers an effective way to simulate real-world researcher-participant interactions. This process not only creates an opportunity for students to develop and refine their research skills but also exposes them to practical, low-risk settings. A detailed discussion of PEARL's practical application within a graduate-level program evaluation course in the field of teacher education will be presented in the upcoming sections.

This study focuses on a graduate teacher education course, aimed at developing program evaluation skills. The course emphasizes evaluation as a discipline and profession, focusing on program evaluation rather than individual assessments. It seeks to foster evaluative thinking, understanding of evaluation's nature, and practical skills for becoming proficient in program evaluation. A key aspect is using PEARL, a persona-generating tool, to enhance real-world understanding through simulations and experiential learning. The study's primary goal is to assess PEARL's effectiveness in enriching the learning experience in a program evaluation course, answering the question: "How does PEARL enhance students' learning in program evaluation?"

Amidst growing adoption of GenAI in education, this study fills a gap in existing research by exploring the potential impacts of GenAI-powered chatbots on student learning, especially within the context of a program evaluation course. The goals of this research are articulated as follows:

1   To determine how interactions with GenAI-personas can enrich students' comprehension and practical application of program evaluation concepts and techniques.
2   To examine students' perceptions regarding the authenticity of their interactions with GenAI-personas and the subsequent effects of this perceived authenticity on their learning experiences.
3   To identify potential barriers and facilitators associated with the effective use of GenAI-personas like PEARL within the program evaluation course, based on the students' personal experiences.

Review of literature

Generative AI, particularly in higher education, has seen significant advancements in recent years. García-Peñalvo and Vázquez-Ingelmo (2023) highlight these developments, emphasizing the progress in understanding and defining Generative Artificial Intelligence. This period has witnessed the emergence of new GenAI models that notably incorporate considerations of ethics within the realm of GenAI. This evolution is highlighted by the development of new GenAI models, which now tackle ethical issues related to GenAI. One of the key areas of evolution has been the use of GenAI technologies like ChatGPT in the context of higher education (Farazouli et

al., 2023). Research by Wang et al. (2023) shows that rapidly advancing GenAI capabilities are transforming higher education. This evolution requires learning developers and instructional designers to adapt, ensuring effective integration of GenAI chatbots in educational settings, signifying a shift in educators' and administrators' roles.

GenAI is considered a tool in enhancing learning experiences in a higher education context for several reasons. Firstly, it can improve learning efficiency and provide customized learning experiences (Chan & Hu, 2023). This is particularly relevant in the development of GenAI chatbots, which can enhance learning experiences by providing personalized and interactive learning environments (Bahroun et al., 2023). Secondly, GenAI technologies like ChatGPT can be responsibly and effectively integrated into various teaching and learning contexts, offering a broad spectrum of impacts and opportunities as discussed in Chen (2023). Thirdly, GenAI has the potential to revolutionize education by deciphering the emotional context of written materials, thereby enhancing learning, life achievements, and mental and physical health support for individuals and society at large (Escotet, 2023). Lastly, GenAI can be a powerful tool for advancing higher education institutions by providing useful and valuable examples of how GenAI can be used in teaching and learning contexts (Kamalov et al., 2023). However, it is important to note that the use of GenAI should be considered with caution, ensuring that it is used to enhance, rather than replace the human aspects of learning environments (Brew et al., 2023).

Persona-generating programs are not new in the human-computer interaction literature and have continued to evolve with the development of artificial intelligence applications (Salminen et al., 2021). In educational settings persona-generating programs emerged in response to the intersection of advancing GenAI capabilities and evolving academic requirements. The rise of GenAI, particularly with tools like ChatGPT proficient in simulating human interactions, introduced transformative possibilities into academia. Historically, academic research involving human participation faced hurdles: ethical considerations, participant recruitment, data authenticity, and time constraints (Cypress, 2018). GenAI's evolution promised solutions. Kocaballi's (2023) study epitomized this potential. Here, ChatGPT simulated human roles in a design project, showcasing GenAI's capability to support human-centered activities, albeit with some limitations in output richness and diversity. The creation of persona-generating programs was not just about circumventing traditional research challenges. As academia progressively leaned towards personalized learning, GenAI-powered personas emerged as tools for tailoring education to individual needs. Despite the benefits in using GenAI, there are also criticisms and limitations of persona designs. In a literature review of data-driven persona development from 2005-2020, Salminen et al. (2021) reported challenges with data quality, data availability, weaknesses with the methods used, and bias. Persona-generating programs in education, such as PEARL, stemmed from a blend of technological innovation and a desire to refine teaching methods for the development of research skills. As the educational community continues to integrate these tools and address limitations and the complexity of persona designs, understanding their full implications remains a priority.

