

Enhancing Team-Based Medical Simulations: Learning Through Reflection with Analytics and AI Tools

Crina Damsa (chair), University of Oslo, Norway, crina.damsa@iped.uio.no
 Vanessa Echeverria (co-chair), Monash University, Australia, Vanessa.Echeverria@monash.edu
 Vitaliy Popov (co-chair), University of Michigan, vipopov@umich.edu
 Olga Chernikova (co-chair), Ludwig-Maximilian-University, Germany, o.chernikova@psy.lmu.de
 Klas Karlgren (Discussant), Karolinska Institute, Sweden, klas.karlgrén@ki.se
 Mikaela Milesi, Monash University, Australia, Mikaela.Milesi@monash.edu
 Ariel Jin, Monash University, Australia, Ariel.Jin@monash.edu
 Linxuan Zhao, Monash University, Australia, Linxuan.Zhao@monash.edu
 Riordan Alfredo, Monash University, Australia, Riordan.Alfredo@monash.edu
 Lixiang Yan, Monash University, Australia, Lixiang.Yan@monash.edu
 Dragan Gasevic, Monash University, Australia, Dragan.Gasevic@monash.edu
 Roberto Martinez-Maldonado, Monash University, Roberto.Martinez-Maldonado@monash.edu
 Yaxuan Li, University of Michigan, yaxuanli@umich.edu
 Kunpeng Huang, University of Michigan, kunpeng@umich.edu
 Alanson Sample, University of Michigan, apsample@umich.edu
 Bryan Harmer, University of Michigan, edubryharne@med.umich.edu
 James Cooke, University of Michigan, cookej@med.umich.edu
 Michael Cole, University of Michigan, mcolemd@med.umich.edu
 Xuechen Zhang, Ludwig-Maximilian-University, Germany, X.Zhang@campus.lmu.de
 Frank Fischer, Ludwig-Maximilian-University, Germany, Frank.Fischer@psy.lmu.de
 Meltem Akoyak-Yildiz, AMRITA University, India, meltem.alkoyak@ammachilabs.org
 Daniel Sanchez, University of Oslo, daniel.sanchez@iped.uio.no
 Rachelle Esterhazy, University of Oslo, rachelle.esterhazy@oslomet.no
 Hugh Alexander von Arnim, University of Oslo, hughav@imv.uio.no
 Lena Günterberg Heyn, lena.heyne@usn.no
 Edda Aslaug Johansen, edda.johansen@usn.no
 Susanne Olesen, susanne.olesen@usn.no
 Ellen Wesenberg Larsen, ellen.w.larsen@usn.no
 Carmen Louwerens, carmen.louwerens@usn.no
 Hilde Eide, hilde.eide@usn.no

Abstract: Simulation-based learning has become a cornerstone of healthcare education, fostering essential skills like communication, teamwork or decision-making in safe, controlled environments. However, participants' reflection on simulations often rely on subjective recollections, limiting their effectiveness in promoting learning. This symposium explores how multimodal analytics and AI can enhance simulation-based education by automating teamwork analysis data, providing structured feedback, and supporting reflective practices. The papers examine real-time analytics for closed-loop communication in cardiac arrest simulations, multimodal data use to refine feedback in ICU nursing simulations, generative AI-powered chatbots facilitating nursing students' interpretation of multimodal learning analytics dashboards, and culturally sensitive, AI-based scenarios for Breaking Bad News in an Indian context. Collectively, these contributions highlight the transformative potential of using data and AI-enhanced solutions, emphasizing personalization, cultural sensitivity, and human-centered design, and invite dialogue on the pedagogical, technological and ethical implications of introducing data-based practices and AI-based tools in medical education.

Introduction

Current teaching and training methods in the healthcare sector, such as medical and nursing training, often involve (team-based) simulations, followed by reflection sessions in which participants analyze their individual and teamwork performance. This type of training allows learners to practice their teamwork or collaborative skills, such as communication, teamwork, leadership, and decision-making, which are essential for providing safe and effective patient care.

