Nurse Education in Practice 90 (2026) 104648

Contents lists available at ScienceDirect

Nurse Education in Practice

journal homepage: www.elsevier.com/locate/issn/14715953

Faculty perceptions of AI-versus human-summarized narrative exit survey data across three nursing programs

Staci S. Reynolds *, Elaine D. Kauschinger, Allen Cadavero, Stefanie Conrad, Jacquelyn M. McMillian-Bohler, Michelle Webb

Department of Nursing, Duke University School of Nursing, 307 Trent Drive, Durham, NC, USA

ARTICLE INFO

Keywords:
Generative artificial intelligence
Nursing education
Data mining
Evaluation
Faculty perceptions

ABSTRACT

Aim: The purpose of this study was to compare faculty perceptions of the quality of artificial intelligence (AI)-generated versus human-generated summaries of narrative exit survey data to assess the feasibility of AI integration into program evaluation processes.
Background: Generative AI tools are increasingly used in higher education to streamline data analysis. In nursing education, student evaluations offer rich insights but are time-consuming to summarize. AI tools like Microsoft Copilot offer potential efficiencies but raise concerns about reliability, bias and the preservation of reflective pedagogy and student voice.
Design: A cross-sectional, descriptive pilot study design was used.
Methods: Five faculty members independently rated summaries generated by Microsoft Copilot and by human analysis using a 7-point Likert scale. Ratings were based on accuracy, clarity, bias and relevance.
Results: Quality ratings of the AI-generated summaries were higher (mean=5.9) compared with the human-generated summaries (mean=5.0).
Conclusions: This pilot project suggests integrating AI as a supportive tool rather than a replacement for human review. The overall intent was to assist faculty in improving efficiency in program evaluations by using AI, in conjunction with human review, to maintain fidelity to the student voices and context.

1. Introduction

In academic assessment and evaluation, narrative student evaluations provide invaluable insight into nursing education programs; however, they require significant time and effort to summarize accurately. Summarizing these data are often completed by a human which can be a time consuming, tedious process prone to error and biases (Roy and Rambo-Hernandez, 2021; Shaikh and Doudpotta, 2019). One solution to expedite this task may be to use an artificial intelligence (AI) tool to summarize the data. Also known as text mining, reports of using AI to summarize large amounts of textual data from student evaluations have been published in the literature (Rybinski and Kopciuszewska, 2021; Shaikh and Doudpotta, 2019; Takaki and Dutra, 2024). While this approach is not new, some may be reluctant to use AI due to concerns with accuracy and quality of the subsequent data summaries. Replacing human analysts with AI-generated summaries raises important considerations about accuracy, interpretive fidelity and educational validity. For faculty who rely on student feedback to inform teaching and program improvement, questions remain about how AI-generated summaries compare to those created by humans and what implications this shift holds for pedagogy and evaluation practice. Additionally, studies investigating faculty perceptions of the accuracy of AI-generated reports on student surveys are limited.

The integration of generative AI into student evaluation represents more than a technical innovation; it introduces pedagogical and theoretical considerations that extend beyond efficiency. Narrative comments are central to understanding the learning environment and reflect co-constructed meaning between students and faculty (Bosun-Arije, 2023). Automating interpretation may identify broad patterns efficiently but risks overlooking less-represented perspectives or subtle affective details that are important for reflective pedagogy, particularly in nursing education, where relational practice and contextual nuance are

* Corresponding author.
E-mail addresses: Staci.reynolds@duke.edu (S.S. Reynolds), Elaine.kauchinger@duke.edu (E.D. Kauschinger), Allen.cadavero@duke.edu (A. Cadavero), Stefanie.conrad@duke.edu (S. Conrad), Jacquelyn.mcmillianbohler@duke.edu (J.M. McMillian-Bohler), Michelle.webb@duke.edu (M. Webb).

https://doi.org/10.1016/j.nepr.2025.104648
Received 14 May 2025; Received in revised form 6 October 2025; Accepted 17 November 2025
Available online 19 November 2025

emphasized (Hitch, 2024; Plathe et al., 2021; Williams, 2024). Li et al. (2024) demonstrated in a healthcare qualitative data analysis that GPT-4 can detect key themes when properly prompted; however, human reviewers identified a richer diversity of subthemes and captured more nuanced insights.

