JUTLP
Journal of University Teaching & Learning Practice

Enhancing Methodological Integrity with GenAI: A Multi-case Study of Experiential Learning using Sequential Augmented Analysis

Manuel Etesse[a], Alexandra Shimabukuro[a], Piero Beretta[b] & Jorge Li[a]
[a] Pontificia Universidad Católica del Perú; [b] Grupo de Análisis para el Desarrollo

Editors

Section: Special Issues
Editor-in-Chief: Dr Joseph Crawford
Senior Editor: Dr Cassandra Colvin

Publication

Submission: 21 June 2025
Revised: 3 March 2026
Accepted: 3 May 2026
Online First: 9 June 2026

Abstract

The rapid expansion of Generative Artificial Intelligence (GenAI) in higher education presents a critical pedagogical challenge for research training: how to integrate these tools without undermining methodological integrity. In qualitative research, unstructured GenAI use may encourage overreliance, superficiality, and unreflexive judgment among novice researchers. Despite growing debate, limited empirical evidence shows how GenAI can be deliberately designed to strengthen rigor in undergraduate qualitative data analysis. This study proposes and analyze the Sequencial Augmented Analysis a structured instructional model that embeds the use of chatbots within Human-Centered AI in Education and Experiential Learning frameworks. Using an exploratory multiple-case design with final-year pre-service teachers, we examined how GenAI-enhanced investigator triangulation and guided reflexivity support methodological integrity. Findings indicate that introducing chatbots after manual analysis stimulated collective reconsideration of decisions, systematic returns to original data, and clearer justification of methodological choices. It also surfaced personal biases, methodological assumptions, and ethical concerns regarding authorship and disclosure. Rather than replacing human judgment, GenAI functioned as a catalyst for dialogue and critique. The study offers a replicable pedagogical design for integrating GenAI into qualitative research courses while reinforcing academic integrity.

Practitioner Notes

1. Embedding GenAI in qualitative research training should prioritize triangulation and reflexivity to reinforce methodological integrity.
2. Educators can design sequential activities combining manual analysis, teamwork, and AI-supported comparison to strengthen learning.
3. Facilitators should position GenAI as a complementary mediator rather than a substitute, ensuring oversight and transparency.
4. Using authentic datasets in experiential learning helps students confront ethical issues such as authorship, disclosure, and accountability.

Keywords

qualitative data analysis, pre-service teacher education, ChatGPT, undergraduate research training

Citation: Etesse, M., Shimabukuro, A., Beretta, P. & Li, J. (2026). Enhancing Methodological Integrity with GenAI: A Multi-case Study of Experiential Learning using Sequential Augmented Analysis. *Journal of University Teaching and Learning Practice, 21*(6). https://doi.org/10.53761/tmq0ks13

Introduction

Generative Artificial Intelligence (GenAI) is rapidly reshaping higher education, influencing how students complete academic tasks like writing and analyzing data (Xia et al., 2024; Alier et al., 2024). In response, universities are debating not only how to regulate these tools but also how to integrate them without compromising academic integrity. Nowhere is this tension more visible than in qualitative research methodology courses, where rigor depends on interpretive depth, transparency, and reflexive judgment (Lincoln & Guba, 1985; Eisenhart & Jurow, 2011).

While GenAI tools such as ChatGPT can assist with several procedures in qualitative data analysis (QDA), their unstructured use risks superficial analysis, overreliance, and uncritical acceptance of automatic outputs (Salvagno et al., 2023; Nguyen et al., 2023). These risks are particularly acute among undergraduate students conducting research projects for the first time (Wagner et al., 2019; Talbott & Lee, 2020). The central problem, therefore, is not whether GenAI should be used in research methodology courses but how it can be intentionally designed to strengthen—rather than weaken—methodological integrity. With appropriate pedagogical scaffolding, it may act as a catalyst for deeper analysis, collective deliberation, and explicit justification of methodological decisions.

This study addresses that gap by embedding GenAI within a Human-Centered AI in Education (HCAIE) framework (Shneiderman, 2020; Yang et al., 2021), operationalized through Experiential Learning (Kolb & Kolb, 2005) and anchored in two established qualitative reliability practices: investigator triangulation and reflexivity (Creswell & Miller, 2000; Levitt et al., 2017). The purpose of this study is to examine how GenAI-enhanced investigator triangulation and reflexivity practices foster methodological integrity in undergraduate teacher education in Lima (Peru). In authentic research settings and across two qualitative projects and their respective datasets last year, students followed structured, team-based Sequential Augmented Analysis (SAA) (Etesse et al., 2026). Rather than positioning ChatGPT as an analytic authority, in SAA it has the role of a mediator introduced after manual analysis to foster comparison, verification, and collective review.

The study addresses a general research question: How does GenAI-enhanced investigator triangulation and reflexivity practices within experiential learning help students strengthen the reliability of findings and contribute to methodological integrity? The article first reviews relevant literature on GenAI in higher education and outlines the conceptual framework integrating HCAIE, experiential learning, and methodological integrity. It then describes the exploratory multiple-case design and learning intervention. The Results section presents findings across two dimensions—investigator triangulation and reflexivity—drawing on focus groups, reflexive writings, and survey data. The Discussion situates these findings within current debates on GenAI and academic integrity in higher education. The paper concludes with implications for teaching research practices and for responsible GenAI integration in university settings. The purpose of this study is to examine how GenAI-enhanced investigator triangulation and guided reflexivity practices foster MI within teamwork EL settings framed by HCAI in Education. Accordingly, the study addresses two specific questions:

> **RQ1.** How does GenAI-enhanced investigator triangulation within experiential learning help students strengthen the reliability of findings and contribute to MI?
>
> **RQ2.** In what ways do GenAI-enhanced reflexivity within experiential learning enhance students' critical awareness of assumptions and contribute to MI?

This study makes three contributions to research on GenAI integration in higher education methodology training. First, it provides empirical evidence on how GenAI-enhanced triangulation and reflexivity strengthen methodological integrity in qualitative research instruction—an area underexplored in existing literature. Second, it offers a replicable pedagogical model grounded in Human-Centered AI in Education and Experiential Learning frameworks, moving beyond generic GenAI adoption toward principled instructional design that aligns GenAI use with established qualitative reliability practices. Third, and most critically, it reframes GenAI not as a threat to academic integrity but as a structured catalyst for triangulation and reflexivity; demonstrating empirically that, when embedded in SAA, a structured pedagogical design, GenAI can deepen methodological integrity. Together, these contributions offer actionable and theoretically guidance for integrating chatbots into qualitative research training.

Literature

GenAI in Higher Education: Tensions and Evidences

Recent literature flags substantial ethical concerns about GenAI in higher education—plagiarism, cheating, misleading/outdated content, overreliance, bias, privacy/security issues, and role ambiguity—requiring clear policies (Nguyen et al., 2023; Xia et al., 2024; Tlili et al., 2023; Chan & Lee, 2023; Kusner & Loftus, 2020; Sivill, 2019; Nguyen, 2023; Guilherme, 2019). Yet intentional integration yields gain in writing quality, higher-order thinking, creativity, comprehension, productivity, and metacognition—especially via instructor–ChatGPT feedback cycles and critique of AI outputs (Liu et al., 2024; Michalak & Ellixson, 2025; Exintaris et al., 2023; Zhao et al., 2023; Murillo-Ligorred et al., 2023).

GenAI broadens creative spaces by generating novel ideas and fostering fresh concepts through iterative analysis and reflection (Zhao et al., 2023; Isiaku et al., 2024). Studies also report comprehension gains and reduced cognitive load as GenAI decomposes complex information and answers advanced queries—supporting autonomy, lowering anxiety, and improving preparedness (Diaz et al., 2025; Annamalai & Nasor, 2025). Students note higher productivity, personalized learning, and enhanced linguistic capability, though they still call for guidance (Zhou et al., 2024; Isiaku et al., 2024). Evidence further shows stronger research-related learning, from tutorials to methodological development in STEM, improving understanding and engagement (Kirwan, 2023; Kong et al., 2023).

Teaching Research with GenAI

The benefits from scaffolding pedagogical procedures with GenAI can be useful for research training in higher education. First, GenAI enables deeper error analysis and verification through comprehensive reviews, improving revisions and clarity (Liu et al., 2024; Kuramitsu et al., 2023; Uddin et al., 2023, 2024). Second, within pedagogical frameworks, GenAI augments—not replaces—human judgment: students shift from passive use to critical appraisal, adopting bias-detection and validation strategies and reframing AI literacy as multidimensional (creativity, ethics, and critical reasoning) (Park & Doo, 2024; Kumar et al., 2024; Rana et al., 2025). Third, GenAI accelerates iterative ideation and refinement; curricular integrations requiring explicit engagement are effective and popular (French et al., 2023). Fourth, AI boosts efficiency (Wang et al., 2024; Arosio, 2025) as, for instance, prompt engineering reduces routine questions and enhances retrieval and insights.