Team-based simulation involves groups of learners working together in a simulated environment to manage a clinical scenario and can be used to train both students and healthcare professionals, including nurses, physicians, and emergency medical technicians. Teamwork can encourage individuals to externalise their personal perspectives to reach a shared understanding (Häkkinen, et al., 2017). Studies have demonstrated the positive effects of teamwork, such as increased knowledge of complex pathologies and patient profiles, problem identification, clinical reasoning, coordination, and communication skills (Considine et al., 2020). The scenarios used in team-based simulation are often complex and challenging, and they are designed to mimic real-life situations as closely as possible. Team-based simulation offers a safe and controlled environment for learners to practice their skills and make mistakes without putting patients at risk (Ballangrud et al., 2014). The fidelity of the simulation can vary, from low-fidelity simulations using simple manikins to high-fidelity simulations using sophisticated computer-controlled simulators.

Regardless of the fidelity, the key to effective team-based simulation is the debriefing session that follows the simulation. Reflection after simulation sessions is valuable in learning patient-care behaviour and stimulating knowledge consolidation but often uses limited data (e.g., video recordings). During the debriefing, learners reflect on their performance, identify areas for improvement, and receive feedback from their peers and instructors. Effective reflection, grounded in objective recollections of (team) activities, is crucial for students' learning and professionals' development, especially in environments where collaboration is key. Stressful teamwork situations, cognitive dissonance or communication and coordination challenges can lead to errors and poor recollection of collaborative efforts, obstructing meaningful reflection and learning from valuable experiences from the simulations. While viewed as highly valuable for learning, reflection sessions are frequently compromised by their reliance on participants' subjective recollections and the fact that it is difficult to use data from the simulation as an objective display of the performance under the simulation.

Collocated teamwork and mediated by digital technologies form an effective strategy for team skills training. While there is abundant evidence of simulation and teamwork's importance in learning, ways of automating measures of teamwork, automated feedback and support for reflection are scarce, especially in authentic settings (Echeverria, et al., 2024). Learning designs that embed multimodal analytics and AI-based tools offer opportunities to facilitate sensemaking of multimodal data related to such phenomena. These techniques and approaches allow the automatic or semi-automatic capturing, analysis, and distillation of data about collaborators' speech, actions, and interactions.

The contributions and their collective significance

The contributions in this symposium emphasize the importance understanding how communication and teamwork skills are being reflected on high-stakes environments, such as ICU simulations, resuscitation training, and Breaking Bad News (BBN) scenarios, and how different analytics and AI-powered designs can support this reflection. Each study also highlights the role of structured personalized feedback, whether through automated analytics, multimodal data, or AI-powered tools, in fostering reflective practices and skill development. The first contribution centers on closed-loop communication in VR-based cardiac arrest simulations, using NLP techniques to provide real-time analytics and actionable feedback. The second contribution introduces a GenAI-powered chatbot designed to guide students in interpreting multimodal learning analytics dashboards and reveals that the chatbot enhances reflective practices for nursing students, supporting constructive feedback and actionable improvements. The third explores cultural sensitivity and adaptive learning for BBN within Indian contexts, using Large Language Models (LLMs) to create realistic scenarios that address socio-demographic nuances. The fourth focuses on multimodal analytics in ICU nursing simulations, integrating physiological and behavioral data to enhance the precision and contextual relevance of feedback.

The symposium shares a common focus on using advanced technology, particularly different types of artificial intelligence (e.g., LLM, learning analytics), to enhance the potential of simulation-based learning in medical and nursing education. Unique aspects of symposium contributions emerge in the specific methodologies and empirical focus, creating the space for discussing best practices and challenges in training medical and healthcare professionals. Together, the contributions illustrate the transformative potential of AI and data-driven feedback in medical and healthcare education. They collectively underscore the importance of personalization, cultural sensitivity, and multimodal approaches to improve learner engagement, reflection, and skills in diverse medical training scenarios.

This symposium directly addresses the conference theme by exploring how data and AI-driven and simulation-based learning tools can transform healthcare education to meet global challenges. By using AI and multimodal analytics, the contributions demonstrate innovative, sustainable solutions to enhance education for healthcare professionals, equipping them with skills to navigate complex and culturally diverse scenarios.

