Nurse Educator

An AI-Facilitated Virtual Teaching Assistant for Undergraduate Nursing Honors Research: An Embedded Mixed-Methods Evaluation

Jamie Qiao Xin Ng, BSc, RN; Jia Ning Chew, BSc, RN; Thitinut Akkadechanunt, PhD, RN; Joelle Yan Xin Chua, PhD, RN; Sook Mun Ng; and Shefaly Shorey, PhD, RN

ABSTRACT

Background: Artificial intelligence (AI) is increasingly integrated into higher education. However, evidence on theory-informed AI interventions supporting nursing research training remains limited.

Purpose: To evaluate an AI-facilitated teaching assistant (INSPIRE-AI) on final-year undergraduate nursing honors students' research self-efficacy, motivation, and research interest, and explore students' experiences of using INSPIRE-AI.

Methods: An embedded mixed-methods study comprising a one-group quasi-experimental pretest/posttest design with postintervention semistructured qualitative interviews. Nursing students received access to INSPIRE-AI throughout the honors year. Quantitative survey data (N = 146) were analyzed using paired *t*-tests and repeated-measures general linear models.

Results: Research self-efficacy improved significantly ($P < .001$). Students reported that INSPIRE-AI supported structured thinking and reduced uncertainty, though engagement varied due to trust concerns, perceived surveillance, and preference for familiar AI tools.

Conclusions: Together, these findings suggest that INSPIRE-AI has the potential to support research self-efficacy through structured scaffolding; however, this interpretation should be considered alongside the broader educational support that students received throughout the honors program.

Keywords: educational technology, generative artificial intelligence, nursing students, TIDieR checklist

Cite this article as: Ng JQX, Chew JN, Akkadechanunt T, et al. An AI-facilitated virtual teaching assistant for undergraduate nursing honors research: an embedded mixed-methods evaluation. *Nurse Educ.* 2026;XX(XX):XX-XX. doi: 10.1097/NNE.0000000000002294

Undergraduate research training is central to contemporary nursing education because it prepares nursing graduates to critically engage with evidence and develop competencies for research-informed practice.[1] Honors programs cultivate these competencies through independent research projects; however, students often encounter challenges integrating methodological knowledge, academic writing, and sustained self-directed learning.[2] Variability in supervisory practices and limited faculty capacity may further contribute to inconsistent support, reducing students' confidence, motivation, and interest in research, potentially limiting their future scholarly engagement.[3]

Research self-efficacy, motivation, and interest are important outcomes in independent research learning. Drawing on social cognitive theory, self-efficacy influences individuals' willingness to initiate and persist in complex tasks.[4] Higher research self-efficacy has been associated with stronger engagement, honors performance, and future research aspirations.[5] Motivation and interest are similarly important because they

Author Affiliations: Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Singapore (Ms Ng, Ms Chew, Dr Chua, and Dr Shorey); Faculty of Nursing, Chiang Mai University, Chiang Mai, Thailand (Dr Akkadechanunt); Artificial Intelligence Centre for Educational Technologies, National University of Singapore, Singapore (Ms Ng); Yong Loo Lin School of Medicine, National University of Singapore, Singapore (Dr Shorey); and Yeo Boon Khim Mind Science Centre, National University of Singapore, Singapore (Dr Shorey).

Jamie Qiao Xin Ng and Jia Ning Chew share joint first authorship for this manuscript.

The authors thank the AI Centre for Educational Technologies (AICET), National University of Singapore, for their tireless efforts in the development of INSPIRE-AI and support in the implementation of this study. We also acknowledge our alumni, Miss Brenda Ng Jinyu, Mr Travis Lanz-Brian Pereira, Mr Jeremy Seow Soon Leong, Miss Abielle Liu, and Miss Arielle Chezka Goh, for their contributions to the needs analysis and user-testing of INSPIRE-AI. We further thank Chiang Mai University for the Visiting Professor program that has made this collaborative work possible.

This research is supported by the National Research Foundation Singapore under the AI Singapore Program (AISG Award No: AISG2-TC-2023-009-AICET).

The authors declare no conflicts of interest.

Correspondence: Dr Shorey, Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Level 5, Centre for Translational Medicine, Block MD6, 14 Medical Drive, Singapore 117599 (nurssh@nus.edu.sg).

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (www.nurseeducatoronline.com).

Accepted for Publication: July 27, 2026

Early Access:

DOI: 10.1097/NNE.0000000000002294

shape students' effort and sustained involvement in research.[6,7] Multiple studies have demonstrated that higher research motivation and interest are linked to greater research self-efficacy and scholarly productivity.[8,9] We posit that interventions that enhance these interrelated constructs will see the most immediate effects on research self-efficacy and potential secondary effects on research motivation and interest.