From a socio-constructivist perspective, effective teaching and learning in nursing education requires collaboration between students and educators, fostering interactive environments where knowledge is co-constructed through shared experiences and engagement (Bosun-Arije, 2023). This approach, however, is not yet fully replicable by AI systems at this time. In parallel, critical data studies remind us that AI systems are not neutral; outputs are shaped by training data and design choices that can privilege certain voices while marginalizing others (Kaminski, 2022). Empirical work suggests that while faculty often value the organization and efficiency of AI-generated summaries, concerns persist regarding interpretive accuracy, representational fidelity and the loss of nuance in comparison to human analysis (Lyu et al., 2025). Pedagogically, reliance on AI raises risks of reduced faculty engagement with student voices. Overreliance on algorithmic summaries may weaken reflective practice and diminish opportunities for faculty to interpret feedback in ways that align with disciplinary and contextual knowledge (Saleh et al., 2025).

AI can provide value as a supportive tool by supporting reflective engagement by assisting in organizing and visualizing data, highlighting themes and identifying novel insights, thereby enabling researchers to review and act on feedback more effectively while maintaining human oversight and reflexivity in the analysis process (Cook et al., 2025). Strengthening faculty trust and encouraging adoption requires AI systems to include transparent explanations of how themes are derived, representative excerpts from student responses and clear indicators of uncertainty (Karran et al., 2025; Rosenbacke et al., 2024).

Taken together, these perspectives suggest that AI is best positioned as a complementary resource within program evaluation. Not only must there be a "human in the loop," the human must be *in charge of* the loop (Cook et al., 2025). Faculty development in AI literacy will be essential to ensure that educators can critically assess AI outputs, maintain interpretive oversight and preserve the central role of student voice in decision-making (Blomquist et al., 2025; Cucci et al., 2025). For nursing education, where reflective and relational pedagogy are integral, safeguarding interpretive fidelity should remain a priority as AI becomes more embedded in evaluative processes (Levin et al., 2025).

1.1. Purpose

The purpose of this study was to compare the perceived quality of summaries of narrative data from student exit surveys between those developed using an AI-powered tool (Microsoft Copilot) and summaries that were human-generated. Authors hypothesized that AI-generated summaries would be perceived by faculty as equal to or superior in quality to human-generated summaries based on their accuracy, clarity and relevance to decision-making. Overall, the usation of AI was intended to assist faculty in improving efficiency in program evaluations. This is among the first studies to explore faculty perception specifically, as opposed to student perceptions or comparative accuracy alone.

2. Methods

2.1. Study design, setting and participants

A cross-sectional, descriptive pilot study design was used. The STROBE reporting guidelines were used for structuring this report (von Elm et al., 2007). This study took place at a private school of nursing in the Southeastern United States. Narrative exit survey data were evaluated from students who graduated in December 2023.

All students slated to graduate in December 2023 received an anonymous exit survey to provide feedback regarding their experiences while a student at the school of nursing. The survey was available for 49 days between November 2023 and January 2024. Of the 206 students who received the survey, 134 completed it (65 % response rate); response rates per program follow: 77.1 % (54/70) for the pre-licensure nursing program; 57.8 % (63/109) for the Master of Science in Nursing (MSN) program; and 63 % (17/27) for the Doctor of Nursing Practice (DNP) program. Students voluntarily completed the survey and anonymity and confidentiality were maintained. Narrative data were evaluated based on participants' responses to two open-ended narrative questions: (1) What aspects of the program did you enjoy the most? and (2) What changes would you recommend for the program for current and future students?

For the current study, the project leader contacted six faculty members with diverse experience across the three programs via email to gauge interest in participating in this scholarship opportunity; four agreed to participate. These four faculty members, along with the project leader, reviewed the narrative data. These five faculty members represented each of the programs, with two faculty members having primary teaching responsibilities in the pre-licensure nursing program (AC, JMB); two in the MSN program (MW, EK) and one in the DNP program (SSR). Faculty had a mean 16.4 years (SD=9.3, range 6–30) of experience in higher education; 4 were female and 1 was male. Two had very little prior knowledge of AI, with 2 having moderate and 1 a high level (self-rated) of AI knowledge. Narrative data from student exit surveys were de-identified.