Organization of the symposium

In our symposium, we will use innovative approaches and organize the symposium as follows. Before the symposium, we will create a symposium repository with curated resources. We produce 5-minute videos introducing each project. This will allow us to share the background of each study and other materials that can enrich in-person discussions during the symposium. We will distribute these resources 2-4 weeks before the symposium. The in-person symposium will start with a short co-chairs' introduction to the topic and an overview of the symposium (7 minutes). It will be followed by a 40-minutes session with pitches of each paper. The remaining time will be dedicated to the discussant's response synthesizing critical insights and posing follow-up discussion questions (10 minutes) and interactive discussion between the discussant, three presenters and the audience (20 minutes). The symposium will conclude with a short synthesis (3 minutes). The reminder of this text presents the four empirical contributions featured in this symposium.

Leveraging AI for automated evaluation of closed-loop communication in VR-based acute care team training

Vitaliy Popov, Yaxuan Li, Kunpeng Huang, Alanson Sample, Bryan Harmer, James Cooke, Michael Cole

Introduction

High-reliability teams need frequent training opportunities and precision feedback to develop complex skills such as communication (Mathieu et al., 2019; Baker et al., 2006). In healthcare specifically, resuscitation teams utilize closed-loop communication (CLC). Closed-loop communication (CLC) is a structured three-stage protocol that ensures clear information exchange and shared understanding among team members, vital for effective teamwork in high-reliability organizations including healthcare, ensuring that critical information is communicated and confirmed, reinforcing cohesive teamwork. Despite the widespread recognition of CLC's importance and its promotion by organizations like the American Heart Association (Cheng et al., 2020), our ability to systematically assess and improve CLC quality remains limited by subjective, non-standardized evaluation methods.

Simulation-based training is the most common method used to teach CLC to healthcare teams. However, evaluating CLC during these simulations relies heavily on subjective instructor observations, manual video reviews, or self-reported surveys, which limits the efficiency, precision of feedback, and scalability of this approach. The assessment of closed-loop communication (CLC) in clinical simulation training presents a complex Instructor Noticing challenge. Through the lens of Teacher Noticing framework (van Es & Sherin, 2002): 1) instructors must attend to multiple team interactions simultaneously, which limits their ability to notice and interpret CLC patterns (*attending* challenge); 2) lack of standardized frameworks and recall bias limit instructors' ability to draw consistent inferences about CLC quality (*interpreting* challenge); and 3) without systematic and timely CLC assessment, instructors struggle to create targeted learning opportunities that address specific communication gaps (*shaping* challenge). To address these challenges, our research goal was to develop and evaluate an automated system that support all three dimensions of teacher noticing: attending (by highlighting relevant CLC instances), interpreting (by providing an analytical framework for quality assessment), and shaping (by enabling rapid, targeted feedback based on observed patterns).

This study addresses two research questions: 1. How can an LLM-based CLC assessment be designed to provide targeted feedback? 2. How do instructors perceive and integrate LLM-based CLC assessment into their evaluation and debriefing practices?

Methods

We performed a three-part study that included: system development, validation, and evaluation. First, we developed a GPT-powered system (Generative Pre-trained Transformer, a class of NLP models developed by OpenAI) that detects, computes, and classifies CLC instances. Leveraging the GPT-4 large language model (LLM), the pipeline accepts transcripts of clinical training sessions as inputs, predicts CLC instances, and generates a report with selected examples of CLC and performance metrics. This approach effectively identifies call-outs, predicts corresponding check-backs, and classifies the quality of each instance based on the check-back's specificity. Second, we conducted a technical accuracy assessment by comparing the tool's predictions against human annotations. Three annotators with background in emergency medical services independently rated 108 good CLCs, 105 poor CLCs, and 102 unresponded CLCs. Performance metrics were calculated using standard information retrieval techniques, computing Precision, Recall, and Inter-rater Reliability for Call-Outs and Check-Backs to quantify the AI's detection accuracy and annotator agreement. Finally, we conducted an instructor evaluation study involving eleven cardiac arrest instructors from three training sites in the Midwest, United States. In 45-60-minute interviews, participants received a CLC overview and shared their current evaluation practices

before observing and assessing a 10-minute VR cardiac arrest simulation using their typical teaching methods. The participants then reviewed an automated CLC report generated by the tool and discussed how it influenced their evaluation of team communication quality and potential feedback approaches. Finally, participants reflected on the tool's impact on their noticing of team communication and provided feedback on its usefulness, usability, and potential improvements. We conducted thematic analysis of participant feedback, evaluating perceived accuracy and utility of our automated CLC assessment approach.