Artificial intelligence (AI) presents new opportunities to provide on-demand, personalized academic support, and emerging evidence in nursing education suggests potential benefits for clinical nursing competencies, motivation, and self-efficacy.[10,11] However, most educational AI tools are general-purpose systems rather than purpose-built interventions for nursing research training.[10] Without pedagogical grounding, AI-generated guidance may remain fragmented or superficial, potentially limiting students' development of metacognition, critical thinking, and independent inquiry. Evidence remains limited on theory-informed AI interventions for undergraduate nursing research education.

To address this gap, we developed and evaluated the Integrated Nursing Support for Progress in Independent Research & Education (INSPIRE-AI), a theory-informed, AI-facilitated virtual teaching assistant for undergraduate nursing honors students. INSPIRE-AI combined structured learning pathways with on-demand research guidance to support students across the honors research process. Because educational technologies' success depends on both design and implementation context, an embedded mixed-methods design was used to examine changes in research learning outcomes and explore students' experiences, adoption, and interpretation of INSPIRE-AI within the broader honors learning environment.[12]

Aims

This study had 2 aims: (1) to evaluate changes in research outcomes (research self-efficacy, research motivation, and research interest) of undergraduate nursing honors students with access to INSPIRE-AI, and (2) to explore students' experiences and perceived acceptability of using INSPIRE-AI.

Methods

Study Design

An embedded mixed-methods design was employed, comprising a one-group quasi-experimental pretest/posttest quantitative component and postintervention qualitative component. Qualitative findings were used to contextualize quantitative outcomes.[12] Reporting followed the Mixed Methods Reporting in Rehabilitation & Health Sciences (MMR-RHS) guidelines (Supplemental Digital Content, Checklist https://links.lww.com/NE/C452).[13]

Setting and Participants

This study was conducted between July 2025 and February 2026 at the Alice Lee Centre for Nursing Studies, National University of Singapore. Final-year honors students complete an independent primary or secondary research thesis under a thesis advisor's guidance. Alongside INSPIRE-AI, students had access to the redesigned honors program, including a pre-semester bridging workshop, a centralized comprehensive resource guide, librarian and biostatistics consultations, postgraduate students and alumni mentoring, and supervisor advisory support.[14]

INSPIRE-AI Intervention

INSPIRE-AI is an AI-facilitated virtual teaching assistant delivered through a conversational chatbot interface. It operates in two modes within the same chat: (1) a structured curriculum mode, guiding students through sequenced research learning activities, and (2) an on-demand query mode for questions specific to students' individual thesis projects. The system used ChatGPT-4o, author-developed prompt engineering, and retrieval-augmented generation (RAG) to ground responses in curated undergraduate-level academic materials.

The intervention was grounded in three educational frameworks: (1) Vygotsky Zone of Proximal Development,[15] (2) Bloom taxonomy,[16] and (3) the load reduction instruction framework,[17] to provide adaptive scaffolding, support progression toward higher-order thinking, and reduce cognitive load. Detailed implementation specifications are described in Supplemental Digital Content 1, TIDieR Checklist https://links.lww.com/NE/C493. Chatbot pathways and examples of structured curriculum modules, prompt engineering, RAG knowledge sources, and representative interface screenshots are provided in Supplemental Digital Content, Appendix A1 https://links.lww.com/NE/C453.

Data Collection

Quantitative data were collected in July 2025 (pretest) and January 2026 (posttest). Students received an orientation covering appropriate use, academic integrity, and the intended role of INSPIRE-AI as a scaffold rather than a substitute for independent work, curriculum support, or supervisor input. Usage analytics were collected, and a daily usage limit was implemented.

Eligible students were recruited via email and through in-person briefings by an independent research assistant. Interested students completed the survey through a secure ethics-approved Qualtrics survey link (Provo, Utah). Names and contact information were collected for follow-up and subsequently de-identified. For the qualitative component, maximum variation purposive sampling was used based on differences in posttest research self-efficacy scores and INSPIRE-AI usage.

Online semistructured interviews were conducted using Zoom (Zoom Communications, 2026), including students who did not use INSPIRE-AI to explore reasons for nonuse. The interview guides were informed by past literature (Supplemental Digital Content, Appendix B https://links.lww.com/NE/C453), pilot tested in the first interview, and retained without changes.[14,18] The duration of the interviews ranged from 45 to 67 minutes.