Methodology

A qualitative action research approach was adopted for this study to provide a deeper, more holistic understanding of the students' experiences and perceptions regarding the utilization of PEARL in a program evaluation context. Details on the action research phases and instructional design can be found in a separate publication by (AUTHORS). The qualitative method used as part of the action research study, specifically semi-structured interviews, enables the collection of rich, in-depth data that would not be possible through quantitative measures alone. This choice was instrumental in understanding the nuances, emotions, and intricacies of the students' experiences and thus was consistent with the goals of the research. Ethical approval for this study was obtained after applying for an ethics certification, which was reviewed by the University's Research Ethics Board.

Participant selection

The participants for this study were selected from an educational research graduate level course titled "Program Evaluation and Practice" offered at a large Canadian university, where the principal investigator (PI) of this study served as the instructor. The co-investigator presented the research to students during a lecture slot. While the GenAI program was available for all 38 students in the two sections of the course, for the purposes of this research, four graduate students who expressed willingness to be interviewed by the co-investigator volunteered to be participants. The co-investigator collected and anonymized data prior to sharing with the PI-instructor after the completion of the course. This decision for the participants to engage or refrain in the study was entirely voluntary, ensuring no undue influence on their academic grades or status.

Learning activity integration

A new interactive hands-on activity was incorporated into the curriculum using GenAI personas, as imaginary people that were part of the program under evaluation. In this activity, student teams, formed in the third week of the course, were challenged to complete one out of five available incomplete program evaluations as practice cases. These evaluations presented a unique opportunity for the students to utilize an innovative artificial intelligence program, designed to practice data collection. The overarching objective of this activity was to offer students an authentic experience of undergoing a program evaluation process with practice cases, commencing from the data collection phase to the final step of proposing insightful recommendations.

The learning activity was structured as follows:

- Provided Material: At the start, student formed groups were presented with five incomplete program evaluations. Each group had the liberty to select one evaluation for completion. These evaluations encompassed various sections such as the executive summary, introduction, the purpose of the evaluation, posed questions, and the methodology.
- GenAI Interviews: A pivotal aspect of this activity revolved around the use of PEARL, which was adept at simulating personas representing stakeholders in the practice cases. Armed with this tool, each group set about designing interview questions and subsequently engaging in interviews with the GenAI personas. This exercise was instrumental in gathering critical data, essential for their selected practice case and evaluation.
- Data Analysis: Post the GenAI-facilitated interviews, each group reviewed the transcripts which they scrutinized to extricate data that aligned with their evaluation queries.
- Completion of the Report: With the analyzed data in hand, groups then embarked on the final part of this activity - completing the remaining segments of their selected program evaluation.

Data collection and interview process

Once the students had experienced interacting with PEARL and completed their interactions throughout the semester, the students were prompted to reach out to the co-investigator to schedule their interviews if they were interested in participating in the study. Four semi-structured interviews (one representative of each group) were led by the co-investigator and sought to probe deeper into the students' experiences and perceptions.

The interviews focused on three main areas:

1 Understanding the perceived influence of interacting with the GenAI-persona on the students' comprehension of program evaluation methods and concepts.
2 Gaining insights into the students' impressions about the authenticity of the GenAI-persona interactions and how it affected their overall learning experience.
3 Identifying any challenges or facilitators the students encountered while using the GenAI-persona in their coursework.

With the students' explicit consent, all interviews were recorded to ensure accuracy in data capture. The recordings were transcribed verbatim and sent to the participants for review to ensure accuracy. One of the four participants was unable to meet synchronously for an interview and opted to submit written responses to the interview questions. Although there was no opportunity for follow-up questions, the participant's responses were substantive and provided valuable insights, and thus were considered as an integral part of the data collected for the study.

