Co-Designing AI-Enabled Learning in Nursing Education: A Learning Engineering Approach Using Faculty and Student Insights

Shah, M., Calacci, M., Lander, L., & Coochwytewa, C.

Academic-Industry Partnerships   Generative AI   Human-centered Design

learning engineering   Nursing Education   Technology Integration

Generative AI tools are rapidly entering nursing education. Yet little empirical work examines how faculty and students actually use these tools in authentic learning contexts or what design requirements emerge from their lived experiences. This practitioner-oriented study investigates an academic-industry partnership through the adoption of Sherpath AI (SPAI), a generative AI chatbot for nursing education, across 45 prelicensure course sections at Arizona State University's Edson College of Nursing and Health Innovation. Through semi-structured interviews with seven faculty members and focus groups with ten students, we identified patterns in AI appropriation for studying, clinical reasoning, and pedagogical design. Drawing on learning engineering (LE) frameworks emphasizing human-centered, iterative, and contextually grounded design, we surface key themes: AI as a cognitive and

pedagogical partner, trust calibration through evidence pathways, prompting as a metacognitive scaffold, and academic-industry learning loops. Findings reveal opportunities - personalized explanations, confidence building, and instructional coherence - and priorities, including verification pathways in specialized domains and scaffolded prompting support. We conclude with empirically derived design requirements for scalable, equitable, and trustworthy AI-enabled learning systems in safety-critical professional education.

Introduction

Generative AI has infiltrated nursing education at an unprecedented pace, promising to personalize learning, reduce faculty workload, and prepare students for technology-enabled practice. Yet, this rapid integration occurs within a high-stakes professional domain where conceptual, ethical, and procedural knowledge must converge to ensure safe, quality patient-centered care. Nurse educators face a precarious challenge: how to harness AI's potential for explanation, practice, and reasoning support while maintaining the rigorous standards required for clinical competence, professional judgment, and ethical practice.

This challenge is amplified by several factors unique to nursing education. First, the discipline requires integration across multiple knowledge areas, including pathophysiology, pharmacology, clinical procedures, ethical reasoning, and patient-centered communication— all of which must be applied in rapidly changing, high-pressure clinical contexts (Benner, Sutphen, Leonard, & Day, 2010). Second, nursing education operates under strict accreditation standards and licensure requirements, including the National Council Licensure Examination (NCLEX), which creates tension between pedagogical innovation and regulatory compliance. Third, the profession is experiencing severe faculty shortages and growing student enrollments, intensifying workload pressures while simultaneously demanding high-quality, individualized instruction (American Association of Colleges of Nursing, 2023).

Against this backdrop, generative AI tools offer both promise and peril. They can provide personalized cognitive scaffolding, generate practice materials aligned to learning objectives, and support self-directed study. However, they also raise concerns about accuracy in specialized clinical contexts, as well as overreliance that could undermine the development of clinical reasoning, academic integrity, and equal access. Most critically, there exists little empirical evidence about how nursing faculty and students use these tools in authentic learning contexts, the breakdowns they encounter, and what design features would support effective and safe integration.

Literature Review

Learning engineering (LE) provides a principled approach to addressing these challenges. Mishra (2025) frames LE as a human-centered, interdisciplinary, iterative, and values-driven practice that bridges learning sciences, instructional design, data analytics, and domain expertise. Rather than viewing technology adoption as the implementation of pre-designed tools, Kolodner (2023) suggested LE conceptualizes it as ongoing design-in-use, in which practitioners surface breakdowns and refinement opportunities through authentic activity.

Additionally, Saxberg (2017) emphasizes that learning innovations succeed only when grounded in cognitive task design; that is, understanding the specific knowledge structures, reasoning patterns, and performance demands of a domain. For nursing, this means attending to how students develop clinical reasoning (hypothesis generation, evidence evaluation, decision-making under uncertainty), integrate theoretical knowledge in clinical situations, and form professional identity. Goodell, Kolodner, and Kessler (2022) extend this by arguing that LE requires iterative cycles of design → implementation → data collection → refinement, with close collaboration among educators, learners, researchers, and developers.