Outcomes

The primary outcome was research self-efficacy, while secondary outcomes included research motivation, research interest, academic anxiety, and perceptions of the research training environment. The questionnaire included 5 validated instruments and an author-developed questionnaire on the students' sociodemographic information, prior research training, and engagement with curriculum components. The instruments have been validated in other international contexts with acceptable or higher internal consistency.[19–22]

Research Self-Efficacy

Research self-efficacy was measured using the 38-item Librarian Research Confidence Scale, developed by Brancolini and Kennedy.[23] The Librarian Research Confidence Scale was selected because it provides comprehensive coverage of the research competencies expected within our undergraduate honors program, including literature searching, systematic review methods, research design, critical appraisal, and scholarly writing. Compared with existing nursing-specific research self-efficacy instruments, it more comprehensively captures literature searching and evidence synthesis skills that constitute a substantial component of our honors curriculum. Items were rated on a 5-point Likert scale from 1 (Not at all confident) to 5 (Very confident). Total scores were summed, with a higher score indicating higher research self-efficacy. For this study, the scale demonstrated excellent internal consistency for pretest ($\alpha = 0.969$) and posttest ($\alpha = 0.965$).

Research Motivation

Research motivation was assessed using the 20-item Research Motivation Scale, measuring intrinsic reward, extrinsic reward, and failure avoidance.[24] Items were rated on a 5-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). Subscale scores were summed, with a higher score demonstrating higher research motivation. Internal consistency was good across subscales, with pretest ($\alpha = 0.679$-0.873) and posttest ($\alpha = 0.770$-0.899).

Research Interest

Research interest was measured using the 16-item Interest in Research Questionnaire.[25] Items were scored on a 5-point Likert scale ranging from 1 (very low interest) to 5 (very high interest). Higher total scores indicated greater research interest. The scale demonstrated excellent internal consistency for pretest ($\alpha = 0.946$) and posttest ($\alpha = 0.956$).

Research Training Environment

Perceptions of the research training environment were measured using the 18-item Research Training Environment-Short Form, comprising interpersonal and instructional subscales.[26] Items were rated on a 5-point Likert scale from 1 (Strongly Disagree) to 5 (Strongly Agree). Higher total scores indicated a more conducive research environment. Internal consistency was good for the pretest ($\alpha = 0.858$) and posttest ($\alpha = 0.871$).

Academic Anxiety

Academic anxiety was measured using the 11-item Academic Anxiety Scale.[27] Items were rated on a 4-point Likert scale from 1 (Not at all typical of me) to 4 (Very typical of me). Higher total scores indicated higher academic anxiety. The scale demonstrated good internal consistency at the pretest ($\alpha = 0.885$) and posttest ($\alpha = 0.889$).

Sample Size Calculation

Given the exploratory nature of this educational evaluation and the fixed size of the honors cohort, a pragmatic sample was used. A priori calculations suggested that approximately 90 participants would be required to detect a small within-subject effect with 80% power at a two-sided alpha of 0.05.[28] Allowing for a 30% attrition rate, 110 participants were targeted. However, the eligible cohort size limited recruitment, and 76 students completed the baseline (pretest) assessment.

Analysis

Quantitative Analysis

Statistical analyses were conducted using SPSS version 29 (IBM Corp, Armonk, New York). Descriptive statistics summarized participant characteristics, honors program experiences, and INSPIRE-AI use. Pre- to posttest changes were examined using paired *t*-tests. Effect sizes were calculated using Hedges' *g* to account for small-sample bias. To explore whether outcome trajectories differed by INSPIRE-AI use, repeated-measures general linear models were fitted with time as the within-subject factor and INSPIRE-AI use as the between-subject factor. Covariates were selected a priori based on conceptual relevance and observed associations with the outcomes. Complete-case analyses were conducted; hence, results should be interpreted as complete-case estimates. INSPIRE-AI use was categorized based on self-reported engagement. As uptake was not randomized, these analyses are exploratory and should not be interpreted as causal. Continuous outcomes were judged to be sufficiently normally distributed to support parametric analyses based on visual

inspection of histograms of the *Q*-*Q* plots. Statistical significance was set at a two-sided *P* value of .05.