Results and discussion

Our findings demonstrate the potential of AI-assisted assessment tools to provide actionable feedback that can be used in simulation-based training. The proposed system achieved technical performance (78.90% precision for call-outs, 74.27% for check-backs), approaching human expert benchmarks with a 5% difference. The system's ability to quantify communication patterns and provide categorized examples of varying quality (good/poor/unresponded) effectively scaffolds instructors' assessment practices, which currently rely on unstructured note-taking and live observations. Instructors prefer using AI-generated insights as reference points rather than definitive assessments. They suggested adding visual charts describing CLC frequencies and options for viewing transcripts of all CLC instances.

These findings contribute to CSCL literature by demonstrating how automated assessment can support instructors and learners in complex collaborative learning scenarios. This is the initial step in developing and validating our approach. Future work will involve experimental studies and potentially expand to other critical care scenarios, such as trauma resuscitation and operating room procedures.

AI-Supported Reflection in Nursing Education: Student Perceptions of a GenAI-Powered Chatbot and Multimodal Teamwork Analytics Dashboards

Vanessa Echeverria, Mikaela Milesi, Ariel Jin, Linxuan Zhao, Riordan Alfredo, Lixiang Yan, Dragan Gasevic, Roberto Martinez-Maldonado

Background and aims

Team-based learning (TBL) in nursing education immerses students in realistic, simulated clinical environments, providing a safe space to develop critical skills such as communication, teamwork, and task management (Maran & Glavin, 2003). Reflection is essential in TBL, as healthcare professionals often face high-stakes situations requiring effective teamwork. Reflection helps students internalize and apply cognitive and social skills by identifying areas for improvement and delineating actions for future practice. This reflection typically occurs after the simulation (i.e., reflection-on-action, Schön, 1992), allowing students to assess their performance and its impact on patient outcomes. Enabling students with personal reflection spaces could enhance learning outcomes. To support this reflection, providing evidence of team dynamics from the simulation would be beneficial, as it may spark meaningful insights and discussions using this evidence.

Multimodal learning analytics dashboards (MMLADs) provide evidence and a comprehensive view of team performance and interactions. By capturing, analyzing, and visualizing students' learning traces and behaviors, MMLADs can facilitate meaningful reflection (Giannakos et al., 2022). When paired with GenAI-powered chatbots, these tools further enhance independent reflection by offering grounded explanations and improving interpretability (Okonkwo, 2021). However, as GenAI-powered chatbots are relatively new in educational contexts, their role in supporting reflection, particularly when integrated with MMLADs, remains largely unexplored. Building on prior work using MMLADs for reflection on teamwork (Authors, 2024), this study explores students' perceptions of using a GenAI-Chatbot to support the interpretation of MMLADs and facilitate reflection. Specifically, it addresses the question: *How do students perceive the role of the GenAI-Chatbot in facilitating reflection?*

Methods

The learning scenario was a team simulation where four students, role-playing as primary and secondary nurses, worked together to manage patient care. An AI-powered multimodal learning analytics (MMLA) system captured team dynamics using lapel microphones for audio and location tags to track movements and body orientation (Authors, 2024). We generated four visualizations and developed a contextualised *GenAI-Chatbot* to support the interpretation of visualisations and answer scenario-related prompts (Fig. 1). Of 134 third-year nursing students (33 teams), 23 voluntarily participated in post-hoc reflection sessions, which included exploration of the MMLAD and use of *GenAI-Chatbot*, and providing their perceptions using a Likert scale (1–5; strongly disagree - strongly agree; *GenAI-Chatbot could enhance my reflective practices*) and open-ended responses (*GenAI-Chatbot helped*