Mishra, Warr, and Islam (2023) describe generative AI as protean (taking many forms), opaque (lacking transparency in reasoning), unstable (outputs varying across iterations), and social (shaped by and shaping human interaction). These characteristics demand that educators develop what they term a "teaching compass" or competencies: agency in shaping AI use, attention to well-being and workload sustainability, and knowledge, skills, values, and ethics for navigating AI-enabled systems (Mishra & Oster, 2023). These framing positions educators not as passive adopters but as active designers who must integrate technical, pedagogical, and contextual knowledge across learning strategies.

Recent systematic reviews provide mixed evidence on AI's impact on nursing education. Ma et al. (2025) analyzed 47 studies and found that AI tools can enhance knowledge acquisition, self-directed learning, positive attitudes toward technology, and simulation performance. However, the effects on clinical reasoning and critical thinking were inconsistent. Similarly, Ramírez-Baraldes and colleagues (2025) identified opportunities for AI to support personalized learning pathways and reduce faculty administrative burden, while highlighting the risks related to accuracy, bias, academic integrity, and overreliance on the output.

Notably, most existing research examines AI-enabled simulations, intelligent tutoring systems, or predictive analytics for student success, not generative AI chatbots for open-ended learning support. The handful of studies on generative AI in nursing education (predominantly focused on ChatGPT) reveal faculty concerns about accuracy in specialized domains like pediatric dosing calculations, potential for students to bypass critical thinking. Moreover, by uncritically accepting AI-generated outputs, educators may struggle to determine whether students have developed true conceptual understanding versus only surface-level familiarity (Chan & Lee, 2024; Phillips & Lee, 2024).

Crucially, existing literature lacks empirical investigation on how nursing faculty and students appropriate generative AI tools in authentic course contexts, what strategies they develop, how they calibrate trust, and what design features would support rather than undermine professional competence development. This gap is particularly significant given nursing education's emphasis on situated cognition, the principle that knowledge must be developed and demonstrated in contexts approximating clinical practice (Lave & Wenger, 1991).

Research Questions

This study addresses these gaps by examining early adoption of Sherpath AI (SPAI), Elsevier's generative AI chatbot for nursing education, at Arizona State University's Edson College of Nursing and Health Innovation. We ask:

1. How do nursing students and faculty use generative AI in authentic learning contexts across diverse course types?
2. What opportunities, constraints, and risks do they identify through their lived experiences?
3. How can their insights inform learning-engineered design requirements for future AI-enabled learning systems in professional education?

By positioning faculty and students as co-designers whose everyday use reveals design requirements, we demonstrate how learning engineering principles translates into specific features for AI-enabled learning systems in safety-critical domains.

Methods

Context and Participants

Arizona State University's Edson College of Nursing and Health Innovation, a large public university serving a diverse student body, adopted Sherpath AI (SPAI) in Summer 2024 as an optional learning resource. SPAI is a generative AI chatbot designed for nursing education, featuring evidence-linked citations from Elsevier's nursing textbooks (Lewis's Medical-Surgical Nursing, Hockenberry's Wong's Nursing Care of Infants and Children), competency alignment to nursing curricula, and transparent citations to authoritative sources. SPAI was available to all participants via their course learning management system (Canvas). In-site (just-in-time) onboarding and on-demand support were provided by co-authors at ASU. Data were collected from nursing faculty (N = 7) and students (N = 10) who used SPAI across 45 prelicensure nursing course sections (Summer 2024–Fall 2025). Faculty participants taught courses such as pediatric nursing, critical care, nursing research foundations, and experiential/clinical coordination. Student participants spanned second through fourth year with varying engagement levels and prior AI experience.