Qualitative Analysis

Interviews were audio-recorded, transcribed verbatim, and analyzed using a reflexive thematic analysis framework.[29] Two independent researchers developed inductive codes after repeated readings of transcripts. Codes were iteratively reworked into broader themes by identifying shared meanings, and regular analytic discussions were held to interrogate assumptions, refine interpretations, and enhance reflexivity. Rather than aiming for data saturation, sampling ceased when the dataset was judged to offer sufficient breadth and depth to address the research questions.[30] Trustworthiness was supported through reflexive memoing, an audit trail of code and theme development (Supplemental Digital Content, Appendix C https://links.lww.com/NE/C453), thick description of the educational context, and reflexive journaling.[31–33]

Ethical Considerations

Ethical approval was obtained from the Department Ethics Review Board at the Alice Lee Centre for Nursing Studies (No. 000151). Students were informed that nonparticipation or withdrawal would not influence their access to the INSPIRE-AI intervention or their grades. Participant confidentiality was maintained using unique identification numbers.

Results

Demographic Data and Honors Experience

A total of 146 students, 76 at baseline and 70 immediately postintervention, completed the pre- and posttest survey, respectively, with 16 interviewed. Six students were lost to follow-up and excluded from analyses (Supplemental Digital Content, Figure 1 https://links.lww.com/NE/C454). The mean age of participants was 22.9 years (SD = 1.13), and most were female (n = 65, 85.5%). Full sociodemographic details are listed in Supplemental Digital Content, Table 1 https://links.lww.com/NE/C455.

More than half of the students reported difficulty balancing thesis demands with personal life (n = 48, 68.6%). Attendance for in-person and online classes was high. Most honors program components were rated favorably, with slightly lower ratings for the Librarian MOOC and Biostatistics Support (Supplemental Digital Content, Appendix D1 https://links.lww.com/NE/C453). Out of 70 students, 25 (35.7%) did not use INSPIRE-AI in their honors year. INSPIRE-AI was most used to address thesis-specific questions (n = 41, 57.7%), followed by clarifying research concepts (n = 34, 47.9%). Eleven students reported difficulties with use, primarily related to response lag or interface navigation (Supplemental Digital Content, Appendix D2 https://links.lww.com/NE/C453).

Main Analysis

Paired *t*-tests showed a statistically significant increase in research self-efficacy from pretest to posttest (mean difference [MD] 30.6, 95% CI 24.7-36.6, $P < .001$), indicating a meaningful improvement in students' confidence in conducting research tasks, including conceptualizing research questions, applying methodology, and academic writing. In contrast, no significant changes were observed in research motivation, research interest, perceived research training environment, or academic anxiety (Table 1). These findings suggest that while students' confidence improved over time, their underlying motivation and perceptions of the learning environment remained relatively stable throughout the honors year. Given that INSPIRE-AI was implemented alongside broader enhancements to the curriculum, these quantitative findings should be interpreted as reflecting changes across the overall educational experience.

Subgroup Analysis

To explore whether outcomes differed by INSPIRE-AI uptake, participants were categorized as users or nonusers. INSPIRE-AI use was entered as a fixed factor in the subsequent GLM analyses. In the univariate analyses, users differed from nonusers in significantly higher failure avoidance scores (MD = 2.673, 95% CI 1.247-4.10, $P < .001$) and significantly lower perceived instructional research training environment scores (MD = –2.029, 95% CI –0.6821 to 2.953, $P = .046$). In adjusted analyses, INSPIRE-AI use was nonsignificant for research self-efficacy scores (MD = 3.58, 95% CI –5.54 to 12.69, $P = .436$) or research interest (MD = 1.22, 95% CI –1.90 to 4.33, $P = .438$), while there was no statistically significant association between INSPIRE-AI use and intrinsic motivation (MD = –3.50, 95% CI –1.97 to 1.27, $P = .069$). Only failure avoidance remained statistically significantly higher among users (MD = 1.73, 95% CI 0.51-2.95, $P = .006$). As uptake was self-selected, observed differences may reflect preexisting learner characteristics rather than the effects of INSPIRE-AI itself, and reverse causality cannot be excluded. Detailed subgroup model outputs are provided in Supplemental Digital Content, Appendix E https://links.lww.com/NE/C453.

Qualitative Results

Students' experiences with INSPIRE-AI were encapsulated in 3 themes and 9 subthemes (Supplemental Digital Content, Figure 2 https://links.lww.com/NE/C456).

Theme 1: Uncertainty and Risk Perception in Using INSPIRE-AI

Students' engagement with INSPIRE-AI was shaped by uncertainty about its legitimacy, scope, and potential risks.