2.2. Measures

In February 2024, a project analyst manually reviewed all narrative comments for each of the three programs from the two open-ended survey questions and developed a summary of findings to share with program faculty. The project analyst was instructed to identify recurring themes and topics for each of the two open-ended questions. If a topic was mentioned by fewer than two or three students, it was included in an "other" category. If a topic was mentioned by more than three students, the topic was included as a "theme." For each "theme" the project analyst was instructed to include the number of students who mentioned the topic and provide a sample of corresponding student quotes in a list format. The manual human-generated summaries took approximately 30 h to complete.

Microsoft Copilot was used within the institution's secure environment to analyze de-identified narrative exit survey responses. Per the university's instructional technology policy, Microsoft Copilot was used as the AI tool as it keeps the data entered into the prompt behind the university technology firewall, protecting the information from being widely shared across the internet. For each of the six datasets (responses to the two open-ended questions for each of the three programs), a single standardized prompt, developed collaboratively with input from two AI experts, was entered into Microsoft Copilot to generate a summary; the prompt is shown in Table 1. The prompt instructed the model to identify themes grounded in the text, organize findings by program and question, include representative quotations and avoid unsupported inferences. Prior research indicates that structured prompts enhance the reliability and clarity of large language model outputs (Meskó, 2023; Yan et al., 2024). Each dataset was processed in a single execution on October 25, 2024. Using Microsoft Copilot, summaries were generated in approximately 15 s (x 6 = 90 s total). Examples of the human- versus AI-generated summaries can be found in supplement 1.

Five faculty members received three datasets, including the (1) raw narrative data, (2) human-generated summary and (3) AI-generated summary for the three nursing programs (nine datasets total). Per instructions, faculty thoroughly reviewed each dataset and completed an online survey assessing their perceptions of the quality of the summaries. The author-developed survey consisted of 14 statements, including six regarding the human-generated summary, seven regarding

Table 1
Prompt entered into Microsoft Copilot to generate the AI summary of student narrative exit survey data.

You are an expert in evaluating student exit surveys. Please analyze the responses provided by the [pre-licensure, MSN, or DNP program] exit surveys and categorize them into clear thematic areas, such as:

- ***Curriculum*** *(e.g., course content, flow of courses),*
- ***Clinical experiences*** *(e.g., quality of clinical placements, preceptor support),*
- ***Faculty support*** *(e.g., availability, mentorship),*
- ***Leadership and practice readiness,***
- ***Interprofessional collaboration,*** *and*
- ***Do students mention any other significant areas?***

Ensure you:

- *Summarize the feedback in bullet points under each theme.*
- *Include the frequency of specific comments or suggestions within each category (e.g., '7 out of 15 comments mentioned a need for more simulation experiences').*
- *Highlight both* ***positive feedback*** *and* ***opportunities for improvement.***
- *Identify and* ***prioritize actionable recommendations*** *for improvement based on frequency and potential impact on student satisfaction, outcomes, or readiness for nursing practice.*

the AI-generated summary and one free-text field for comments. Survey questions appear in supplement 2. The 13 quantitative statements were rated on a 7-point Likert scale (1 =strongly disagree, 7 =strongly agree). The authors developed and assessed the survey for face validity. Faculty members provided their ratings between October 2024 and February 2025.

2.3. Ethical considerations

This study was approved by the university's Institutional Review Board and determined exempt, not meeting the definition of human subjects' research (Pro00115947).

2.4. Analysis

Descriptive statistics, including mean and standard deviation (SD), were calculated to compare quality ratings between summary types. A thematic analysis of open-ended responses provided context for the observed quantitative trends.

3. Results

Overall means and SDs appear in Table 2. Faculty rated the overall quality of the AI-generated summaries 18 % higher than the human-generated summaries (mean=5.9, SD=1.2 vs mean=5, SD=1.7). Responses regarding the AI-generated summaries were the same or higher for all statements rated, except for two; the statement "compared with the raw data, the summary accurately reflected the student's comments" was rated higher for the human-generated summary (mean=5.4, SD=1.5) versus the AI-generated summary (mean=5.2, SD=1.5) for the pre-licensure program data and the statement "the summary was free from bias" was rated higher for the human-generated summary (mean=5.6, SD=0.9) versus the AI-generated summary (mean=5.4, SD=1.5) for the MSN program data.