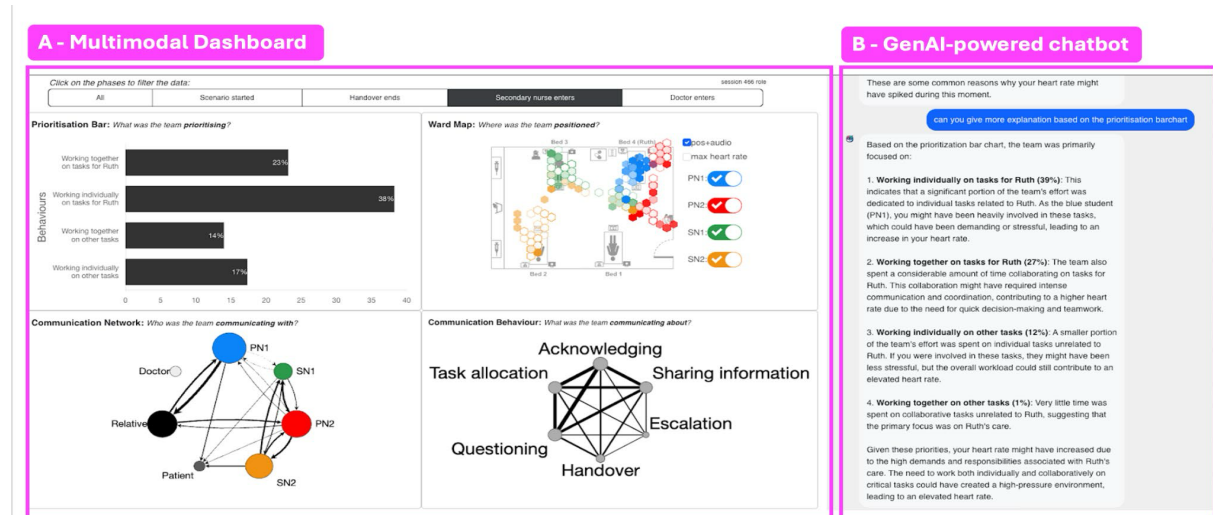
you to reflect on your team's performance and nursing practice? Why or why not?). Quantitative data were summarized with medians and percentages, while qualitative responses were analyzed using a thematic analysis approach and open-coding (Braun & Clarke, 2012) to identify emerging themes. Two key themes emerged: the *GenAI-Chatbot's role in facilitating reflection* and *limitations* in supporting reflective practices.

Findings

GenAI-Chatbot's role was perceived positively by many students, with 95% agreeing and strongly agreeing that it could enhance their reflective practices (Median = 5). Qualitative analysis revealed that 11 students found the *GenAI-Chatbot* **helpful for facilitating reflection**, with six emphasizing its ability to explain what went well, what went wrong, and why, shifting their perspective toward constructive reflection. Five students highlighted its usefulness in generating actionable suggestions for improvement, while three appreciated its role in interpreting visualizations and abstract concepts like "prioritization." However, four students expressed reservations, noting **limitations**, such as its inability to capture personal experiences, lack of response depth for detailed analysis, and challenges in framing meaningful questions without guidance.

Figure 1

Interactive MMLA Dashboard



Conclusion and implications

This study explored students' perceptions of using *GenAI-Chatbot* and MMLADs to support reflection on nursing practice. GenAI chatbots can offer support on reflective practices by providing actionable feedback and fostering constructive self-criticism. However, additional strategies and mechanisms are required to facilitate and promote deeper levels of reflection. This work represents a step toward scaling reflective practices in healthcare education through the integration of MMLADs and GenAI-powered solutions.

Empowering simulation-based learning in medical education with AI: designing a personalized and context sensitive tutor for Breaking Bad News