Empirical studies on teaching research with GenAI remain scarce, yet evidence from practice shows transferable benefits for qualitative data analysis (QDA). GenAI streamlines workflows while preserving human centrality, with heuristic gains during transcript exploration as GenAI generates provisional outputs that complement immersion (Lopezosa & Codina, 2023; Etesse, 2025a, 2025b). GenAI facilitates work with extensive qualitative data—interviews, open replies, and social media—supporting inductive coding (Gao et al., 2023; Qiao et al., 2024), thematic analysis (Perkins & Roel., 2024; Goyanes et al., 2025), and content analysis (Bjiker et al., 2024). Workflows with GenAI assistance can yield fresh insights and alternative perspectives, enriching discovery and fostering reflexivity in QDA; close comparisons between manual and ChatGPT outputs—verbatim selection and coding—surface biases and interpretive tendencies, enhancing reflexive practice (Wachinger et al., 2025; Etesse et al., 2026).

Conceptual framework

Our conceptual framework integrates three strands: Human-Centered AI in Education (HCAIE) as the normative and design stance, Experiential Learning (EL) as the pedagogical vehicle for situated, team-based practice, and Methodological Integrity (MI) operationalized through investigator triangulation and reflexivity in QDA.

Human-Centered AI in Education (HCAIE)

HCAIE adapts the broader paradigm (Shneiderman, 2020)—which places people, ethics, and human values at the center of human-machine collaboration—to educational design and practice (Yang et al., 2021). It functions both as a normative stance and as a design commitment to the sustained, responsible use of AI while ensuring human control (Li & Gu, 2023; Díaz & Nussbaum, 2024). Within this framework, AI may automate routine tasks but does not replace the educational judgment of teachers and students (Xu et al., 2021). This orientation implies ethically aligned design, technology enhancement, and human-factor considerations to guarantee explainable and comprehensible AI solutions; transparency and interpretability are therefore essential conditions for trust and effective workflows in teaching and learning (Zhang et al., 2025).

Within HCAIE, a critical concern is the systematic identification and management of risks. Li and Gu (2023) propose an eight-dimension framework for educational processes. This framework underscores that adopting GenAI is not merely a technical innovation but an educationally situated challenge requiring safeguards that are both ethical and pedagogical.

Experiential Learning (EL)

EL frames learning as a continuous, situated cycle through which students construct meaning in authentic contexts (Kolb & Kolb, 2005). Morris (2019) refined the cycle for higher education by emphasizing context-rich experience, critical reflection, context-specific conceptualization, and pragmatic experimentation. Our approach is inherently collaborative: teamwork around shared problems and peer challenge deepens reflection, sustains motivation, and strengthens accountability for quality (Healey & Jenkins, 2000). In line with students-as-partners, collaboration positions learners as co-analysts (Cook-Sather et al., 2014; Salinas-Navarro et al., 2024).

The collaborative orientation is particularly suited to the teaching of QDA, where students must negotiate consensual criteria, resolve discrepancies, and justify decisions as a team, thereby enacting the very practices of methodological integrity in research teamwork. Experiential structures ensure that such collective inquiry is not abstract but embedded in authentic, context-rich tasks, allowing learners to build transferable analytic skills while cultivating shared responsibility for rigor and transparency (Teixeira-Poit et al., 2011).

Methodological integrity (MI)

MI, a dimension of academic integrity, denotes fidelity to the phenomenon, aligning procedures with epistemic commitments and aims; reliability criteria serve as benchmarks for assessing MI in QDA (Levitt et al., 2017; Creswell & Creswell, 2017). Two anchors—reflexivity and investigator triangulation—support reliability (Creswell & Miller, 2000). Investigator triangulation—one of four types—involves multiple investigators in collection, analysis, or interpretation (Denzin, 2008). Diverse perspectives mitigate single-viewpoint limits, deepen insight, and speed detection/correction of inconsistencies (Campbell et al., 2020). It serves three aims—convergence, complementarity, and productive divergence—enhancing interpretive depth and reliability (Bans & Tiimub, 2021; Stamenkov, 2022).

Reflexivity, a cornerstone of QDA, is the ongoing examination of how researchers' subjectivity and context—across methodological and institutional dimensions—shape exploration and interpretation, including assumptions, positionalities, and bias (Macbeth, 2001; Finlay, 2002; Russell & Kelly, 2002). Rather than weakening inquiry, reflexivity strengthens transparency and rigor, enabling judgments and ethical decisions; practices include reflective memos and team discussions (Mauthner & Doucet, 2003; Berger, 2015; Olmos-Vega et al., 2023).

Building on these principles, QDA sequences emphasize engagement and systematic reduction (Morse, 1994). In phenomenology, this begins with selecting key verbatims, then horizons comparison and thematization (Moustakas, 2010). The sequence is effective in education, entailing case-by-case reading before reducing it to core themes (Morrow et al., 2023; Bonyadi, 2023; Tavakol & Sandars, 2025; Eddles-Hirsch, 2015).

Method

Design

This study adopted an exploratory multiple-case study as a qualitative design (Kass et al., 2024). Drawing on two educational research projects—each with unique aims, questions, and anonymized datasets—the cases enabled authentic EL, as students in small teams applied an AI-assisted workflow during QDA sessions. This dual-case design allowed for an in-depth exploration of how and why the phenomenon unfolded, yielding more robust insights than a single-case study (Creswell, 2007). The use of a small, comparable set of cases was further justified by its potential to produce context-specific conclusions (Lieberson, 2000). In line with Yin's (2018) replication logic, the research team systematically reproduced core procedures across both cases, allowing for the identification of recurring patterns and their relevance to each project's goals and evidence.

Participants

Participants were grouped in two case studies. In total, 18 pre-service teachers from preschool, primary, and secondary Education took part. All were in their final year, though at different stages of undergraduate research. Participants had formal training in basic QDA, ensuring a shared methodological foundation. The two cases were conducted in parallel by different facilitators, in separate settings, and developed independently.

Table 1

Demographic characteristics of participants

		n	%
Total		18	100
Sex	Female	14	77.8
	Male	4	22.2
Age	20–24	12	66.7
	25–29	4	22.2
	30+	2	11.1
Program of study	Elementary ed.	10	55.6
	Preschool ed.	5	27.8
	High-school ed.	3	16.7
First year in university	2021	5	27.8
	2020	5	27.8
	2019	3	16.7
	2018	3	16.7
	2016 or before	2	6.2

Case A. Research focus: educational trajectories and disability. The project examined the primary-school experiences of six adults with disabilities (three mobility, three visual; gender balanced; aged 35–55) from Cusco, Puno, and Lima. Data source: Retrospective narrative-biographical telephone interviews on salient experiences. Student participants: Eight undergraduates (7 women, 1 man). Most majored in primary education (n=7), with one in secondary. Six were writing theses, one doing fieldwork, and one was preparing to start. Learning setting: Conducted under the EL session format.

Case B. Research focus: Teachers' perceptions of standardized testing. The project explored views of six primary-school teachers in Lima (three public, three private; five women, one man; aged 40–50). Data source: In-person semi-structured interviews guided by a thematic protocol on classroom practices, norms, and culture. Student participants: Ten undergraduates (7 women, 3 men). Representation was highest in preschool (n=5), followed by primary (n=3) and secondary (n=2). Eight were writing theses, and two were preparing to begin. Learning setting: Conducted under the EL session format.