From the free-text comments, faculty appreciated the organization, thematic structure and actional feedback; faculty comments included "I liked the organization of the AI-generated report by themes and by positive feedback and opportunities for improvement. I think this increased the readability of the evaluation summaries" and "The AI-generated summary was well-organized and provided actionable next steps, making it useful for decision-makers."

However, some faculty perceived the AI-generated report as either having failed to capture key elements of the data or having provided redundancies in themes (e.g., "The AI-generated response did not seem to capture the spirit of the responses, with not as much discussion about the nuances of the frustration with the flipped classroom" and "[The AI-generated report] omitted some critical details and introduced themes that were not a major focus for students"). Despite these limitations, most faculty recommended that AI summaries could replace human-generated summaries if combined with expert review (mean= 5.5, SD=1.6).

4. Discussion

4.1. Summary

Overall study results showed that AI-generated summaries were of higher perceived quality and could replace human-generated summaries. Both quantitative and qualitative feedback was extracted from faculty, which adds richness to the findings. The authors' initial hypothesis that AI-generated summaries could provide a level of quality equal to or higher than that of the human-generated summaries was supported by the findings. Interestingly, the means of the AI-generated summaries for the programs increased between the pre-licensure to the MSN to the DNP program. This could be because this was the order prompts were entered and summaries obtained from Microsoft Copilot,

Table 2
Means and Standard Deviations (SD) of faculty survey responses.

Question	Pre-licensure Mean (SD)		MSN Mean (SD)		DNP Mean (SD)		Total Mean (SD)	
	Human-generated	AI-generated	Human-generated	AI-generated	Human-generated	AI-generated	Human-generated	AI-generated
Compared to the raw data, the summary accurately reflected the student's comments.	5.4 (1.5)	5.2 (1.5)	5.2 (1.3)	5.2 (1.5)	5.2 (2.5)	6 (0.7)	5.3 (1.7)	5.5 (1.2)
The summary provided actionable next steps.	4.4 (2.4)	6 (0.7)	4.4 (1.8)	6.4 (0.5)	4.2 (1.6)	6.4 (0.5)	4.3 (1.8)	6.3 (0.6)
The summary was free from bias.	4.8 (1.6)	5.4 (1.1)	5.6 (0.9)	5.4 (1.5)	5.2 (1.1)	5.2 (1.3)	5.2 (1.2)	5.3 (1.2)
The summary provided helpful details found within the raw data.	4.8 (2.2)	5.6 (1.1)	5.2 (1.1)	6.2 (0.8)	5.2 (1.9)	6.2 (0.8)	5.1 (1.7)	6 (0.9)
The summary was appropriate (ie, results made contextual sense).	5.6 (1.1)	5.6 (0.5)	5.2 (1.9)	6.2 (0.4)	5.8 (1.6)	6.2 (0.4)	5.5 (1.5)	6 (0.5)
Please rate the overall quality of the summary.	5.4 (1.5)	5.4 (1.5)	5 (1.7)	6.2 (0.8)	4.6 (2.1)	6 (1.2)	5 (1.7)	5.9 (1.2)
The AI-generated summary could take the place of the human-generated summary.	–	5 (2.0)	–	5.2 (1.6)	–	6.2 (1.3)	–	5.5 (1.6)

indicating that the quality of the AI-generated summaries may have improved with subsequent usage. Copilot may have learned how to develop high-quality summaries over time, leading to the higher means for AI-generated summaries for the DNP program compared with the pre-licensure program. Indeed, generative AI has been shown to continuously learn, adapt and improve responses over time (Pescapè, 2024).

Data from this project supports the use of Microsoft Copilot to generate summaries of narrative data from student exit surveys at this institution. Previously, these data were manually summarized by an individual and took a substantial amount of time; however, using Microsoft Copilot, summaries were generated in mere seconds. Using AI tools for these types of administrative tasks can help improve efficiency in higher education, providing more time for staff and faculty to focus on other tasks that require human attention. The time savings between the human-generated summaries (30 h) versus AI-generated summaries (90 s) is substantial and demonstrates how AI can improve institutional processes. In addition to saving time, AI speeds up data organization and theme identification. The themes produced by AI provide a quick overview of patterns in the data. Human review is still necessary to confirm accuracy, but the time is spent on analysis and validation rather than an initial review of the data.