Olga Chernikova, Xuechen Zhang, Frank Fischer, Meltem Akoyak-Yildiz

Background and aims

Simulation-based learning represents a powerful pedagogical tool in higher education, particularly in medical training, as it provides controlled environments for practicing complex, ethically and culturally sensitive tasks like Breaking Bad News (BBN, see Mahendiran et al., 2023). It offers a practical alternative to real interactions with standardized patients, which are limited by ethical concerns, high costs, and resource-intensive training requirements. However, creating content for such simulations, enabling continuous practice opportunities remains a challenging task. The high incidence of fatal diseases, inequitable access to health care, and socioeconomic disparities in India, as well as high diversity in population regarding cultures, religions, and varying patient expectations, can complicate BBN formal training and require a lot of culturally specific knowledge and experience from the healthcare professionals (e.g., Martis & Westhues, 2013). Internationally recognized

communication protocols might introduce even more cultural dilemmas (e.g., Holmes & Illing, 2021). On the one hand, recent technological developments in artificial intelligence (AI) models might contribute to generating any amount of practice scenarios, offering a promising solution, but, on the other hand, many of these models might inherit and propagate gender and cultural biases, disadvantaging the minority population (Shrestha & Das, 2022).

In this study we aim at (1) contributing to basic fundamental research on cultural aspects of BBN in India to create information that can be used to train Large Language Models (LLM) in creating realistic BBN scenarios; and (2) identifying features of typical and atypical situations that can be encountered in BBN to address the possibility of introducing typicality scaffold in BBN simulations (see Fischer et al., 2022). Typical and atypical BBN scenarios are defined by patients' emotional (e.g., distress, denial or acceptance), cognitive (e.g., understanding and engagement in decision making), and socio-demographic (e.g., age, gender, religion, education) attributes, all of which shape patient expectations and communication preferences. These attributes are crucial for identifying culturally sensitive approaches in BBN (e.g., Foronda, 2008), ensuring that healthcare providers can address diverse patient needs effectively.

To address the goals of the study we rely on a professional conversation competence model (Gartmeier et al., 2015), focusing on establishing supportive relationships, advancing problem solutions, and structuring conversations transparently; and in-depth structured interviews with Indian doctors, who share their experience and observations. The interviews focus on patient emotional reactions, and individualized communication strategies, used by the doctors and therefore contribute to cultural sensitivity of BBN scenarios, promoting knowledge, understanding, and respect, which leads to effective communication. We explore two primary research questions: (1) What are the features of typical and atypical BBN settings in Indian context considering emotional, cognitive, and socio-demographic attributes. (2) To what extent can AI driven tutor prompted with expert information create realistic typical and atypical BBN scenarios and support developing learners' skills.

Methods

The study involved interviews with four experienced Indian physicians (three males, one female) from a South Indian multi-specialty tertiary care hospital, averaging 19 years of experience as physicians and 18.5 years in delivering bad news. Selected for their proficiency in English and expertise in BBN within Indian cultural context, these physicians provided insights into typical and atypical experiences with patients. Using thematic analysis (see Braun & Clarke, 2006), the study examines key themes in their communication practices (e.g., patient reactions), offering an in-depth look at the nuances of breaking bad news in an Indian medical context. The insights from the interviews were translated into prompts for an LLM and several typical and atypical scenarios were generated. The experts have evaluated quality and authenticity of typical and atypical scenarios created by LLM, which can be embedded in simulation-based learning environment.

Findings and implications

To outline the typical and atypical cases in BBN scenarios, research can focus not only on disease characteristics, but also on the emotional and cognitive patient dimensions. Many identified sub-themes (e.g., anxiety and distress, emotional outbursts, denial, suspicion, and aggressiveness) aligned with other research results on doctor-patient communication issues (e.g., Sander, 2010). This study first examined typical and atypical BBN scenarios in the Indian healthcare system. It then explored the use of LLMs for generating BBN scenarios. By interviewing experienced Indian physicians, we collected 9 scenarios capturing both typical and atypical BBN dimension confinements to inform the development of AI-generated training materials. We practiced a well-defined prompting strategy to train the model in creating similar scenarios to represent different cultural as well as cognitive and emotional attributes of the patients and reduce occurrences of gender or cultural bias. Additionally, this study discusses possible uses of obtained data to develop prompts for individualized feedback based on professional conversation competence model and SPIKES protocol, as well as prompts to support peer feedback on BBN performance in culturally sensitive context.

We conclude that AI-powered adaptive simulations offer a promising solution, allowing for personalized feedback and scenario progression that aligns with each learner's specific needs. The findings underscore a significant opportunity: AI-driven adaptive simulations specifically trained to address nuanced cultural contexts could set a new standard in simulation-based learning, particularly in medical education, by creating highly effective, culturally attuned training environments for sensitive communication tasks.