Data analysis

Upon collecting the data, the first step was familiarizing ourselves with the content. We thoroughly reviewed the transcribed interviews multiple times, which aided in comprehending the depth of the responses and fostering initial thematic ideas from the qualitative data (Saldaña, 2021). Following this immersion, we began generating initial codes. Each significant portion of the interviews was systematically coded, encapsulating its primary essence. The approach to coding was twofold: inductive, where codes arose naturally from the data, and deductive, where the process was steered by our research questions. Subsequent to the coding process, we began searching for overarching themes. Data extracts that bore similarity or interrelated ideas were collated into these potential themes. Once collated, it was crucial to review these tentative themes against the backdrop of the coded data extracts and the dataset as a whole. Through this review, themes were honed further, with some being merged, refined, or even separated to ensure they truly mirrored the insights from the data. With a clearer picture of the prominent themes, the next phase was to crystallize them further by defining and naming each. This step ensured that each theme was delineated with precision and clarity, representing its core idea effectively. Concluding the

thematic analysis process, we synthesized the themes into a comprehensive account, making sure they provided an authentic and vivid representation of the data in the context of our research questions.

Findings

In the course of our investigation into students' experiences and perceptions surrounding the utilization of PEARL, several critical insights emerged. Students expressed varied opinions about the advantages and challenges of using GenAI personas as a part of their course work in the academic environment, especially within the framework of program evaluation. By analyzing these responses, three dominant themes were identified which encapsulate the essence of their collective sentiment. In the succeeding sections, we delve deep into each of these themes, underscoring the benefits and potential pitfalls associated with the integration of GenAI into a learning design.

Theme 1: The integration and impact of GenAI in learning

A significant portion of the feedback pivoted around the profound impact GenAI has on learning within a research context. The responses in this category were bifurcated into its advantages and the authenticity-related challenges. Many participants lauded the advantages they observed when interacting with GenAI personas. They felt that this tool expedited the learning process, eliminated cumbersome bureaucratic procedures, and offered a platform to hone their skills through repetitive practice. One student aptly summarized this sentiment, stating, " [Using GenAI] was also easier because finding people might be difficult because of the short term. And then if you had to actually go and look for people, like you are looking for people to participate, it's a lot of work, right? You have to sign consent forms. So, I think it was really wonderful with that experience."

However, the enthusiasm around the potential of GenAI was not unanimous. Some participants pointed out limitations. They felt that while GenAI was innovative, its technology had its confines, particularly when attempting to simulate real-life scenarios. Highlighting this, a student remarked, "I found that there was generally consistency in the AI responses, but it was limited by the number of available tokens. I would have liked to explore with more follow-up questions to see how specific the responses might be, but we maxed out. For this reason, it didn't feel as authentic as it could have." Additionally, another participant mentioned a distracting facet of these GenAI personas: "I felt that the AI personas were detrimental sometimes because instead of focusing on the actual program evaluation process, we were more focused on trying to get the AI personas to behave in the way we wanted them to." This theme underscores the delicate balance between the potential of GenAI to transform educational practices and the need for refinement to make its integration seamless and genuinely enriching.

Theme 2: The importance of collaboration, feedback, and improvement in program evaluation

The salience of collaboration and the ongoing feedback loop in program evaluation were recurring themes in the students' reflections on their experiences with the GenAI personas. These elements were crucial not only in the context of interacting with GenAI personas but also in the broader realm of real-world program evaluation practices.

An essential element that students frequently highlighted was the significance of a collective effort, especially when deciphering data and ensuring accurate coding and evaluation. The GenAI personas presented different lived experiences in the practice cases that were being examined as part of the evaluation. The student group had to then discern themes from four different GenAI personas, discuss, and synthesize them to arrive at a coherent understanding. One participant described how peers worked as a team to interview and discern themes, "So, with the AI personas when we interacted with them, we were able to get different themes. And based on the course, we were able to derive them. We had common themes that were common to 4 of them because we interviewed 4 personas and based on that we were able to get the feedback that you know, matched what they were actually talking about."

The learning journey, replete with its challenges and triumphs, became a focal point for many students. Their reflections underscored the importance of approaching the GenAI as more than just a digital tool but rather as a simulated stakeholder, thereby aiming for an authentic experience. As one student elaborated, "I would say like that mentality, got us to think, okay, we're not going to start right into asking our questions as if it's like an Amazon chatbot... We would want to introduce ourselves, have space for them to introduce themselves, like how they're doing, like a check-in, explain the process to them, like what we're doing, what the purpose is, and before we start the interview, do we have their consent?" This sentiment reinforces the idea of approaching the GenAI personas with the same professionalism and research acumen as one would in a traditional face-to-face interview with a participant.