Additionally, while the human-generated summary provided overall positive and negative themes from student comments based on the prompt used, the AI-generated summaries went further by including a list of actionable next steps. Faculty found this component particularly helpful because the purpose of the exit survey questions is to elicit feedback that informs program improvement. Providing practical, actionable recommendations helped guide faculty in identifying areas for future enhancement. Additionally, AI may not be as influenced by affective judgments or selective attention; it processes all available data points without omitting comments based on subjective impressions. While AI models are influenced by training data, prompt framing and potential algorithmic biases that can affect the tone and meaning, using a comprehensive approach may help ensure that feedback is not overlooked simply because it was missed during manual review. By offering a second, systematic perspective, AI reduces the likelihood of gaps in interpretation and supports a more complete and objective analysis of student input.

4.2. Implications

AI offers promising opportunities to support the analysis of narrative student evaluations by reducing faculty workload and making it more feasible to review qualitative feedback at scale. This efficiency can encourage more consistent use of student voices in program evaluation and instructional improvement. At the same time, the findings of this study and the broader literature highlight that interpretive fidelity, contextual nuance and faculty trust remain critical considerations.

AI-generated summaries can effectively identify broad themes and recurring patterns but may omit less-represented perspectives or affective details that are important for reflective pedagogy (Hitch, 2024). Faculty perceptions of AI are influenced by concerns over transparency, trust and the perceived fairness of its outputs (Karran et al., 2025; Lyu et al., 2025). As such, AI is best positioned as a supportive tool rather than a replacement for human analysis. Hybrid approaches, where AI generates preliminary thematic maps and faculty refine them, align with the findings of Rosenbacke et al. (2024), which highlight the importance of combining AI outputs with human expertise to enhance efficiency while maintaining interpretive depth and preserving relational and contextual understanding.

Moving forward, faculty development in AI literacy will be essential to ensure that educators can critically appraise AI outputs, recognize both strengths and limitations, maintain interpretive oversight and align AI integration with pedagogical goals while addressing ethical, technical and infrastructural challenges (Blomquist et al., 2025; Cucci et al., 2025). In nursing education, in particular, where relational pedagogy and professional judgment are central, protecting interpretive fidelity must remain a priority (Levin et al., 2025).

4.3. Comparison with previous studies

While previous literature has used text mining through AI tools to summarize student evaluation data (Rybinski and Kopciuszewska, 2021; Shaikh and Doudpotta, 2019; Takaki and Dutra, 2024), there is a dearth of evidence regarding faculty perceptions of the perceived quality of AI-generated reports. Fuller et al. (2024) explored the level of agreement between instructor-identified and AI-identified themes from student course evaluations; results showed there was high agreement between the instructor and AI (ChatGPT) results, with AI taking less time to analyze the data. Similar to the current study, Fuller et al. (2024) identified a need to use AI as a tool to assist with analysis and to not rely solely on its outputs for conclusions.

Er et al. (2024) assessed student perceptions of instructor versus AI-generated feedback. Contrasting with findings from the current study, Er et al. (2024) found that students perceived instructor feedback as significantly more useful than the AI-generated feedback; authors posited that AI models need to be further trained on data specific to educational contexts. In the study conducted by Hostetter et al. (2024), faculty assessed the quality of student versus AI-written reflections, finding both rated similarly by participants, similar to the current study. Lastly, Demoulin and Coussement (2020) evaluated the acceptance of text-mining systems by information systems managers, finding that the quality of information influences behavioral intentions and use of text-mining tools, highlighting that support from managers play a key role in determining usage of text-mining tools.

While faculty may use AI for coursework or teaching (Kim et al., 2025), previous studies that assessed faculty's perceived quality of AI-generated summaries were not identified. As such, the current study contributes to this gap in evidence by showing that AI-generated summaries are of high perceived quality and its use is supported by faculty.

4.4. Limitations

This pilot study has several limitations. First, only five faculty members evaluated the datasets; review by additional faculty members may have yielded different results. Due to this small sample size, we were unable to conduct inferential statistical tests to determine differences between groups. Data were reviewed from one school of nursing, which may limit generalizability. Faculty completed an author-developed survey to assess quality of the summaries; while this survey underwent review for face validity, no further psychometric testing was completed. As the data were not blinded, faculty reviewers were aware of whether the summaries were human- or AI-generated, which may have introduced bias into their evaluations. Faculty may have consciously or unconsciously favored or penalized summaries based on authorship. Although faculty were instructed to focus solely on the content, prior research suggests that preconceived notions about the reliability of AI-generated content may have influenced their perceptions. Khlaif et al. (2023) found that while ChatGPT-generated scientific content was rated as high quality, concerns related to authorship and research integrity affected reviewer judgment. Similarly, Cheng et al. (2023) reported that evaluators often identified AI-generated abstracts and rated them as less credible despite comparable quality to human-written content. Additionally, faculty perceptions do not necessarily equate to validity, especially given the potential influence of bias or novelty effects in AI tools. These findings suggest that faculty biases may have influenced the assessments in the current study.