Table. Paired *t*-Tests Between Continuous Outcome Measures at Pretest (n = 76) and Posttest (n = 70)

Variable	Pretest (n = 76)		Posttest (n = 70)		Paired *t*-Test			
Continuous	Mean	SD	Mean	SD	Mean Difference (95% CI)	*t*	Hedges' g (95% CI)	*P* value
Research self-efficacy	82.8	22.3	113.5	22.9	30.6 (24.7 to 36.6)	10.3	1.22 (0.90 to 1.52)	<.001
Research motivation								
Intrinsic motivation	31.3	4.83	32.1	5.66	0.80 (–0.09 to 1.69)	1.80	0.21 (–0.02 to 0.45)	.076
Extrinsic motivation	16.9	3.70	16.8	3.43	–0.13 (–0.79 to 0.53)	–0.39	–0.04 (–0.28 to 0.19)	.698
Failure avoidance	19.1	3.33	19.0	3.67	–0.10 (–0.88 to 0.68)	–0.25	–0.03 (–0.26 to 0.20)	.800
Interest in research	46.2	10.4	47.9	11.8	1.64 (–0.56 to 3.85)	1.48	0.18 (–0.06 to 0.41)	.142
Research training environment	61.5	7.96	61.9	8.46	0.40 (–1.29 to 2.09)	0.47	0.06 (–0.18 to 0.29)	.639
Interpersonal	27.4	3.92	27.0	4.49	–0.33 (–1.21 to 0.57)	–0.74	–0.09 (–0.32 to 0.15)	.462
Instructional	34.2	4.68	34.9	4.43	0.73 (–0.28 to 1.74)	1.44	0.17 (–0.06 to 0.40)	.156
Academic anxiety	23.1	6.50	24.6	7.14	1.47 (–0.28 to 3.22)	1.68	0.20 (–0.04 to 0.43)	.098

Epistemic Uncertainty About INSPIRE-AI's Legitimacy and Scope

Some students hesitated to rely on INSPIRE-AI as they were unsure who developed it, with one student questioning its credibility because they "don't know who designed it." Others perceived it to be confined to institutional materials, reducing confidence in its suitability for honors-level research guidance.

Perceived Surveillance and Fear of Academic Misconduct

Students were conscious that the school might be "watching" their interactions with INSPIRE-AI, hence more "careful" because they "don't know if there are penalties" for inappropriate use. To avoid "being judged," some therefore diverted ambiguous questions to mainstream AI platforms or avoided using INSPIRE-AI altogether.

Interface-Related Uncertainty and Cognitive Load

INSPIRE-AI's requirement for students to select study-design-specific chatbots disrupted familiar help-seeking habits. One student expressed, "it's cognitively overwhelming … I just go into Chat (ChatGPT), open one chat and then I just ask a question." Daily usage limits further disrupted engagement and momentum, making the tool feel less convenient than mainstream AI platforms.

Theme 2: INSPIRE-AI as a Catalyst for Independent Research Learning

INSPIRE-AI was regarded as a facilitative yet personalized scaffold that helped structure students' thinking according to their needs.

Developing Cognitive Frameworks for Research Through Question-Based Scaffolding

Students valued INSPIRE-AI early in their thesis journey as its prompts provided a "clear direction" and a "framework of what to think." Its questioning approach made systematic thinking more intuitive and transferable across assignments and thesis writing.

Adaptive Scaffolding That Regulates Cognitive and Emotional Load

Students described INSPIRE-AI as well-balanced, supporting sustained engagement without becoming overwhelming. Additionally, students appreciated INSPIRE-AI's ability to organize fragmented and preliminary ideas and provide individualized stepwise guidance according to their level of understanding.

Recognizing INSPIRE-AI as a Trusted, Reliable Source of Guidance

Some students viewed INSPIRE-AI as a reliable source as it provided more specialized, academically grounded guidance. One student shared that "it's very targeted to my needs … I know that it's developed by our school." This reduced extensive cross-checking, enabling them to progress more efficiently.

Tensions Between Self-Directed Learning and AI Reliance

A contrasting account from a nonuser suggested that resistance to INSPIRE-AI could also reflect a desire to preserve independent learning. The student expressed concerns that reliance on AI could undermine deeper learning and acknowledged that they were inclined to uncritical acceptance of AI-generated outputs when

under time pressure: "Whatever answer it gives me, I'll take it as the truth."

Theme 3: The Ecosystem of Academic Support Shaping INSPIRE-AI Use

Students' decisions to adopt, limit, or avoid INSPIRE-AI were shaped by the broader honors support ecosystem.

Navigating Multiple Sources of Academic Guidance

Students did not view INSPIRE-AI as replacing supervisors, whose input remained essential for contextual judgment. Some felt they were "doing double work" when checking INSPIRE-AI's responses with advisors, while others preferred mainstream AI platforms for faster outputs or direct answers. One student expressed, "It just takes longer on INSPIRE-AI than it would take for me to use it on Gemini."