Procedures

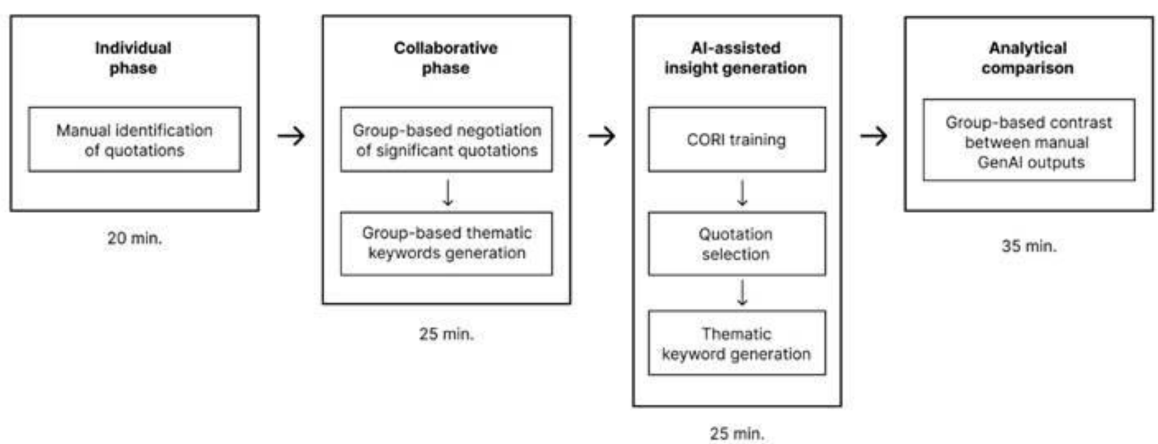
Main procedures were developed in EL sessions. The immersive environment—PUCP research facilities rarely accessed by students—was designed to foreground EL's authentic atmosphere

(Yardley et al., 2012). To deepen immersion, participants were called analysts and organized into teams, strengthening identification with research roles and goals. Each session formed two teams to analyze an anonymized transcript via a structured analytic pathway using ChatGPT-4 within a case-by-case phenomenological approach. Teams worked with minimal facilitator intervention following a brief introduction to the research project and dataset. They also received training in CORI-f prompting—a structured approach that configures chatbots for qualitative data analysis by specifying Context, Objective, Role, and Information relevant to the research task (Etesse, 2024).

The experience was positioned at the “augmentation” level in Puentedura’s SAMR framework, where technology enhances analysis without changing its nature (Hamilton et al., 2016). Also, following Xu and Ouyang (2022), ChatGPT acted as a direct mediator after a manual, individual, and collective analysis is done, enabling interaction and optimized resource use and fostering peer-to-peer exchanges and self-directed engagement. Sessions followed a structured workflow called Sequential Augmented Analysis (Etesse et al., 2026), with four main phases (see Figure 1): (1) Individual analysis (20 min) to examine a transcript extract and identify meaningful quotations; (2) Collaborative analysis (25 min) to choose three significant excerpts and agree on two thematic keywords for each; (3) AI-assisted insight generation (25 min) repeating the process with ChatGPT— quotation selection and thematic keyword generation; and (4) Analytical comparison (35 min) to contrast manual and GenAI outputs. In the las stance, guiding questions-oriented reflection on similarities, differences, and concrete improvements to human-generated analysis—first for quotation selection, then for thematization.

Figure 1

Session workflow for Sequential Augmented Analysis (SAA)



Data

Three data sources informed this multi-case study, fostering concrete data triangulation (Yin, 2018). First, using the activity guide, each team produced reflexive writings and results for each collective work phase, later systematized and incorporated into the dataset. Second, after the EL sessions, focus groups were conducted with each team to discuss overall experiences, group dynamics, and methodological choices for a) quotation selection and b) keyword generation. Four focus groups (one per team) were recorded, transcribed, and line-edited from TurboScribe's automatic transcripts by re-listening to recordings. Third, participants completed a nine-item questionnaire on their experiences, focusing on participation, perceptions of GenAI, and its role in academic research, using five-point Likert-scale responses. All three anonymized datasets and corresponding instruments to collect data are publicly available in the PUCP repository (see Supplementary material).

Analysis

According to the exploratory focus of the research design, thematic analysis was applied (Braun & Clarke, 2006), offering a flexible yet rigorous framework to identify and report patterns. The process began with familiarization through transcript readings and reflexive writings, followed by systematic coding. Codes were consolidated into broader categories, forming the basis for theme development. To enhance transparency, raw verbatims were structured into case-specific matrices with focus groups in columns and thematic dimensions in rows (Averill, 2002), enabling systematic cross-case comparison. Additionally, the protocol of Goyanes et al. (2025) for AI-assisted thematic analysis was used to contrast and confirm findings, identifying researcher biases and reinforcing reliability. The analysis addressed the research questions on investigator triangulation and reflexivity. For each, the iterative process yielded three salient themes. Teams produced analytic memos, examined discrepant cases, and held peer-debriefing sessions with the four authors (see Supplementary material).

Ethical standards

The study was approved by the Dean's Office of the Faculty of Education within the ethical framework of the PUCP Research Ethics Committee (PUCP, 2016) and in alignment with the Declaration of Helsinki (WMA, 2013). Students received a detailed explanation of the project before participation, with opportunities to clarify questions. Engagement was voluntary, with handwritten informed consent obtained and the right to withdraw at any time guaranteed. Confidentiality was assured through anonymized and securely stored data. Protocols ensured equal treatment, while the research team emphasized benevolence and respect to foster a supportive environment.

Results

This section is organized into two parts addressing the study's central dimensions: investigator triangulation and reflexivity. The study identified nine key themes (see Table 2). All were found across cases, though themes 1.2 and 2.2 showed greater salience in cases A and B, respectively, due to data characteristics.

Table 2

Dimensions, themes and presence in cases

Dimensions	Themes	Case
1. Investigator triangulation	1.1. Collective methodological decisions prompted with GenAI 1.2. Thematization collective revision induced by GenAI 1.3. Original data revisit during participants' exchanges	1.1. Both 1.2. Both, more salient in A 1.3. Both
2. Reflexivity	2.1. Identification of personal bias enhanced with GenAI 2.2. Reflexivity on individual methodological decisions 2.3. Ethical interrogations made explicit and formal	2.1. Both 2.2. Both, more salient in B 2.3. Both

Collective methodological decisions prompted with GenAI

ChatGPT provided input that helped reopen discussions among participants regarding previously made methodological decisions. For example, ChatGPT facilitated an alternative way of responding to the task, which led to collective reflection and an adjustment of criteria. A participant reflection was: "So, I think it helped us, in a way, to have a much broader view, but it was also somewhat debatable whether we should go with what Chat suggested or stick with what we had done manually (CaseA- A2)." Another representative statement is:

> "I liked it. Because, for example, you can discuss everything with a group, but it's different when you see what you did compared to something a machine gives you, right? I mean, it gives you a compilation of everything aligned with the main objective. Then we agreed on what we considered relevant—looking at what we selected plus what ChatGPT suggested—and seeing what elements we didn't include in our group work that ChatGPT did highlight." (CaseB- B1)

The survey data further reinforced these qualitative insights, showing that the collective dimension of decision-making was significantly strengthened by the interplay with GenAI. A large majority of participants valued highly the exchange of perspectives with peers when analysing excerpts, with more than three-quarters situating its usefulness at the upper end of the scale (Figure 2). Similarly, nearly two-thirds of respondents acknowledged a strong influence of ChatGPT's outputs on the group's methodological decisions (Figure 3).

Thematisation collective revision induced by GenAI

After obtaining the materials generated by ChatGPT, participants agreed to conduct a deeper group review of the themes initially proposed for each meaningful quote. ChatGPT's construction of thematic keywords yielded more specific terms, prompting participants to analyse and reach consensus on the desired granularity of keywords—that is, how general or specific the terms should be. One participant stated: "For example… We had included the keyword 'autonomy,' and it [ChatGPT] had proposed 'active autonomy.' This narrows the range much more" (CaseB-B1). Another participant added:

> "About some words that we perhaps had in mind but did not distil into a single term, ChatGPT helped us with that. But I also feel that… there were moments like in the third quotation, where it gave us a term that was more global and, in the

> second, a more specific one that was contained within the third… It seemed like it was repeating itself, but it was actually more specific. Another case for example is where ChatGPT used the keyword 'inclusion,' but for us it was not 'inclusion,' it was simply 'integration,' because there was nothing else, at least in the text, that showed the teacher was practicing inclusion" (CaseA-A1)

It is worth noting that this collective revision of thematization was more marked among the two teams in case A. This difference can be plausibly linked to the nature of the data they analysed: less structured, biographical narratives that were also more distant from their own experience, which required greater effort in negotiating meaning and refining thematic precision.