Further deepening the learning experience was the emphasis on revisitation and iteration. After the initial round of GenAI interviews and thematic identification, peer groups reconvened to discuss any potential gaps, ambiguities, or areas that might benefit from further exploration. "I would say that it helped in that case. And then when we did identify the themes and our group met. We're like, so what does this mean? What information are we missing? And then we were talking about how it is possible in the real world to have a secondary interview with participants, to seek further clarification or elaboration of questions. Could we go back to the participants, to ask these questions?" This feedback loop, and the possibility of revisiting the data source (in this case, the GenAI personas interview transcripts), reflected the students' maturation in the process and a deeper appreciation of the intricacies of program evaluation.

Theme 3: Navigating the realities of program evaluation with GenAI integration

Delving deeper into the students' experiences brought forth a nuanced understanding of the role of GenAI in program evaluation and the overarching challenges and intricacies associated with the process.

A recurring and pivotal topic was the ethical considerations in integrating GenAI into the research-skill development process. The reflections showed an earnest commitment to ensuring that research, even with GenAI personas, upheld ethical standards. A student recounted their experience: "For our individual, that's where, the questions that we had created, like the general broader questions, to give us a similar set approach, to the 4 different personas…Then the testing, like the demographic questions, the program related questions, we did that on Zoom just to assess, will they answer these questions? Will the personas be able to give us more information, beyond the experiences of their program? So, we did that together on Zoom and we're like, oh my goodness, yes, they answer the question." This quote shows how the group intentionally tested questions with the GenAI personas to see if they would answer sensitive demographic questions, and to practice treating them with the same ethical standards as human subjects. This reflection underscores the careful thought and rigorous approach that the students adopted, ensuring that the GenAI personas were not reduced to mere tools but were interacted with ethically even in a practice environment.

Central to the program evaluation process is the acquisition of practical skills. Students reflected on the significance of real-world application and understanding the practical implications of the tools and techniques employed. As one participant aptly pointed out, "The nature of [this GenAI-powered task] provided insight as to how to conduct an actual program evaluation, and thus the AI tech helped. I was able to see how to code data in a semi-structured interview and the importance of working with a team to analyze that data." This sentiment emphasizes the hands-on experience gained through the GenAI-based evaluation, reinforcing the importance of practical exposure in academic settings.

Another crucial facet was the need for comprehensive information to ensure a thorough evaluation. Students expressed a desire for a more detailed scope to aid them in framing their questions better. One of them mentioned, "I didn't like that I didn't have more information about the scope of the program so that I could then be more intentional in posing questions. There were also some questions we asked off the record." Such reflections shed light on the nuances of information gathering, and the challenges posed when certain pieces of the puzzle are missing. It underscores the importance of equipping researchers with a full understanding to facilitate a meaningful and thorough evaluation.

Discussion

This study explored students' experiences and perceptions surrounding the use of a GenAI chatbot, PEARL, within a graduate-level program evaluation course in teacher education. The key objectives were to determine how simulated GenAI-persona interactions enriched comprehension and application of concepts, examine perceptions of authenticity and effects on learning, and identify challenges and facilitators with integrating GenAI. Through semi-structured interviews with students and thematic analysis, several crucial insights emerged centered around three dominant themes: the integration and impact of GenAI on learning, the significance of collaboration and feedback loops in evaluation, and navigating the intricacies of evaluation with GenAI. This discussion will provide a reflective examination of these central themes in light of the research goals.

Integration and impact of GenAI in learning

The integration of GenAI through the PEARL chatbot introduced noteworthy advantages but also some limitations in terms of authenticity. Students widely acknowledged the benefits of GenAI in streamlining processes, overcoming recruitment challenges for practice purposes, and enabling repeated practice. This aligns with

literature emphasizing improved efficiency and accessibility with GenAI (Bahroun et al., 2023). However, despite the positives, some students felt interactions lacked authenticity due to technological constraints. The absence of true dynamic dialogue and audio limited perceived realism. This substantiates existing critiques of GenAI's inability to fully capture nuanced human interactions (Kocaballi, 2023). Nonetheless, students recognized PEARL's innovation in providing low-risk opportunities to practice and hone skills. This carries valuable implications for designing GenAI curricula and learning designs that strategically target growth areas while acknowledging current technological bounds.