Habits, Social Norms, and Peer Validation Shaping Uptake

Uptake of INSPIRE-AI was also influenced by familiarity and peer norms. Mainstream AI tools felt "easier" and "automatic," while INSPIRE-AI felt "relatively new." Limited peer use discouraged experimentation, while positive peer endorsement encouraged some students to try the tool.

Integration

Guided by the framework proposed by Younas, Fàbregues, Munce, and Creswell,[34] integration of findings indicated that the improvement in research self-efficacy was supported by qualitative accounts of INSPIRE-AI facilitating structured thinking, uncertainty reduction, and guided engagement. In contrast, statistically insignificant changes in research motivation and interest were contextualized by uneven uptake, trust concerns, perceived academic risk, and preference for familiar AI tools. These findings indicate that the educational value of INSPIRE-AI is co-determined by both the intervention design and the broader learning environment. The qualitative findings, therefore, explain not only why research self-efficacy improved but also why broader motivational outcomes remained unchanged despite access to the intervention. The integrated findings are presented in Supplemental Digital Content, Table 2 https://links.lww.com/NE/C457.

Discussion

This embedded mixed-methods study examined research learning outcomes of undergraduate nursing honors students who had access to INSPIRE-AI, alongside their experiences of using it. Research self-efficacy improved significantly over time, whereas research motivation, research interest, academic anxiety, and perceptions of the research training environment showed little change. As students were concurrently exposed to several honors program supports, these findings should be interpreted within the broader learning system rather than the effects attributable to INSPIRE-AI alone. Qualitative findings suggest that INSPIRE-AI may have supported confidence by making research reasoning more explicit through structured prompts, stepwise guidance, and organization of ideas. The improvement in research self-efficacy is consistent with Bandura social cognitive theory, where guided support and mastery experiences strengthen confidence in performing complex tasks.[4] Students consistently described INSPIRE-AI as being most useful when it made research reasoning explicit through structured questioning and stepwise guidance, suggesting that its educational value lay primarily in supporting cognitive scaffolding rather than providing answers. However, the absence of significant changes in motivation and research interest may be partly explained by uneven uptake and engagement due to uncertainty about INSPIRE-AI, which may have limited its integration into their learning processes and reduced its influence on motivation and sustained interest.[35]

Failure avoidance was higher among INSPIRE-AI users, suggesting that students with greater performance concerns may have been more likely to use the tool as a safety net. One possible explanation is that students with greater concerns about academic performance preferentially adopted INSPIRE-AI as an additional source of support. Because platform use was self-selected, this pattern should be interpreted cautiously, and reverse causality cannot be excluded. Students who preferred alternative large language models offering more direct responses may have been less inclined to engage with scaffolded learning processes, although this interpretation remains speculative and warrants further investigation.[36] These findings underscore the need to design AI-facilitated interventions that promote critical engagement rather than task completion alone.

Students' engagement with INSPIRE-AI was shaped by the broader research training ecosystem, including supervisory availability, peer networks, and institutional expectations. These observations highlight the co-determination of learning outcomes by both pedagogical design and implementation context, emphasizing that AI tools alone cannot guarantee improved learning outcomes. Rather, students' interpretations of INSPIRE-AI's legitimacy, safety, and usefulness relative to existing sources of support appeared to shape how they engaged with the tool.[37]

Limitations and Strengths

Several limitations should be considered. The one-group pretest/posttest design precludes causal attribution, as students were simultaneously exposed to multiple honors program supports, making it difficult to isolate the intervention's independent effects. Study participation and INSPIRE-AI uptake were self-selected, introducing

potential selection bias, confounding by learner characteristics, and reverse causality in user/nonuser comparisons. The single-school context and modest sample size limit transferability and power for subgroup analyses. Outcomes relied primarily on self-report measures, and follow-up was limited to the honors year.

Despite these limitations, the embedded mixed-methods design strengthened interpretation by linking outcome changes with students' experiences within the broader research training ecosystem. The study also contributes early evidence on a theory-informed, retrieval-augmented AI teaching assistant for undergraduate nursing research education. A further strength of this study is the transparent reporting of the intervention using the TIDieR checklist together with detailed descriptions of chatbot pathways, prompt engineering, and RAG resources, facilitating reproducibility and future adaptation by nurse educators.