Data Collection and Analysis

Following IRB approval (#00022071), the research team conducted semi-structured faculty interviews (individual, 45-60 minutes) and student focus groups (3 groups of 3-4 students, 60 minutes) between Fall 2024 and Fall 2025. Semi-structured protocols allowed flexibility to pursue emergent themes while ensuring systematic coverage of key domains: (1) awareness and implementation; i.e., how participants described and integrated SPAI across courses; (2) use patterns; i.e., types of prompts, workflows, and refinement strategies; (3) impact; ie., specific examples of learning/teaching support; (4) trust and quality; i.e., evaluation of accuracy, verification practices, and factors influencing confidence; (5) barriers and improvements; i.e., challenges encountered and desired features; and (6) additive value; i.e., comparison to traditional resources and influence on practice. Sessions were audio-recorded, transcribed verbatim, and member-checked.   We employed thematic analysis (Braun & Clarke, 2006) to identify patterns in SPAI appropriation, trust reasoning, prompting strategies, and challenges. Four themes emerged: (1) AI as a cognitive and pedagogical partner, (2) trust calibration and evidence pathways, (3) prompting as a learning engineering process, and (4) academic-industry learning loops. Analysis attended to 'breakdowns' revealing design requirements (Kolodner, 2023) and 'teaching compass' competencies (Mishra & Oster, 2023).

Results

AI as a Cognitive and Pedagogical Partner

Students consistently used SPAI to deconstruct complex concepts, compare diagnoses, and rehearse NCLEX-style reasoning. One student explained, "Sometimes I just need it to explain something at my level—like third semester—not a whole textbook chapter." This personalization was particularly valued in courses where textbooks provided comprehensive but overwhelming detail. Another student noted that SPAI helped build confidence before class participation: "I don't want to ask a dumb question in front of everyone. I [will] try it in Sherpath AI first." Faculty appropriated SPAI for pedagogical design. They used it to generate case studies, create NCLEX-style practice questions with appropriate distractors, and scaffold active learning activities. One instructor shared, "I used it to check that my explanations matched what's in the textbook and what students would see on their exams." SPAI also supported pedagogical coherence across instructors by offering standardized, evidence-linked explanations.

Trust Calibration, Accuracy, and Evidence Pathways

Participants consistently emphasized SPAI's evidence-linked citations as the primary source of trust, distinguishing it from general-purpose AI tools. A faculty member noted, "If I see Lewis or Hockenberry cited, I know where it's coming from." Students echoed this sentiment: "I check the evidence tab every time. ChatGPT never tells me where it gets anything."

This transparency enabled what we term "trust calibration," where participants developed nuanced judgments about when to rely on SPAI and when additional verification was needed. Yet, faculty also identified limits, especially in pediatrics: "Sometimes it gets the nuance wrong, so I double-check against clinical practice." These tensions reflect nursing education's need for clinically aligned, continuously updated content, aligned with Mishra's principle (2025) that AI systems inevitably embed trade-offs requiring human judgment and oversight.

Prompting as a Learning Engineering Process

Early use revealed prompting challenges: unclear requests, overly broad queries, and outputs mismatched to the learner's level. Through experimentation, students and faculty learned to refine prompts: "Explain this again but simpler," "Give me distractors a second-semester student would choose," "Create a two-part case study with increasing complexity." This iterative prompting became a metacognitive scaffold. As one student admitted, "I realized if I can't ask a good question, it's because I don't understand what I'm stuck on." Such iterative refinement exemplifies LE cycles of design → feedback → redesign, and points to design needs such as built-in level controls, clarifying questions, suggested follow-up questions, and competency alignment.

Academic–Industry Learning Loops

Faculty identified additional needs: multimodal learning assets (audio summaries, diagrams, side-by-side comparisons), competency-aligned question generation, instructor insight filters, and prompt exemplars. Notably, SPAI functioned as what Star and Griesemer (1989) term a "boundary object," i.e., a shared referent that faculty, clinical partners, and students used to triangulate conceptual understanding and align instruction across classroom and clinical settings. One instructor noted, "I take the Sherpath AI-generated STEMI summary to my ICU colleagues and ask, 'Is this what you're seeing clinically?'" These patterns reflect Kolodner's (2023) claim that LE is fundamentally collaborative and that effective systems emerge from distributed expertise, not tool adoption alone.