Additionally, after receiving ChatGPT's outputs, participants debated their approach to choosing thematic keywords. It was identified that the manual work leaned more toward description, whereas the work with ChatGPT was more conceptual, that is, deeper and more explicitly grounded in the responses. A participant commented:

> "We read and rely only on what we understand in the first or second reading. In contrast, ChatGPT even provides a rationale for why it should be that way or how the research benefits from that perspective. In other words, it offers support that helps you understand why yes and why no. We could even have added (as a prompt): 'OK, based on the case and our quotations, what rationale could you help us find to use it as a keyword?' " (CaseA- A2)

Another participant added:

> "I think ChatGPT did help us. Above all, because of what I mentioned about having a broader view, knowing that, perhaps, we can not only choose short extracts but also somewhat longer paragraphs (from the text) that provide a more precise perspective and better evidence for the keyword or code being used." (CaseA- A1)

Collectively revisiting original data

The study's results show that materials generated by ChatGPT prompted a collective review of the original data to identify information and topics GenAI deemed relevant for selecting quotations and thematic keywords, contrasting with participants' materials. This process strengthened methodological rigor and encouraged deeper examination of textual information through discussion and read-aloud sessions. On this point, one participant stated: "ChatGPT proposed aspects we had not considered. For example, the suggestion of 'exclusion' as a thematic keyword, which I think would help us review the text again and see in what other parts that is also discussed." (CaseA- A1)

Likewise, reviewing the original data enabled new group consensus on identifying relevant aspects that could be addressed through the diversity of perspectives sharpened by ChatGPT's outputs. One participant noted:

> "I think the four of us agreed that the quotation selected by ChatGPT was important because, as the objective says, significant experiences are also something that marks a person, and we set that aside. So I think… the AI gives you certain things you haven't seen, and you can reread to see what else is happening, in what other

parts of the interview or the observation the same thing appears, and perhaps you didn't take it into account" (CaseA- A2)

Another participant stated:

> "What most caught our attention from our position as teachers was that it was more critical… I think that when we moved to working with ChatGPT, we realized that perhaps our bias as teachers did not allow us to see other things that the AI could see" (CaseB-B1)

Figure 2

How valuable was the exchange of perspectives with other group members when analysing the interview excerpt?

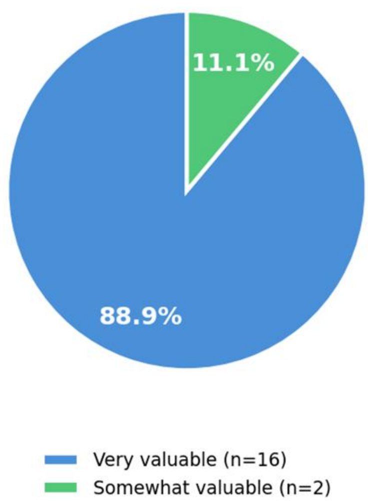
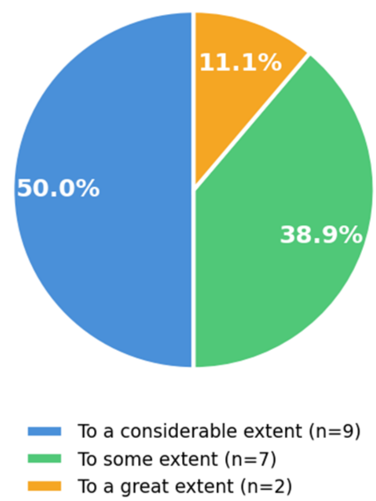


Figure 3

To what extent did the inputs generated by artificial intelligence influence the group's decision-making process?



Identification of personal bias and positionality enhanced with GenAI

ChatGPT offered input that helped students consider perspectives not addressed in their manual analysis. This reflects the research quality criterion of reflexivity, fostering a broader, more objective view. A key contribution was identifying personal bias in data selection and interpretation. For instance, one group admitted to choosing emotionally charged quotes, while ChatGPT highlighted more neutral segments:

> "What we mentioned [ChatGPT] was that, rather than being mechanical, it was actually more objective, because we aligned it with the specific research objectives, whereas we allowed ourselves to be guided a little by our emotions as researchers" (Case B- B1)

In another case, students noticed a preference for positive experiences, which contrasted with ChatGPT's selection of quotes that also included more negative or challenging situations.

> "I think we unconsciously selected only the positive ones. But while we were talking and looking at what Chat was showing us, we said, 'Wow, yes, this could also be meaningful because it helped the girl experience this reality check—from having everything come easy to needing to make an effort or do something for it.' It's also meaningful to consider not only the 'positive,' but also the 'negative,' in quotes, that any case might have" (CaseA-A2)

This recognition of bias was most evident in Case B teams. Their data—testimonies from in-service teachers—led participants to identify more strongly with interviewees. Beyond detecting bias, students developed explicit awareness of their positionality and defended it as an advantage. As one team put it: "As future teachers we can critically justify the results of the standardized tests; we have the sensitivity of educators, while the AI only analyzes it mechanically" (Case B-Reflexivity-writing-B2). This shows participants affirmed that educator sensitivity, grounded in contextual and relational understanding, can enrich the interpretation of situated qualitative data in ways GenAI cannot.

Survey results support this finding: 77.8% of respondents indicated that GenAI use enabled them to acknowledge their assumptions to a considerable or great extent, while 22.2% reported this occurred only to some extent (Figure 4). Likewise, 66.7% perceived that AI integration enhanced the robustness of findings to a considerable or great extent, while 33.3% reported this occurred to some extent (Figure 5).

Reflexivity on individual methodological decisions

The comparison between manual analysis and that conducted with ChatGPT enabled participants to broaden their perspective on the data and reflect on the methodological foundations of their decisions. In both cases, they acknowledged that juxtaposing manual and GenAI-supported approaches led them to question the solidity of their analytical choices and identify gaps in their research training. This was evident when participants questioned the possible criteria used by ChatGPT to select a significant quotation. As one participant explained:

> "We had this doubt about whether we should only select the quotations as they were, whether we could cut the sentences, or whether we could even join fragments together to form a single quotation—and the tool did that" (CaseA-A1).

While participants tended to fragment the text into shorter excerpts, GenAI often merged fragments into longer quotations, producing contrasts that highlighted different interpretative logics. Similarly, the choice of thematic keywords became a point of methodological interrogation. Participants recognized that "ChatGPT helped us broaden the conception of our (thematic) keywords. In our case, we limited ourselves to choosing only one keyword, which did not necessarily reflect what we were really thinking as researchers" (CaseB-Reflexivity-writing-B2).

Some participants took this reflection further, situating it at a structural level and linking it to gaps in their academic training. One student pointed directly to the curricular deficiencies behind these difficulties in conceiving a keyword composed of more than one term:

> "I want to point out that there is a deficiency in our curriculum regarding our research skills. How is it possible that we are in the tenth semester and it hadn't occurred to us that key themes could be more than one word?" (CaseB- B1).

Formulation of explicit ethical interrogations

Beyond methodological learning and critical questioning, the use of GenAI fostered reflection on MI. Participants stressed that ChatGPT cannot be treated merely as a technical resource; its use involves ethical considerations about the researcher's role, authorship of analysis, and the need for institutional guidelines.

A visible tension emerged from participants' initial suspicion. They doubted how far one could rely on GenAI without compromising personal authorship. As one participant noted, there was concern about "what is ethical" and the need to distinguish between using GenAI as an instrumental aid and the temptation to treat it as an "investigative voice":

> "As a group we leaned toward a positive perception of using AI. But at a personal level, I was afraid of what is ethical. You have to separate AI as a tool from AI as if it were a person helping me in the research. That did scare me" (CaseB-B1)

At the same time, other participants reinforced the principle that ultimate responsibility rests with the researcher. For them, GenAI may serve as a useful input but can never replace authorship or the academic commitment to producing original work:

> "I think it is important to mention the ethical part, because the role I am giving it, I think, is the correct one. I don't ask it to… It may provide an analysis, but it does not end up being definitive in my assignment, because the research carries my name. So I think that is very important, because it is quite ethical. I can use it as needed for the research, but beyond that, it has to be our part" (CaseB-B1)

An additional concern that emerged was the question of transparency in disclosing GenAI use. Some participants asked whether its use should be explicitly acknowledged in reports or theses: "I asked about the matter of disclosure—whether in that case it was necessary to state that I had used it for the analysis" (CaseA-A1).

Figure 4

To what extent did the use of artificial intelligence enable you to recognize your own assumptions or interpretations during the analysis?