Sense-making through Reflection supported by Multimodal Dashboards in Team-based Simulations in Nursing Education

Daniel Sanchez, Crina Damsa, Rachele Esterhazy, Hugh von Arnim, Lena Heyn, Edda Johansen, Susanne Olesen, Ellen Larsen, Carmen Louwerens

Background and aims

Collocated teamwork, where participants share the same physical space while using advanced technologies to solve complex problems, remains prevalent across various domains. Hence, research about how automated feedback using teamwork analytics generated through computational means can assist students and teachers is crucial for good learning and effective guidance when applying knowledge into situations in practice (Cukurova, 2018). Simulations of teamwork offer a bridge between theory and practice, providing students with realistic environments to develop critical clinical skills. Reflection after these sessions is essential for learning patient-care behavior and consolidating knowledge but relies on limited data, such as video recordings (Verkuyt et al., 2020).

A significant gap exists in understanding how students and learners interpret and utilize feedback derived from diverse data sources presented in multimodal formats within these environments, as feedback's value lies in how well learners comprehend and apply it (Winstone et al., 2017). Although the research in this area is relatively small, it is growing. For example, Zhao et al. (2022) integrated positional and audio data in nursing simulations to distinguish team communications, enhancing our understanding of team dynamics under stress. Similarly, Yan et al. (2023) found that physiological synchrony and arousal offer insights into how stress levels affect team dynamics and individual performance, and a need to explore physiological measures in educational practices.

This study investigates the gap within the framework of deliberative practice (Ericsson, 2008). Deliberative practice in nursing simulations involves timely feedback, specific goal setting, and opportunities for repetitive exercises—key components essential in the high-stakes environment of ICU simulations, where nuanced understanding and skill application are critical. The central research question is: How do nursing students make sense of and use multimodal feedback in ICU simulations? This inquiry aims to deepen our understanding of feedback utilization in high-fidelity nursing education and training.

Conceptualizations and methods

This study was conducted with students enrolled in a master-level nursing course at a Norwegian university. Thirteen students were divided into three simulation teams to ensure optimal participation and engagement. Each participant was assigned a specific role such as Leading Nurse; Supporting Nurses; Handover Responsible; and Observer. Ethical approval and informed consent were obtained in accordance with GDPR and the university's research guidelines. After each simulation session, a debriefing was conducted where instructors facilitated student reflection, and critical thinking about their performance. Before the second simulation, students engaged in a data-based reflection session, using the multimodal inputs from the first simulation as discussion prompts. The data included video recordings, audio captured with unidirectional microphones, and physiological data using a medical-grade sensor (eq02+ LifeMonitor). The audio recordings were transcribed and coded in collaboration with an experienced nurse to identify moments when students applied a communication protocol (closed-loop communication). These moments were captured into video clips with captions and paired with coded transcripts. Heart rate and net acceleration were visualized as stress indicators, including pivotal moments (from manual log files) of the scenario. These visualizations also incorporated baseline heart rate estimates from pre-simulation data to identify potential stress moments during the scenario enactment.

Both the debriefing and reflection sessions were recorded, and a thematic analysis (Braun & Clarke, 2006), summarized in Table 1, was conducted on the transcripts. It identified and categorized meaningful instances concerning feedback. This comprehensive approach afforded a nuanced understanding of student interactions with and interpretations of the multimodal feedback in the reflection.

Findings

The findings reveal the nuanced reactions of nursing students to multimodal feedback in simulation, team-based learning environments. The data session demonstrated that students gained a deeper understanding and appreciation for precise, goal-oriented feedback, preferring actionable insights with clear directives for skill improvement. This actionable feedback was effectively integrated during the debriefing session, showing deliberate implementation by students in the second simulation. Moreover, evidence-based feedback highlighted discrepancies between students' self-assessments and actual performance, underscoring the dual role of multimodal feedback in reinforcing teacher observations and challenging students' self-perceptions.