Implications for Nursing Education and Future Research

AI-enabled learning support should be integrated as formative scaffolds within existing educational infrastructures rather than as replacements for supervisors or established learning resources. Successful implementation requires clear guidance on appropriate AI use, transparent governance regarding data access, monitoring, and assessment practices, and open communication that fosters psychological safety by clarifying how student interactions are managed. Alignment with institutional academic integrity policies, assessment expectations, and transparent AI governance is essential to promote trust, engagement, and responsible AI use. Likewise, educator guidance should reinforce AI-supported tools as formative learning resources that support exploration and iterative learning, rather than tools that evaluate students' interactions or performance, to preserve psychological safety. Future research should employ larger comparative and longitudinal designs to evaluate the sustained effects, objective academic outcomes, and implementation conditions influencing uptake and educational value.

Conclusions

This study provides preliminary evidence that a theory-informed AI-facilitated teaching assistant can support undergraduate nursing research learning, particularly by strengthening research self-efficacy through structured scaffolding. However, its value depended on uptake, trust, perceived academic risk, usability, and the broader learning environment. Future research should employ comparative and longitudinal designs that incorporate objective educational outcomes and implementation evaluations to determine effectiveness, identify mechanisms of action, and support broader adoption across nursing education settings.

References

1. Liang BJ, Chua JYX, Shorey S. Perceptions of nurses and nursing students regarding evidence-based nursing: a qualitative systematic review. *J Adv Nurs*. 2026;82(4):2846-2862. doi:10.1111/jan.70087
2. Betzer C, Larsen NS, Larsen K. Diversity, contradictions, and distinction – a scoping review of honours programmes in nursing. *J Prof Nurs*. 2023;49:116-125. doi:10.1016/j.profnurs.2023.09.010
3. Shipton EV, Kearney L, Delaney L, Kilgour C. Midwifery and nursing honours programs in Australia and New Zealand: a scoping review of contemporary challenges and opportunities. *Women Birth*. 2024;37(5):101641. doi:10.1016/j.wombi.2024.101641
4. Bandura A. Social cognitive theory: an agentic perspective. *Annu Rev Psychol*. 2001;52(1):1-26. doi:10.1146/annurev.psych.52.1.1
5. Adedokun OA, Bessenbacher AB, Parker LC, Kirkham LL, Burgess WD. Research skills and STEM undergraduate research students' aspirations for research careers: mediating effects of research self-efficacy. *J Res Sci Teach*. 2013;50(8):940-951. doi:10.1002/tea.21102
6. Davis SN, Wagner SE. Research motivations and undergraduate researchers' disciplinary identity. *Sage Open*. 2019;9(3):2158244019861501. doi:10.1177/2158244019861501
7. Smith CR, Martsolf DS, Draucker CB, Shambley-Ebron DZ, Pritchard TJ, Maler J. Stimulating research interest and ambitions in undergraduate nursing students: the research-doctorate pipeline initiative. *J Nurs Educ*. 2016;55(3):133-140. doi:10.3928/01484834-20160216-03
8. Lambie GW, Hayes BG, Griffith C, Limberg D, Mullen PR. An exploratory investigation of the research self-efficacy, interest in research, and research knowledge of Ph.D. in education students. *Innov High Educ*. 2013;39(2):139-153. doi:10.1007/s10755-013-9264-1
9. Sasson I, Miedijensky S. Research performance: a view of research self-efficacy, interest, and gender. *Education Sciences*. 2023;13(12):1166. doi:10.3390/educsci13121166
10. Go J, Lee Y. The effectiveness of AI-based conversational agents in nursing education: a systematic review. *Nurse Educ Pract*. 2026;91:104728. doi:10.1016/j.nepr.2026.104728
11. Ma J, Wen J, Qiu Y, et al. The role of artificial intelligence in shaping nursing education: a comprehensive systematic review. *Nurse Educ Pract*. 2025;84:104345. doi:10.1016/j.nepr.2025.104345
12. Ivankova NV, Creswell JW. Mixed methods. In: Ivankova NV, Creswell JW, ed. *Qualitative Research in Applied Linguistics: A Practical Introduction*. Palgrave Macmillan; 2009:135-161.
13. Tovin MM, Wormley ME. Systematic development of standards for mixed methods reporting in rehabilitation health sciences research. *Phys Ther*. 2023;103(11):pzad084. doi:10.1093/ptj/pzad084
14. Ng JQX, Chua JS, Ang WHD, et al. Redesigned final year research program for undergraduate nursing students: a mixed-methods study. *Nurse Educ Today*. 2025;147:106620. doi:10.1016/j.nedt.2025.106620
15. Chaiklin S. The zone of proximal development in Vygotsky's analysis of learning and instruction. In: Kozulin A, Gindis B, Ageyev VS, Miller SM, ed. *Vygotsky's Educational Theory in Cultural Context*. Cambridge University Press; 2003:39-64.
16. Krathwohl DR. A revision of Bloom's taxonomy: an overview. *Theory Pract*. 2002;41(4):212-218. doi:10.1207/s15430421tip4104_2
17. Martin AJ. *Using Load Reduction Instruction (LRI) to Boost Motivation and Engagement*. British Psychological Society Leicester; 2016.
18. Kong W, Ning Y, Ma T, et al. Experience of undergraduate nursing students participating in artificial intelligence + project task driven learning at different stages: a qualitative study. *BMC Nurs*. 2024;23(1):314. doi:10.1186/s12912-024-01982-1
19. Jeon SJ, Yoo HH. Changes in medical students' research-related perceptions through student-engaged medical research