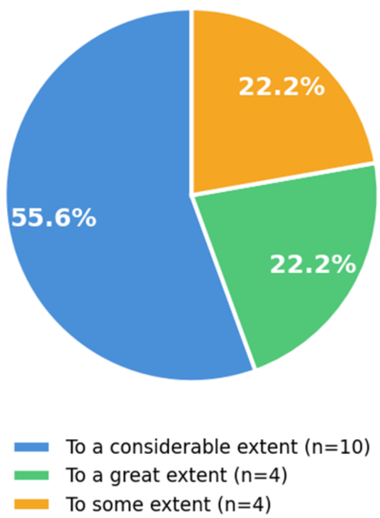
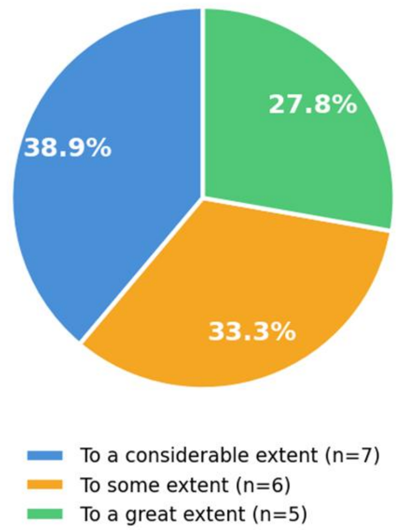


Figure 5

Compared to conducting the analysis exclusively manually, to what extent did the integration of artificial intelligence enhance the robustness of your findings?



Discussion

This study examined how GenAI-enhanced triangulation and reflexivity foster MI in two experiential learning cases. Two research questions guided the inquiry: (1) To what extent does GenAI triangulation strengthen credibility and dependability of findings? and (2) How do GenAI-guided reflexivity practices surface assumptions and biases, enhancing critical awareness?

The findings show that investigator triangulation with Sequential Augmented Analysis (SAA) strengthened reliability through three concrete mechanisms: reopening previously closed methodological decisions, prompting collective revision of thematic keywords toward greater conceptual specificity, and encouraging teams to return to original data to verify or refine their selections. Survey data reinforce this: over three-quarters of students rated peer exchanges highly, and nearly two-thirds acknowledged ChatGPT's positive influence on methodological decisions. Triangulation, defined as engaging multiple viewpoints to strengthen reliability and depth (Denzin, 2008; Creswell & Miller, 2000), was central. Bans and Tiimub (2021) emphasize convergence and complementarity—both evident here. ChatGPT also acted as catalyst for divergence by surfacing discrepancies (e.g., merged vs. fragmented quotations), stimulating dialogue that refined criteria. Collective thematization was stronger in Case A, where biographical narratives required greater negotiation of meaning.

Reflexivity was similarly enhanced across three dimensions: students identified personal biases in data selection, confirming that close comparison between manual and AI-generated outputs surfaces interpretive tendencies that would otherwise probably remain tacit as specialized studies had shown (Wachinger et al., 2025). Participants of the study questioned the methodological foundations of their analytical decisions and formulated explicit ethical interrogations around authorship, disclosure, and the appropriate role of AI in academic work, in line with emerging concerns in the literature on GenAI and academic integrity (Bittle & El-Gayar, 2025; Arar et al., 2025). Quantitatively, 78% of participants reported that GenAI enabled recognition of their own assumptions, and 72% perceived it strengthened the robustness of their findings. Across both cases and all six themes found, participants were more inclined to consider ChatGPT as a complementary mediator or tool rather than an investigative authority. These outcomes carry implications for both theory and practice. These reflexive gains were more salient in Case B for methodological decision-making, consistent with its data type and participants' professional identification.

Theoretical Implications

This study contributes theoretically to the intersection of HCAIE, Experiential Learning, and Methodological Integrity in Academia. First, the findings extend the conceptualization of investigator triangulation by showing that GenAI can stimulates dialogue and cross-referencing, extending Flick's (2018) argument that triangulation serves theoretical development beyond validation (Denzin, 2008; Creswell & Miller, 2000). Second, the study repositions MI as a collective and dialogic achievement: team-based negotiation prompted by AI-generated alternatives suggests that analytical quality emerges from social interaction (Levitt et al., 2017; Campbell et al., 2020). Third, GenAI functioned as a structural catalyst for reflexivity, externalizing and defamiliarizing students' interpretive tendencies at scale, offering a theoretically distinct mechanism that complements but differs from peer dialogue or supervisor feedback, consistent with Berger's (2015) and Finlay's (2002) accounts of reflexivity as requiring a productive "other". Fourth, the enactment of HCAIE principles at the micro-level of student-AI interaction, where transparency and explainability shaped critical appropriation of AI outputs, provides empirical grounding for more granular models of how human centrality is sustained in AI-enhanced learning environments (Yang et al., 2021; Fu & Weng, 2024).

Practical Implications

The findings offer concrete guidance for educators and curriculum designers integrating GenAI into research methodology training. The 4-phase SAA workflow used here (individual analysis, collaborative deliberation, AI-assisted input generation, and comparative reflection) provides a replicable model for instruction that can be adapted across disciplinary contexts, ensuring GenAI augments rather than replaces student judgment (Salinas-Navarro et al., 2024; Salvagno et al., 2023). Institutions can strengthen auditability by incorporating reflective memos and prompting logs as standard course deliverables, making student practice transparent and traceable (Averill, 2002; Olmos-Vega et al., 2023). The ethical tensions surfaced by participants, about authorship, disclosure, and appropriate AI reliance, should be treated as explicit curricular content rather than incidental byproducts; proactive instructional design that positions students as moral agents, supported by clear institutional disclosure policies, is necessary to address the regulatory gaps reported here (Bittle & El-Gayar, 2025; Arar et al., 2025). Finally, brief HCAIE-oriented facilitation, framing GenAI as a complementary mediator within authentic, team-based tasks, proved sufficient to generate meaningful gains in methodological awareness and ethical reasoning, suggesting feasibility even within constrained curricular time (Rana et al., 2025; Michalak & Ellikson, 2025).

Limitations

The study's methodological limitations centre on its exploratory, dual-case design and small participant pool, which restricts the generalizability of findings across diverse research contexts. Practically, the intervention's success is contingent upon students having prior QDA foundations and receiving specific training in CORI-f prompt engineering, potentially limiting its immediate application for absolute novices or in settings without institutional support. Furthermore, it is important to note that the chatbot's free version exhibits greater limitations and biases, which, as specialists have shown (Fleisig et al., 2024), are particularly more pronounced for non-English users. Theoretically, while the study utilizes Human-Centered AI and Experiential Learning, it has a "technological-oriented" focus that treats GenAI primarily as a mediating tool within human-controlled procedures. We think that the interpretation of this data could be further enriched by conceptual lenses acknowledging the complexity of agency re-distribution, where the chatbot is not merely an instrument but possibly an "actor" that reshapes the analytic process. Future research could explore agency to better understand the nature of methodological integrity in human-AI assemblages.

Conclusion

This study examined how GenAI-enhanced triangulation and reflexivity within HCAIE experiential learning contribute to MI. Findings across both cases show that integrating a chatbot acted as an analytic mediator, broadening perspectives, stimulating reflection, and exposing hidden biases.

Regarding the first research question, AI inputs reopened methodological discussions and supported triangulation by prompting convergence and complementarity. These processes led students to refine criteria, revisit original data, and provide more transparent justifications, strengthening reliability. The second research question focused on reflexivity, where ChatGPT encouraged recognition of personal and methodological biases, critical assessment of decision-making frameworks, and engagement with ethical issues such as authorship, disclosure, and transparency.

Importantly, technology's role was not to replace human judgment but to catalyze more rigorous and ethically grounded decision-making. Overall, results show that GenAI, when integrated into EL contexts under a HCAIE framework, contributes to MI through triangulation and reflexivity. At the same time, ethical concerns voiced by participants highlight the need for broader questions about responsibility in academia. Future research should build on these insights by developing replicable pedagogical models for GenAI integration in QDA training and assessing long-term impacts on students' methodological competence across diverse contexts.

Acknowledgements

The authors disclose no actual or perceived conflicts of interest. They have not received funding for this manuscript beyond academic time and logistical support from CISE, Faculty and Department of Education at PUCP. In the preparation of this article, ChatGPT-4 (free version) was used selectively to support language refinement at the level of grammar, syntax, and stylistic clarity. The substantive arguments, conceptual framing, data analysis, and interpretations remain

the sole responsibility of the authors (Crawford et al., 2023). The use of GenAI for grammar editing did not alter the analytical content; however, the iterative revision process prompted reflexive insights regarding the precision of key terms, the coherence of argumentative transitions, and the implicit assumptions embedded in certain formulations. In this sense, grammar-focused interaction with GenAI functioned not only as a linguistic aid but also as a reflective device, encouraging closer scrutiny of conceptual clarity and Global South authorial positioning within the manuscript. As researchers based in South America, we recognize the ethical tension inherent in using chatbots, which are trained on large-scale English datasets of partially opaque and potentially copyrighted provenance. This concern is particularly salient in our context, historically shaped by epistemic inequality and extractive dynamics, where questions of data justice and intellectual ownership are technical and political. While such tools can enhance our participation in global academic discourse, we therefore adopt a reflexive stance. The authors list the following CRediT contributions. Conceptualization: ME; methodology: ME, AS; formal analysis: ME; data curation: ME, AS, PB, JL; validation: ME, AS, PB; writing–original draft: ME, AS, PB, JL; supervision: ME, AS. All authors have read and agreed to the published version of the manuscript. The authors confirm they have met the ethical standards expected in this journal (Purvis & Crawford, 2024).