The importance of collaboration, feedback, and improvement in program evaluation

A prevailing theme was the significance of collaborative analysis, feedback loops, and iterative improvements for meaningful learning. The need to synthesize diverse GenAI outputs reinforced the need for professional collaboration and the role of teams in deciphering complex findings, aligning with literature on collective knowledge construction (Chan, 2023) and the principles of networked learning, such as the importance of human relationships and how technologies are used within a collaborative activity (NLEC, 2021). Furthermore, the openness to revisit GenAI data mirrors the real-world process of clarifying ambiguities through follow-up inquiries. This reflective practice and desire for elaboration highlights the students' deepening insight into the responsive nature of evaluation. For educators, this advocates for integrating collaborative GenAI-powered activities with channels for ongoing peer and instructor feedback.

Navigating the realities of program evaluation with GenAI integration

Despite GenAI's advantages, students highlighted the intricacies in navigating evaluation, from ethics to information needs. The commitment to preserving ethical standards despite interacting with artificial entities in a practice setting reveals a nuanced understanding of researchers' duties. This finding, coupled with the want for comprehensive background data, demonstrates an appreciation of the complexities inherent in program evaluations. It further shows that while GenAI facilitates skill-building, it cannot wholly substitute core domain knowledge. For effective integration, educators must scaffold GenAI with foundational concepts.

Comparison with previous literature

Several key findings aligned with existing literature on GenAI in education, while also yielding some novel insights. The advantages of efficiency, accessibility and repetitive practice confirm GenAI's learning benefits noted across studies (Chen, 2023). However, the authenticity critiques enrich the discourse on GenAI's limitations in fully capturing human complexity (Kocaballi, 2023). This tension between GenAI's assets and constraints mirrors the ongoing debate around replacement versus enhancement framings in academia (Brew et al., 2023). Furthermore, the emphasis on collaboration provides empirical support for GenAI's role in collective knowledge building, as suggested by Chan (2023) and principles of networked learning (NLEC, 2021). Unique to this study was the focus on program evaluation contexts. Students' reflections revealed GenAI's impacts on gaining practical research skills, which affirms claims of augmented experiential learning (Bahroun et al., 2023). At the same time, desires for more information beyond GenAI highlight the importance of foundational domain knowledge, a novel insight. Overall, while the findings validate GenAI's educational promise, they also showcase some domain-specific nuances in evaluation training for novice researchers.

Implications for educators and researchers

Several meaningful implications emerge from this study for educators and researchers interested in integrating GenAI and chatbots into program evaluation or similar higher education contexts. Firstly, the findings showcase the need for a blended approach that thoughtfully combines GenAI with traditional pedagogies. Standalone GenAI risks authenticity perceptions; whereas, situating GenAI in a networked learning frame and combining it with peer collaboration, instructor guidance and domain fundamentals could mitigate this. Secondly, the reflections reveal that GenAI training should be scaffolded, starting with lower-risk repetitive skills then progressing to more ambiguous real-world applications. Lastly, creating channels for ongoing improvement through feedback loops with GenAI could enrich learning. Researchers must also continue examining ethics, authenticity, and content needs within specific disciplines as GenAI capabilities evolve.

Limitations of the study

While providing valuable insights, this study had some limitations. The small sample size of four students from one program evaluation course restricts generalizability. Additionally, as the study focused narrowly on chatbot experiences, other forms of GenAI integration were not examined. Participant views may also have been subject

to biases such as novelty effects of a new technology and inexperience in research activities. Further research with larger, more diverse samples across various GenAI applications could build on these findings.