Discussion and Conclusions

Sherpath AI adoption at ASU unfolded as a LE process: faculty and students iteratively adapted the tool, reflected on mismatches, and reworked instructional strategies. This aligns with Mishra, Warr, and Islam's (2023) arguments that educators must integrate technical, pedagogical, and contextual knowledge within AI-mediated learning systems. It also resonates with Kolodner's (2023) framing of LE as design-in-use, where authentic activity reveals breakdowns that guide redesign. Findings also mirror nursing literature: AI can enhance self-directed learning, confidence, and foundational knowledge, but requires safeguards to support clinical reasoning, ethical judgment, and professional identity formation (Ma et al., 2025). The ASU– Elsevier partnership demonstrates how academic–industry collaboration can enable principled, data-informed, human-centered innovation. As Saxberg (2017) notes, scaling effective learning requires structures for exposure, education, effort, and evaluation. These structures emerged organically through this partnership.

Three insights emerged from this work. First, generative AI tools serve different functions for different users in different contexts: students sought cognitive scaffolding while faculty sought pedagogical design support. Second, trust in AI is not binary but contextual and calibrated, requiring transparent evidence pathways and verification opportunities. Third, knowing what to ask, when to use the tool, and how it supports learning is determined through practice and highlights iterative design needs for scaffolded interfaces that support users at varying expertise levels.

This practitioner-oriented study also has several takeaways for specific audiences. For nurse educators and programs, GenAI can be positioned as a cognitive partner that supports explanation, comparison, and rehearsal without replacing human judgment. Embedding AI literacy (prompting, verification, ethical use) into curricula strengthens both professional formation

and safe clinical reasoning. For academic institutions, AI adoption can function as an infrastructure for learning engineering, where faculty and student feedback regularly informs tool refinement. Intentional structures such as policies, training, and course-level experimentation are needed to support effective and ethical GenAI use in professional programs. For industry partners, faculty, and students act as co-designers and end-users. Academic−industry partnerships should be structured as ongoing learning loops feeding authentic insight into product evolution. For learning engineers, nursing education demonstrates how LE principles such as cognitive alignment, iterative design, contextual constraints, and human-centered values translate into specific AI features such as level controls, competency tagging, and trust signals. This context underscores the need for LE teams to integrate learning sciences, domain expertise, data, and ethics in safety-critical fields.

References

American Association of Colleges of Nursing. (2023). Fact sheet: Nursing faculty shortage. https://www.aacnnursing.org/news-data/fact-sheets

Benner, P., Sutphen, M., Leonard, V., & Day, L. (2010). Educating nurses: A call for radical

transformation. Jossey-Bass.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77-101.

Chan, C. K. Y., & Lee, K. K. W. (2024). Generative AI in nursing education: Opportunities and challenges. Nurse Education Today, 134, 106089.

Goodell, J., Kolodner, J., & Kessler, A. (2022). Learning engineering applies the learning sciences. In Learning engineering toolkit (pp. 47-82). Routledge.

Kolodner, J. L. (2023). Learning engineering: What it is, why I'm involved, and why I think more of you should be. Journal of the Learning Sciences, 32(2), 305-323.

Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. Cambridge University Press.

Ma, J., Wen, J., Qiu, Y., Wang, Y., Xiao, Q., Liu, T., ... & Sun, Z. (2025). The role of

artificial intelligence in shaping nursing education: A comprehensive systematic

review. Nurse Education in Practice, 104345.

Mishra, P. (May, 2025). From artifacts to systems: The evolution of learning engineering.

Keynote. International Education Forum on Learning Engineering. Ankara,

Turkey.

Mishra, P., Warr, M., & Islam, R. (2023). TPACK in the age of ChatGPT and Generative

AI. Journal of Digital Learning in Teacher Education, 39(4), 235-251.

Phillips, A., & Lee, S. (2024). ChatGPT in nursing education: Faculty perspectives on risks and opportunities. Journal of Nursing Education, 63(3), 156-162.

Ramírez-Baraldes, E. L., García-Gutiérrez, D., & García-Salido, C. (2025). Artificial intelligence in nursing: New opportunities and challenges. European Journal of Education, 60(1), e70033.

Saxberg, B. (2017, April). Learning engineering: the art of applying learning science at

scale. In Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale

(pp. 1-1).

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. Social Studies of Science, 19(3), 387-420.

★★★★★