Additionally, other more robust AI tools, such as ChatGPT, may have provided higher quality AI-generated summaries. The prompts used to generate the AI summaries, while developed with input from AI experts to promote clarity and consistency, may represent a limitation. The structure and specificity of the prompts likely influenced the outputs generated by Microsoft Copilot, as the quality and focus of AI-generated content can vary depending on the wording and framing of the prompts. In this study, the prompts emphasized thematic organization and actionable recommendations, which may have shaped how the AI tool categorized and prioritized student feedback. This reliance on prompt design may have affected the overall quality of the summaries and their alignment with faculty perspectives. Future studies may consider evaluating the impact of different prompt structures on AI-generated outputs and exploring strategies to optimize prompt design, thereby supporting program evaluation efforts more effectively. Recent studies have highlighted the effects of prompt engineering on the performance of generative AI in healthcare and education settings, underscoring the need for careful prompt construction to achieve accurate and contextually relevant outputs (Yu et al., 2023; Zaghir et al., 2024). This study contributes to the growing body of work on AI literacy in nursing education. Faculty familiarity with AI capabilities and limitations will be critical as generative tolls are increasingly embedded into educational practice.

4.5. Future research

Future studies on using AI tools to evaluate narrative data may consider using more robust AI tools, additional faculty and staff review and different prompts. A larger sample size of faculty and/or staff to review the data would allow inferential statistical tests to be conducted to determine differences between human- and AI-generated summaries. While AI-generated reports may help improve efficiency in evaluating narrative student exit survey data, higher educational institutions may continue to include a human element in the review process to ensure data quality and appropriateness. Using one structured process for both human-generated and AI-generated summaries would ensure consistency and accuracy. Lastly, future research should investigate longitudinal patterns of faculty use, the impact of AI-supported evaluation on teaching decisions and whether hybrid approaches yield measurable improvements in pedagogy.

5. Conclusion

In conclusion, results showed that AI-generated reports of student narrative exit survey data are of high perceived quality. While AI-generated reports may be able to take the place of human-generated reports based on faculty perceptions, faculty should continue to have a role in evaluations to ensure accuracy in reporting. Using AI to summarize evaluation data is a practical way to integrate AI into daily workflows to improve efficiency. AI-generated summaries of student evaluations are perceived by faculty as high quality and timesaving. Educators should be encouraged to engage with AI tools critically, using them as assistants in creating efficient and contextually sensitive feedback systems.

Funding

None.

CRediT authorship contribution statement

Allen Cadavero: Writing – review & editing, Validation, Formal analysis. **Reynolds Staci:** Writing – original draft, Visualization, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jacquelyn M. McMillian-Bohler:** Writing – review & editing, Project administration, Formal analysis. **Michelle Webb:** Writing – review & editing, Project administration, Formal analysis. **Stefanie Conrad:** Writing – review & editing, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Elaine D. Kauschinger:** Writing – review & editing, Project administration, Formal analysis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to acknowledge Jennie De Gagne PhD, DNP, RN, NPD-BC, CNE, ANEF, FAAN for her assistance with developing the AI prompt used in this study.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.nepr.2025.104648.

References

Blomquist, J., Alderden, J., Llewellyn, S., Haynes, M., Dworak, E., 2025. Student and faculty perceptions and use of generative artificial intelligence in nursing education. J. Nurs. Educ. 64 (8), 519–522. https://doi.org/10.3928/01484834-20250312-06.

Bosun-Arije, F.S., 2023. A constructive approach to learning and teaching in nursing education. Nurse Educ. Pract. 68, 103591. https://doi.org/10.1016/j.nepr.2023.103591.