References

Annamalai, N., & Nasor, M. (2025). Exploring ChatGPT in education: unveiling learners' experiences through the lens of self-determination theory. *Smart Learning Environments, 12*(1), 59. https://doi.org/10.1186/s40561-025-00393-2

Alier, M., García Peñalvo, F. J., & D. Camba, J. (2024). Generative Artificial Intelligence in Education: From Deceptive to Disruptive. International Journal of Interactive Multimedia and Artificial Intelligence, 8(5), 5–14. https://doi.org/10.9781/ijimai.2024.02.011

Arar, K., Ozen, H., Polat, G., & Turan, S. (2025). Artificial intelligence, generative artificial intelligence and research integrity: a hybrid systemic review. *Smart Learning Environments, 12*(1), 44.https://doi.org/10.1186/s40561-025-00403-3

Arosio, L. (2025). Generative AI as a Teaching Tool for Social Research Methodology: Addressing Challenges in Higher Education. *Societies, 15*(6), 157. https://doi.org/10.3390/soc15060157

Athaluri, S., Manthena, S., Kesapragada, V., Yarlagadda, V., Dave, T., & Duddumpudi, R. (2023). Exploring the boundaries of reality: investigating the phenomenon of artificial intelligence hallucination in scientific writing through ChatGPT references. *Cureus, 15*(4). https://doi.org/10.7759/cureus.37432

Averill, J. (2002). Matrix analysis as a complementary analytic strategy in qualitative inquiry. *Qualitative Health Research, 12*(6), 855–866. https://doi.org/10.1177/104973230201200611

Baidoo-Anu, D., & Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of Artificial Intelligence, 7*(1), 52–62. https://doi.org/10.61969/jai.1337500

Bans-Akutey, A., & Tiimub, B. (2021). Triangulation in research. *Academia Letters, 2*(3392), 1–7. https://doi.org/10.20935/AL3392.

Berger, R. (2015). Now I see it, now I don't: Researcher's position and reflexivity in qualitative research. *Qualitative Research, 15*(2), 219–234. https://doi.org/10.1177/1468794112468475

Bijker, R., Merkouris, S. S., Dowling, N. A., & Rodda, S. N. (2024). ChatGPT for Automated Qualitative Research: Content Analysis. *Journal of medical Internet research*, *26*, e59050. https://doi.org/10.2196/59050

Bin-Nashwan, S., Sadallah, M., & Bouteraa, M. (2023). Use of ChatGPT in academia: Academic integrity hangs in the balance. *Technology in Society, 75*, Article 102370.10.1016/j.techsoc.2023.102370

Bittle, K., & El-Gayar, O. (2025). Generative AI and academic integrity in higher education: A systematic review and research agenda. *Information, 16*(4), 296. https://doi.org/10.3390/info16040296

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Bonyadi, A. (2023). Phenomenology as a research methodology in teaching English as a foreign language. *Asian-Pacific Journal of Second and Foreign Language Education, 8*(1), 11. https://doi.org/10.1186/s40862-022-00184-z

Campbell, R., Goodman-Williams, R., Feeney, H., & Fehler-Cabral, G. (2020). Assessing triangulation across methodologies, methods, and stakeholder groups: The joys, woes, and politics of interpreting convergent and divergent data. *American Journal of Evaluation, 41*(1), 125–144. https://doi.org/10.1177/1098214018804195

Chan, C., & Lee, K. (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). https://doi.org/10.1186/s40561-023-00269-3

Chen, Z., Chen, C., Yang, G., He, X., Chi, X., Zeng, Z., & Chen, X. (2024). Research integrity in the era of artificial intelligence: Challenges and responses. *Medicine, 103*(27), e38811. https://doi.org/10.1097/MD.0000000000038811

Chubb, L. (2023). Me and the machines: Possibilities and pitfalls of using artificial intelligence for qualitative data analysis. *International Journal of Qualitative Methods, 22*, https://doi.org/10.1177/16094069231193593

Cook-Sather, A., Bovill, C., & Felten, P. (2014). *Engaging students as partners in learning and teaching: A guide for faculty*. Wiley.

Crawford, J., Cowling, M., Ashton-Hay, S., Kelder, J., Middleton, R., & Wilson, G. (2023). Artificial intelligence and authorship editor policy: ChatGPT, Bard Bing AI, and beyond. *Journal of University Teaching and Learning Practice, 20*(5). https://doi.org/10.53761/1.20.5.01

Creswell, J. W. (2007). Five Qualitative Approaches to Inquiry. In J. W. Creswell (Eds.), Qualitative Inquiry and Research Design: Choosing among five Approaches (pp. 53-84). Thousands Oaks: Sage Publications.

Creswell, J., & Creswell, J. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage.

Creswell, J., & Miller, D. (2000). Determining validity in qualitative inquiry. *Theory into Practice, 39*(3), 124–130. https://doi.org/10.1207/s15430421tip3903_2

Creswell, J., & Poth, C. (2018). *Qualitative inquiry and research design: Choosing among five approaches*. Sage.

Denzin, N. (2008). *The landscape of qualitative research*. Sage.

Diaz, B., Chen, G., Charlotte, N., & Jaselskis, E. (2025). Supporting Generative AI Literacy: Exploring the Pedagogical Roles Students Assign ChatGPT and Impact on Course Grades. *Comunicar. 33(82), 46–61. https://doi.org/10.5281/zenodo.15993999*

Díaz, B., & Nussbaum, M. (2024). Artificial intelligence for teaching and learning in schools: The need for pedagogical intelligence. *Computers & Education*, 217, 105071. https://doi.org/10.1016/j.compedu.2024.105071

Eddles-Hirsch, K. (2015). Phenomenology and educational research. *International Journal of Advanced Research, 3*(8).

Eisenhart, M., & Jurow, A. S. (2011). Teaching qualitative research. *The SAGE handbook of qualitative research*, *4*, 669-714.

El Khoury, E. (2023). An equitable approach to academic integrity through alternative assessment. In S. Eaton (Ed.), *Handbook of academic integrity* (pp. 1–12). Springer. https://doi.org/10.1007/978-981-287-079-7_135-1

Etesse, M. (2024). *Introduction to Qualitative Data Analysis with Artificial Intelligence: Practical Guide to Use ChatGPT in Social and Educational Research.* CISEPA, Pontificia Universidad Católica del Perú. https://doi.org/10.2139/ssrn.5070929

Etesse, M. (2025a). AI and Qualitative Data Analysis in Education Research: Familiarization and Exploration of Interview Transcripts Enhanced with ChatGPT. In *International Conference on Communication and Applied Technologies* (pp. 285–295). Springer. https://doi.org/10.18800/9786124355172

Etesse, M. (2025b). Enseñar a investigar en la universidad: hacia un acompañamiento formativo con inteligencia artificial para el diseño crítico de objetivos de investigación. In Méndez et al. (Ed.) *Formar, transformar, acompañar: experiencias desde distintos espacios educativos* (pp. 771-794). Dykinson.

Etesse, M., Beretta, P., Gómez, F. & Vieira, D. (2026). Integrating Chatbots into Undergraduate Research Training: Developing AI Literacy through Experiential Learning. In *5th International Conference on Advanced Learning Technologies on Education & Research* (pp. 1-6). IEEE. https://doi.org/10.1109/ICALTER69698.2025.11355077.

Exintaris, B., Karunaratne, N., & Yuriev, E. (2023). Metacognition and critical thinking: Using ChatGPT-generated responses as prompts for critique in a problem-solving workshop (SMARTCHEMPer). *Journal of Chemical Education, 100*, 2972–2980. https://doi.org/10.1021/acs.jchemed.3c00481

Feuston, J., & Brubaker, J. (2021). Putting tools in their place: The role of time and perspective in human–AI collaboration for qualitative analysis. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW2), Article 469.https://doi.org/10.1145/3479856

Finlay, L. (2002). Negotiating the swamp: The opportunity and challenge of reflexivity in research practice. *Qualitative Research, 2*(2), 209–230. https://doi.org/10.1177/146879410200200205

Fleisig, E., Smith, G., Bossi, M., Rustagi, I., Yin, X., & Klein, D. (2024, November). Linguistic bias in ChatGPT: Language models reinforce dialect discrimination. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* (pp. 13541-13564). https://doi.org/10.48550/arXiv.2406.08818

Flick, U. (2017). Mantras and myths: The disenchantment of mixed-methods research and revisiting triangulation as a perspective. *Qualitative inquiry*, *23*(1), 46-57.