Conclusion

This study delved into graduate students' experiences with the GenAI chatbot, PEARL, in a program evaluation course, employing semi-structured interviews for a nuanced understanding of GenAI's role in education. The research revealed that learning designs with GenAI technologies like PEARL enhance learning efficiency, accessibility, and skill practice, yet struggle with the authenticity of interactions. A key theme was the balance between GenAI's benefits and its limitations. Importantly, the study underscored the irreplaceable role of human elements in education, such as peer learning and instructor guidance, alongside GenAI tools aligning with principles of networked learning. It noted that while GenAI offers significant potential, it cannot fully substitute for real-life informants or the complex dynamics of human interaction, indicating a need for a carefully balanced educational approach that melds GenAI with human activity and collaborative inquiry. The research also provided practical insights, advocating for blended GenAI curricula that simulate real research conditions, thereby enriching the learning experience. For future research, it emphasized the importance of replicating similar studies across various disciplines and GenAI applications to broaden understanding. Conclusively, as GenAI reshapes learning designs in academia, this study marks an essential step in evaluating its integration into educational contexts, highlighting the importance of reflective, evidence-based pedagogical research in this evolving landscape.

References

Atlas, S. (2023). *ChatGPT for higher education and professional development: A guide to conversational AI*. University of Rhode Island. https://digitalcommons.uri.edu/cba_facpubs/548/

Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming education: A comprehensive review of generative artificial intelligence in educational settings through bibliometric and content analysis. *Sustainability*, *15*(17). https://doi.org/10.3390/su151712983

Brew, M., Taylor, S., Lam, R., Havemann, L., & Nerantzi, C. (2023). Towards developing AI literacy: Three student provocations on AI in higher education. *Asian Journal of Distance Education*. https://www.asianjde.com/ojs/index.php/AsianJDE/article/view/726

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *arXiv Preprint arXiv:2305.00290*.

Chen, T.-J. (2023). ChatGPT and other artificial intelligence applications speed up scientific writing. *Journal of the Chinese Medical Association*, *86*(4), 351−353.

Cypress, B. (2018). Qualitative Research Methods: A Phenomenological Focus. *Dimensions of Critical Care Nursing*, *37*(6), 302. https://doi.org/10.1097/DCC.0000000000000322

Escotet, M. Á. (2023). The optimistic future of Artificial Intelligence in higher education. *PROSPECTS*. https://doi.org/10.1007/s11125-023-09642-z

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

García-Peñalvo, F., & Vázquez-Ingelmo, A. (2023). What do we mean by GenAI? A systematic mapping of the evolution, trends, and techniques involved in generative AI. *International Journal of Interactive Multimedia and Artificial Intelligence*, *7*(4), 7−16. https://doi.org/10.9781/ijimai.2023.07.006

Gourlay, L., Rodríguez-Illera, J. L., Barberà, E., Bali, M., Gachago, D., Pallitt, N., Jones, C., Bayne, S., Hansen, S. B., Hrastinski, S., Jaldemark, J., Themelis, C., Pischetola, M., Dirckinck-Holmfeld, L., Matthews, A., Gulson, K. N., Lee, K., Bligh, B., Thibaut, P., … Networked Learning Editorial Collective (NLEC). (2021). Networked Learning in 2021: A Community Definition. *Postdigital Science and Education*, *3*(2), 326−369. https://doi.org/10.1007/s42438-021-00222-y

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability*, *15*(16). https://doi.org/10.3390/su151612451

Kocaballi, A. B. (2023). *Conversational AI-powered design: ChatGPT as designer, user, and product*. arXiv Preprint.

Malik, A., Khan, M. L., & Hussain, K. (2023). *How is ChatGPT transforming academia? Examining its impact on teaching, research, assessment, and learning.* (SSRN Scholarly Paper 4413516). https://doi.org/10.2139/ssrn.4413516

Networked Learning Editorial Collective (NLEC). (2021). Networked Learning: Inviting Redefinition. *Postdigital Science and Education*, 3(2), 312−325. https://doi.org/10.1007/s42438-020-00167-8

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342−363. https://doi.org/10.37074/jalt.2023.6.1.9

Saldaña, J. (2021). *The coding manual for qualitative researchers*. Sage Publications.

Salminen, J., Guan, K., Jung, S.-G., & Jansen, B. J. (2021). A Survey of 15 Years of Data-Driven Persona Development. *International Journal of Human−Computer Interaction*, 37(18), 1685−1708. https://doi.org/10.1080/10447318.2021.1908670

Wang, T., Lund, B. D., Marengo, A., Pagano, A., Mannuru, N. R., Teel, Z. A., & Pange, J. (2023). Exploring the potential impact of artificial intelligence (AI) on international students in higher education: Generative AI, chatbots, analytics, and international student success. *Applied Sciences*, 13(11). https://doi.org/10.3390/app13116716