Cheng, S.-L., Tsai, S.-J., Bai, Y.-M., Ko, C.-H., Hsu, C.-W., Yang, F.-C., Tsai, C.-K., Tu, Y.-K., Yang, S.-N., Tseng, P.-T., Hsu, T.-W., Liang, C.-S., Su, K.-P., 2023. Comparisons of Quality, correctness and similarity between ChatGPT-generated and human-written abstracts for basic research: cross-sectional study. J. Med. Internet Res. 25 (1), e51229. https://doi.org/10.2196/51229.

Cook, D.A., Ginsburg, S., Sawatsky, A.P., Kuper, A., D'Angelo, J.D., 2025. Artificial Intelligence to Support qualitative data analysis: Promises, approaches, pitfalls. Acad. Med. 100 (10), 1134. https://doi.org/10.1097/ACM.0000000000006134.

Cucci, F., Marasciulo, D., Romani, M., Soldano, G., Cascio, D., De Nunzio, G., Caldararo, C., Rubbi, I., Vitale, E., Lupo, R., Conte, L., 2025. The contribution of artificial intelligence in nursing education: a scoping review of the literature. Nurs. Rep. 15 (8), 283. https://doi.org/10.3390/nursrep15080283.

Demoulin, N.T.M., Coussement, K., 2020. Acceptance of text-mining systems: the signaling role of information quality. Inf. Manag. 57 (1), 103120. https://doi.org/10.1016/j.im.2018.10.006.

Er, E., Akçapınar, G., Bayazıt, A., Noroozi, O., Banihashem, S.K., 2024. Assessing student perceptions and use of instructor versus AI-generated feedback. Br. J. Educ. Technol. https://doi.org/10.1111/bjet.13558.

Fuller, K., Morbitzer, K.A., Zeeman, J.M., M. Persky, A., C. Savage, A., McLaughlin, J.E., 2024. Exploring the use of ChatGPT to analyze student course evaluation comments. BMC Med. Educ. 24, 423. https://doi.org/10.1186/s12909-024-05316-2.

Hitch, D., 2024. Artificial intelligence augmented qualitative analysis: The way of the future? Qual. Health Res. 34 (7), 595–606. https://doi.org/10.1177/10497323231217392.

Hostetter, A.B., Call, N., Frazier, G., James, T., Linnertz, C., Nestle, E., Tucci, M., 2024. Student and faculty perceptions of generative artificial intelligence in student writing. Teach. Psychol., 00986283241279401 https://doi.org/10.1177/00986283241279401.

Kaminski, J., 2022. Theory applied to informatics: Socio-Technical Theory. Can. J. Nurs. Inform. 17 (3–4). ⟨https://cjni.net/journal/?p=10442⟩.

Karran, A.J., Charland, P., Trempe-Martineau, J., Ortiz de Guinea Lopez de Arana, A., Lesage, A.-M., Sénécal, S., Léger, P.-M., 2025. Multi-stakeholder perspective on responsible artificial intelligence and acceptability in education. Npj Sci. Learn. 10 (1), 44. https://doi.org/10.1038/s41539-025-00333-2.

Khlaif, Z.N., Mousa, A., Hattab, M.K., Itmazi, J., Hassan, A.A., Sanmugam, M., Ayyoub, A., 2023. The potential and concerns of using AI in scientific research: ChatGPT performance evaluation. JMIR Med. Educ. 9 (1), e47049. https://doi.org/10.2196/47049.

Kim, J., Klopfer, M., Grohs, J.R., Eldardiry, H., Weichert, J., Cox, L.A., Pike, D., 2025. Examining faculty and student perceptions of generative AI in university courses. Innov. High. Educ. https://doi.org/10.1007/s10755-024-09774-w.

Levin, C., Zaboli, A., Turcato, G., Saban, M., 2025. Nursing judgment in the age of generative artificial intelligence: a cross-national study on clinical decision-making performance among emergency nurses. Int. J. Nurs. Stud. 172, 105216. https://doi.org/10.1016/j.ijnurstu.2025.105216.

Li, K.D., Fernandez, A.M., Schwartz, R., Rios, N., Carlisle, M.N., Amend, G.M., Patel, H. V., Breyer, B.N., 2024. Comparing GPT-4 and human researchers in health care data analysis: qualitative description study. J. Med. Internet Res. 26 (1), e56500. https://doi.org/10.2196/56500.