French, F., Levi, D., Maczo, C., Simonaityte, A., Triantafyllidis, S., & Varda, G. (2023). Creative use of OpenAI in education: case studies from game development. *Multimodal Technologies and Interaction, 7*(8), 81. https://doi.org/10.3390/mti7080081

Fu, Y., & Weng, Z. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centered AI practices. *Computers and Education: Artificial Intelligence, 7*, 100306. https://doi.org/10.1016/j.caeai.2024.100306

Gao, J., Choo, K., Cao, J., Lee, R., & Perrault, S. (2023). CoAIcoder: Examining the effectiveness of AI-assisted human-to-human collaboration in qualitative analysis. *ACM Transactions on Computer-Human Interaction, 31*(1), 1–38. https://doi.org/10.1145/3617362

Goyanes, M., Lopezosa, C., & Jordá, B. (2025). Thematic analysis of interview data with ChatGPT: Designing and testing a reliable research protocol for qualitative research. *Quality & Quantity*, 1–20. https://doi.org/10.1007/s11135-025-02199-3

Guilherme, A. (2019). AI and education: the importance of teacher and student relations. *AI & Society, 34*(1), 47–54. https://doi.org/10.1007/s00146-017-0693-8

Hamilton, E., Rosenberg, J., & Akcaoglu, M. (2016). The substitution augmentation modification redefinition (SAMR) model: A critical review and suggestions for its use. *TechTrends, 60*(5), 433–441. https://doi.org/10.1007/s11528-016-0091-y

Healey, M., & Jenkins, A. (2000). Kolb's experiential learning theory and its application in geography in higher education. *Journal of Geography, 99*(5), 185–195.

Henriksen, D., Woo, L., & Mishra, P. (2023). Creative uses of ChatGPT for education: A conversation with Ethan Mollick. *TechTrends, 67*(4), 595–600. https://doi.org/10.1007/s11528-023-00862-w

Käss, C. Brosig, M. Westner, and S. Strahringer, Short and sweet: Multiple mini case studies as a form of rigorous case study research. I*nformation Systems and e-Business Management*, *22*(2), 351–384, 2024. https://doi.org/10.1007/s10257-024-00674-2

Kuramitsu, K., Obara, Y., Sato, M., & Obara, M. (2023). KOGI: A Seamless Integration of ChatGPT into Jupyter Environments for Programming Education. Proceedings of the 2023 ACM SIGPLAN International Symposium on SPLASH-E, 50–59. https://doi.org/10.1145/3622780.3623648

Isiaku, L., Kwala, A., Sambo, K., & Isiaku, H. (2024). Academic evolution in the age of ChatGPT: An in-depth qualitative exploration of its influence on research, learning, and ethics in higher education. *Journal of University Teaching and Learning Practice, 21*(6), 67–91. https://doi.org/10.53761/7egat807

Jiang, J., Wade, K., Fiesler, C., & Brubaker, J. (2021). Supporting serendipity: Opportunities and challenges for human–AI collaboration in qualitative analysis. *Proceedings of the ACM on Human–Computer Interaction, 5*(CSCW1), Article 94. https://doi.org/10.1145/3449168

Kirwan, A. (2023). ChatGPT and university teaching, learning and assessment: Some initial reflections on teaching academic integrity in the age of Large Language Models. *Irish Educational Studies*, 1–18. https://doi.org/10.1080/03323315.2023.2284901

Kolb, A., & Kolb, D. (2005). Learning styles and learning spaces: Enhancing experiential learning in higher education. *Academy of Management Learning & Education, 4*(2), 193–212. https://doi.org/10.5465/AMLE.2005.17268566

Kong, Z., Adi, V., Segovia-Hernandez, J., & Sunarso, J. (2023). Complementary role of large language models in educating undergraduate design of distillation column: Methodology development. *Digital Chemical Engineering, 9*. https://doi.org/10.1016/j.dche.2023.100126

Kumar, S., Rao, P., Singhania, S., Verma, S., & Kheterpal, M. (2024). Will artificial intelligence drive the advancements in higher education? A tri-phased exploration. *Technological Forecasting and Social Change, 201*, 123258. https://doi.org/10.1016/j.techfore.2024.123258

Kusner, M., & Loftus, J. (2020). The long road to fairer algorithms. *Nature, 578*(7793), 34–36. https://doi.org/10.1038/d41586-020-00274-3

Lan, Y., & Chen, N. (2024). Teachers' agency in the era of LLM and generative AI. *Educational Technology & Society, 27*(1), I–XVIII. https://doi.org/10.30191/ETS.202401_27(1).PP01

Levitt, H., Motulsky, S., Wertz, F., Morrow, S., & Ponterotto, J. (2017). Recommendations for designing and reviewing qualitative research in psychology: Promoting methodological integrity. *Qualitative Psychology, 4*(1), 2. https://doi.org/10.1037/qup0000082

Lincoln, Y., & Guba, E. (1985). *Naturalistic Inquiry*. Sage.

Li, S., & Gu, X. (2023). A Risk Framework for Human-centered Artificial Intelligence in Education: Based on Literature Review and Delphi–AHP Method. *Educational Technology & Society*, 26(1), 187–202. https://www.jstor.org/stable/48707976
Lieberson, Hacer que cuente: la mejora de la investigación y la teoría sociales, Univ. of California Press, 2000.
Liu, R., Zenke, C., Liu, C., Holmes, A., Thornton, P., & Malan, D. (2024). Teaching CS50 with AI. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education* 750–756. https://doi.org/10.1145/3626252.3630938
Lodge, J., de Barba, P., & Broadbent, J. (2023). Learning with generative AI within a network of co-regulation. *Journal of University Teaching and Learning Practice, 20*(7), 1–10. https://doi.org/10.53761/1.20.7.02
Lopezosa, C., & Codina, L. (2023). ChatGPT y software CAQDAS para el análisis cualitativo. *Revista de investigación e innovación en ciencias de la salud. 5*(1), 1-5*.* https://doi.org/10.46634/riics.211
Macbeth, D. (2001). On “reflexivity” in qualitative research: Two readings, and a third. *Qualitative Inquiry, 7*(1), 35–68. https://doi.org/10.1177/107780040100700103
Mauthner, N., & Doucet, A. (2003). Reflexive accounts and reflexivity in qualitative data analysis. *Sociology, 37*(3), 413–431. https://doi.org/10.1177/00380385030373002
Megahed, F. M., Chen, Y. J., Ferris, J. A., Knoth, S., & Jones-Farmer, L. A. (2024). How generative AI models such as ChatGPT can be (mis)used in SPC practice, education, and research? An exploratory study. *Quality Engineering*, *36*(2), 287–315. https://doi.org/10.1080/08982112.2023.2206479
Mennella, T., & Quadros-Mennella, P. (2024). Student use, performance and perceptions of ChatGPT. *Journal of University Teaching and Learning Practice, 21*(1), 95–118. https://doi.org/10.53761/pgwk1a93
Michalak, R., & Ellixson, D. (2025). Fostering Ethical AI Integration in First-Year Writing: A Case Study on Human-Tool Collaboration in Artificial Intelligence Literacy. *Journal of Library Administration*, 65(3), 361–377. https://doi.org/10.1080/01930826.2025.2468136
Morris, T. (2019). Experiential learning – a systematic review and revision of Kolb’s model. *Interactive Learning Environments, 28*(8), 1064–1077. https://doi.org/10.1080/10494820.2019.1570279
Morse, J. (1994). “Emerging from the data”: the cognitive processes of analysis. In J. Morse (Ed.), *Critical issues in qualitative research methods* (pp. 23–43). Sage.
Moustakas, C. (2010). *Phenomenological research methods.* Sage.
Morrow, R., Rodriguez, A., & King, N. (2023). Paul Colaizzi’s descriptive phenomenological methodology. In *Introduction to phenomenology: Focus on methodology* (pp. 19–30). https://doi.org/10.4135/9781071909669.n8
Nguyen, A., Ngo, H., Hong, Y., Dang, B., & Nguyen, B. (2023). Ethical principles for AI in education. *Education and Information Technologies, 28*(4), 4221–4241. https://doi.org/10.1007/s10639-022-11316-w
Nguyen, N. (2023). Exploring the role of AI in education. *London Journal of Social Sciences, (6)*, 84–95. https://doi.org/10.31039/ljss.2023.6.108
Nguyen-Trung, K. (2025). ChatGPT in Thematic Analysis: Can AI become a research assistant in qualitative research?. *Quality & Quantity*, 1–34. https://doi.org/10.1007/s11135-025-02165-z