curriculum experience. *BMC Med Educ*. 2024;24(1):1002. doi:10.1186/s12909-024-06003-y
20. Li S, Huang J, Hussain S, Dong Y. How does supervisor support impact Chinese graduate students' research creativity through research self-efficacy and intrinsic motivation? A multi-group analysis. *Think Skills Creat*. 2025;55:101700. doi:10.1016/j.tsc.2024.101700
21. Liu X, Liu Y, Dai Y, Fu J. Academic stress and university students' dependency on generative artificial intelligence: a multiple mediation model using PLS-SEM. *BMC Psychol*. 2026;14(1):216. doi:10.1186/s40359-026-03986-9
22. Tao KW, Gloria AM. Should I stay or should I go? The role of impostorism in STEM persistence. *Psychol Women Q*. 2018;43(2):151-164. doi:10.1177/0361684318802333
23. Brancolini KR, Kennedy MR. The development and use of a research self-efficacy scale to assess the effectiveness of a research training program for academic librarians. *LIR*. 2017;41(124):44-84. doi:10.29173/lirg760
24. Deemer ED, Martens MP, Buboltz WC. Toward a tripartite model of research motivation: development and initial validation of the Research Motivation Scale. *J Career Assess*. 2010;18(3):292-309. doi:10.1177/1069072710364794
25. Bishop R, Bieschke K. *Interest in Research Questionnaire*. Unpublished Scale; 1994.
26. Kahn JH, Miller SA. Measuring global perceptions of the research training environment using a short form of the RTES-R. *Meas Eval Couns Dev*. 2000;33(2):103-119. doi:10.1080/07481756.2000.12069001
27. Cassady JC. *Academic Anxiety Scale*. Academic Anxiety Resource Center. Accessed June 28, 2026. https://sites.bsu.edu/aarc/research/academic-anxiety-scale/
28. Bujang MA, Baharum N. Sample size guideline for correlation analysis. *WJSSR*. 2016;3(1):37-46. doi:10.22158/wjssr.v3n1p37
29. Clarke V, Braun V. Thematic analysis. *J Posit Psychol*. 2016;12(3):297-298. doi:10.1080/17439760.2016.1262613
30. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qual Res Sport Exerc Health*. 2019;13(2):201-216. doi:10.1080/2159676x.2019.1704846
31. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis. *Int J Qual Methods*. 2017;16(1):1609406917733847. doi:10.1177/1609406917733847
32. Amankwaa L. Creating protocols for trustworthiness in qualitative research. *J Cult Divers*. 2016;23(3):121-127.
33. Meyer K, Willis R. Looking back to move forward: the value of reflexive journaling for novice researchers. *J Gerontol Soc Work*. 2019;62(5):578-585. doi:10.1080/01634372.2018.1559906
34. Younas A, Fàbregues S, Munce S, Creswell JW. Framework for types of metainferences in mixed methods research. *BMC Med Res Methodol*. 2025;25(1):18. doi:10.1186/s12874-025-02475-8
35. Baumeister RF, Vohs KD. Self-regulation, ego depletion, and motivation. *Soci Personality Psych*. 2007;1(1):115-128. doi:10.1111/j.1751-9004.2007.00001.x
36. Wang KD, Wu Z, Tufts L, Wieman C, Salehi S, Haber N. Scaffold or crutch? Examining college students' use and views of generative AI tools for STEM education. *arXiv*. 2025:1-10.
37. Attard-Frost B, Lyons K. AI governance systems: a multi-scale analysis framework, empirical findings, and future directions. *AI Ethics*. 2024;5(3):2557-2604. doi:10.1007/s43681-024-00569-5