Lyu, W., Zhang, S., Chung, T., Sun, Y., Zhang, Y., 2025. Understanding the practices, perceptions and (dis)trust of generative AI among instructors: a mixed-methods study in the U.S. higher education. Comput. Educ. Artif. Intell. 8, 100383. https://doi.org/10.1016/j.caeai.2025.100383.

Meskó, B., 2023. Prompt Engineering as an important emerging skill for medical professionals: Tutorial. J. Med. Internet Res. 25 (1), e50638. https://doi.org/10.2196/50638.

Pescapè, A., 2024. Exploring the current state and future potential of generative artificial intelligence using a generative artificial intelligence. In: Santoianni, F., Giannini, G., Ciasullo, A. (Eds.), Mind, Body and Digital Brains. Springer Nature Switzerland, pp. 37–56. https://doi.org/10.1007/978-3-031-58363-6_4.

Plathe, H., Solheim, E., Eide, H., 2021. Nursing Students' and preceptors' experiences with using an assessment tool for feedback and reflection in supervision of clinical skills: A qualitative pilot study. Nurs. Res. Pract. 2021 (1), 5551662. https://doi.org/10.1155/2021/5551662.

Rosenbacke, R., Melhus, Å., McKee, M., Stuckler, D., 2024. How Explainable Artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: systematic review. JMIR AI 3 (1), e53207. https://doi.org/10.2196/53207.

Roy, A., Rambo-Hernandez, K.E., 2021. There's so much to do and not enough time to do it! A case for sentiment analysis to derive meaning from open text using student reflections of engineering activities. Am. J. Eval. 42 (4), 559–576. https://doi.org/10.1177/1098214020962576.

Rybinski, K., Kopciuszewska, E., 2021. Will artificial intelligence revolutionise the student evaluation of teaching? A big data study of 1.6 million student reviews. Assess. Eval. High. Educ. 46 (7), 1127–1139. https://doi.org/10.1080/02602938.2020.1844866.

Saleh, Z.T., Rababa, M., Elshatarat, R.A., Alharbi, M., Alhumaidi, B.N., Al-Za'areer, M.S., Jarrad, R.A., Al Niarat, T.F., Almagharbeh, W.T., Al-Sayaghi, K.M., Fadila, D.E.S., 2025. Exploring faculty perceptions and concerns regarding artificial intelligence Chatbots in nursing education: potential benefits and limitations. BMC Nurs. 24 (1), 440. https://doi.org/10.1186/s12912-025-03082-0.

Shaikh, S., Doudpotta, S.M., 2019. Aspects based opinion mining for teacher and course evaluation. Sukkur IBA J. Comput. Math. Sci. 3 (1), 34–43. https://doi.org/10.30537/sjcms.v3i1.375.

Takaki, P., Dutra, M.L., 2024. Text mining applied to distance higher education: a systematic literature review. Educ. Inf. Technol. 29 (9), 10851–10878. https://doi.org/10.1007/s10639-023-12235-0.

von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C., Vandenbroucke, J.P., STROBE Initiative, 2007. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ (Clin. Res. Ed.) 335 (7624), 806–808. https://doi.org/10.1136/bmj.39335.541782.AD.

Williams, R.T., 2024. Paradigm shifts: exploring AI's influence on qualitative inquiry and analysis. Front. Res. Metr. Anal. 9, 1331589. https://doi.org/10.3389/frma.2024.1331589.

Yan, S., Knapp, W., Leong, A., Kadkhodazadeh, S., Das, S., Jones, V.G., Clark, R., Grattendick, D., Chen, K., Hladik, L., Fagan, L., Chan, A., 2024. Prompt engineering on leveraging large language models in generating response to InBasket messages. J. Am. Med. Inform. Assoc. 31 (10), 2263–2270. https://doi.org/10.1093/jamia/ocae172.

Yu, P., Xu, H., Hu, X., Deng, C., 2023. Leveraging generative AI and large language models: a comprehensive roadmap for healthcare integration. Healthc. (Basel Switz.) 11 (20), 2776. https://doi.org/10.3390/healthcare11202776.

Zaghir, J., Bjelogrlic, M., Goldman, J.-P., Ehrsam, J., Gaudet-Blavignac, C., Lovis, C., 2024. Human-machine interactions with clinical phrase prediction system, aligning with Zipf's least effort principle? PLOS ONE 19 (12), e0316177. https://doi.org/10.1371/journal.pone.0316177.