Olmos-Vega, F., Stalmeijer, R., Varpio, L., & Kahlke, R. (2023). A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. *Medical Teacher, 45*(3), 241–251. DOI: 10.1080/0142159X.2022.2057287

Park, Y., & Doo, M. (2024). Role of AI in blended learning: A systematic review. *International Review of Research in Open and Distributed Learning, 25*(1), 164–196. https://doi.org/10.19173/irrodl.v25i1.7566

Perkins, M., & Roe, J. (2024). The use of Generative AI in qualitative analysis: Inductive thematic analysis with ChatGPT. *Journal of Applied Learning and Teaching, 7*(1), 390–395. https://orcid.org/0000-0002-4479-4565

PUCP. (2016). *Reglamento del Comité de Ética de la Investigación*. Pontificia Universidad Católica del Perú.

Puentedura, R. (2014). Building transformation: An introduction to the SAMR model [Blog post]. https://www.hippasus.com/rrpweblog/archives/2014/08/22/BuildingTransformation_AnIntroductionToSAMR.pdf

Purvis, A., & Crawford, J. (2024). Ethical Standards in Educational Research and Publication. *Journal of University Teaching and Learning Practice, 21*(9). https://doi.org/10.53761/hqnqr710

Qiao, S., Fang, X., Zhang, R., Li, X., & Kang, Y. (2024). Generative AI for qualitative analysis in maternal health study: coding semistructures interviews using large language models (LLMs). *MedRxiv, 2024-09.* https://doi.org/10.1101/2024.09.16.24313707

Rana, V., Verhoeven, B., & Sharma, M. (2025). Generative AI in design thinking pedagogy: enhancing creativity, critical thinking, and ethical reasoning in higher education. *Journal of University Teaching and Learning Practice, 22*(4), 1–22. https://doi.org/10.53761/tjse2f36

Salinas-Navarro, D., Vilalta-Perdomo, E., Michel-Villarreal, R., & Montesinos, L. (2024). Using generative AI tools to enhance experiential learning for authentic assessment. *Education Sciences, 14*(1), 83. https://doi.org/10.3390/educsci14010083

Salvagno, M., Taccone, F., & Gerli, A. (2023). Artificial intelligence hallucinations. *Critical Care, 27*(1), 180. DOI: 10.1186/s13054-023-04473-y

Sivill, T. (2019). Ethical and statistical considerations in models of moral judgments. *Frontiers in Robotics and AI, 6*, 39. https://doi.org/10.3389/frobt.2019.00039

Shneiderman, B. (2022). Human-Centered AI. Oxford University PressOxford. https://doi.org/10.1093/oso/9780192845290.001.0001

Slotnick, R., & Boeing, J. (2025). Enhancing qualitative research in higher education assessment through AI integration: a path toward meaningful insights and a cautionary tale. *New Directions for Teaching and Learning, 2025*(182), 97–112. https://doi.org/10.1002/tl.20631

Stamenkov, G. (2023). Recommendations for improving research quality: relationships among constructs,verbs in hypotheses, theoretical perspectives,and triangulation. *Quality & Quantity, 57*(3), 2923–2946. https://doi.org/10.1007/s11135-022-01461-2

Talbott, M., & Lee, J. (2020). Ambiguous Terminology: A Challenge in Teaching Social Science Research Methods and Statistics. *International Journal of Teaching and Learning in Higher Education*, *32*(3), 519-527.

Tavakol, M., & Sandars, J. (2025). Twelve tips for using phenomenology as a qualitative research approach in health professions education. *Medical Teacher,* 1–6. https://doi.org/10.1080/0142159X.2025.2478871

Teixeira-Poit, S., Cameron, A., & Schulman, M. (2011). Experiential learning and research ethics: enhancing knowledge through action. *Teaching Sociology, 39*(3), 244–258. https://doi.org/10.1177/0092055X1140734

Tlili, A., Shehata, B., Adarkwah, M., Bozkurt, A., Hickey, D., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*(1), 15. https://doi.org/10.1186/s40561-023-00237-x

Uddin, S., Albert, A., Ovid, A., & Alsharef, A. (2023). Leveraging ChatGPT to aid construction hazard recognition and support safety education and training. *Sustainability, 15*(9), 7121. https://doi.org/10.3390/su15097121

Wachinger, J., Bärnighausen, K., Schäfer, L., Scott, K., & McMahon, S. (2025). Prompts, pearls, imperfections: Comparing ChatGPT and a human researcher in Qualitative Data Analysis. *Qualitative Health Research, 35*(9), 951–966. DOI: 10.1177/10497323241244669

Wang, M., Xu, X., Yang, L., Cai, D., & Yin, M. (2024). Unleashing ChatGPT's power: A case Study on Optimizing Information Retrievalin Flipped Classrooms via Prompt Engineering. *IEEE Transactions on Learning Technologies, 17*, 629–641. DOI: 10.1109/TLT.2023.3324714

Wagner, C., Kawulich, B., & Garner, M. (2019). A mixed research synthesis of literature on teaching qualitative research methods. *Sage Open*, *9*(3), 2158244019861488.

WMA. (2013). World Medical Association Declaration of Helsink: ethical principles for medical research involving human subjectsi. *JAMA, 310*(20), 2191–2194. DOI: 10.1001/jama.2013.281053

Xia, Q., Weng, X., Ouyang, F., Lin, T., & Chiu, T. (2024). A scoping review on generative AI in assessment. *IJETHE, 21*(1), 40. https://doi.org/10.1186/s41239-024-00468-z

Xie, X., & Ding, S. (2023). ChatGPT in higher education. In *Proceedings of the 2023 Int. Conf. on Educational Knowledge and Informatization* (pp. 14–18).

Xu, W. (2019). Toward human-centered AI: a perspective from human-computer interaction. *Interactions, 26*(4), 42–46. https://doi.org/10.1145/3328485

Xu, W., Ge, L., & Gao, Z. (2021). Human-AI interaction: An emerging interdisciplinary domain for enabling human-centered AI. *CAAI Transactions on Intelligent Systems, 16*(4), 605–621. https://doi.org/10.48550/arXiv.2112.01920

Xu, W., & Ouyang, F. (2022). Systematic review of AI role in the educational system based on a proposed conceptual framework. *Education and Information Technologies, 27*(3), 4195–4223. https://doi.org/10.1007/s10639-021-10774-y

Yang, S., Ogata, H., Matsui, T., & Chen, N. (2021). Human-centered artificial intelligence in education: Seeing the invisible through the visible.*Computers and Education: AI, 2*, 100008. https://doi.org/10.1016/j.caeai.2021.100008

Yardley, S., Teunissen, P., & Dornan, T. (2012). Experiential learning: transforming theory into practice. *Medical Teacher, 34*(2), 161–164. https://doi.org/10.3109/0142159X.2012.643264

Yin, R. (2018) Case study research and applications. Sage.

Zhang, H., Wu, C., Xie, J., Lyu, Y., Cai, J., & Carroll, J. (2023). Redefining qualitative analysis in the AI era. *arXiv preprint arXiv:2309.10771*. https://doi.org/10.1016/j.chbah.2025.100144

Zhang, H., Wu, C., Xie, J., Lyu, Y., Cai, J., & Carroll, J. (2025). Harnessing AI in qualitative research. *Computers in Human Behavior: Artificial Humans, 4*, 100144. https://doi.org/10.1016/j.chbah.2025.100144

Zhao, T., Wallace, E., Feng, S., Klein, D., & Singh, S. (2021). Calibrate before use. *arXiv.org.*

Zhao, X., Aydeniz, M., & Yuan, F. (2023). AI in biomedical informatics education. In *2023 IEEE EMBS Int. Conf. on Biomedical and Health Informatics* (pp. 1–4). 10.1109/BHI58575.2023.10313394

Zhou, X., Zhang, J., & Chan, C. (2024). Students' perceptions of AI in higher education. *Journal of University Teaching and Learning Practice, 21*(6), 126–145. https://doi.org/10.53761/xzjprb23