Conclusion and implications

The results suggest that multimodal feedback, when aligned with specific performance goals, of individual and team and provided contextually, enhances students' comprehension and application of feedback. It also fosters the development of refined self-assessment capabilities, aligning with the principles of deliberative practice that emphasize timely, specific feedback and repeated practice for skill proficiency. Additionally, multimodal feedback can challenge and reshape students' self-perceptions, leading to more accurate self-and team-awareness. This study advocates for ongoing research to refine the integration of multimodal feedback within simulation-based learning, enhancing the feedback process and, ultimately, student learning outcomes.

References

- Ballangrud R., Hall-Lord M.L., et al. (2014). Intensive care nurses' perceptions of simulation-based team training for building patient safety in intensive care: a descriptive qualitative study. *Intensive Crit Care Nurs.* doi: 10.1016/j.iccn.2014.03.002.
- Baker, D. P., Day, R., & Salas, E. (2006). Teamwork as an essential component of high-reliability organizations. *Health services research, 41*(4p2), 1576-1598.
- Braun, V., & Clarke, V. (2012). Thematic analysis. In *APA handbook of research methods in Psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57–71).
- Cheng, A., Magid, D. J., Auerbach, et al. (2020). Part 6: resuscitation education science: 2020 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. *Circulation, 142*(16_Suppl_2),
- Considine J, Berry D, Sprogis SK, Newnham E, Fox K, Darzins P, Rawson H, Street M. Understanding the patient experience of early unplanned hospital readmission following acute care discharge: a qualitative descriptive study. *BMJ Open.* doi: 10.1136/bmjopen-2019-034728
- Cukurova, M., Luckin, R., Millán, E., & Mavrikis, M. (2018). The NISPI framework: Analyzing collaborative problem-solving from students' physical interactions. *Computers in Education, 116*, 93–109.
- Echeverria, V., Yang, K., Lawrence, L. E. M., Rummel, N., & Aleven, V. (2023). Designing hybrid human-AI orchestration tools for individual and collaborative activities: A technology probe study. *IEEE Transactions on Learning Technologies, 16*(2), 191-205. <https://doi.org/10.1109/TLT.2023.3248155>
- Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine, 15*(11), 988–994. <https://doi.org/10.1111/j.1553-2712.2008.00227.x>
- Fischer, F., Bauer, E., Seidel, T., et al. (2022). Representational scaffolding in digital simulations – learning professional practices in higher education. *Information and Learning Sciences, 123*(11/12), 645-665. <https://doi.org/10.1108/ILS-06-2022-0076>
- Foronda, C. (2008). A concept analysis of cultural sensitivity. *Journal of Transcultural Nursing, 19*, 207-212. <https://doi.org/10.1177/1043659608317093>
- Gartmeier, M., Bauer, J., Fischer, M.R., et al. (2015). Fostering professional communication skills of future physicians and teachers: Effects of e-learning with video cases and role-play. *Instructional Science, 43*, 443–462. <https://doi.org/10.1007/s11251-014-9341-6>
- Giannakos, M. N., Sharma, K., Pappas, I. O., et al. (2019). Multimodal data as a means to understand the learning experience. *International Journal of Information Management, 48*, 108–119.
- Häkkinen, P., Järvelä, S., Mäkitalo-Siegl, K., Ahonen, A., Näykki, P., & Valtonen, T. (2017). Preparing teacher-students for twenty-first-century learning practices (PREP 21): a framework for enhancing collaborative problem-solving and strategic learning skills. *Teachers and Teaching, 23*(1), 25–41. <https://doi.org/10.1080/13540602.2016.1203772>
- Holmes, S.N., & Illing, J. (2021). Breaking bad news: Tackling cultural dilemmas. *BMJ Supportive & Palliative Care, 11*, 128-132. <https://doi.org/10.1136/bmjspcare-2020-002700>
- Keiser, N. L., & Arthur Jr, W. (2021). A meta-analysis of the effectiveness of the after-action review (or debrief) and factors that influence its effectiveness. *Journal of Applied Psychology, 106*(7), 1007.
- LeBaron, V., Flickinger, T., Ling, D., et al. (2023). Feasibility and acceptability testing of CommSense: A novel communication technology to enhance health equity in clinician–patient interactions. *Digital Health, 9*, 20552076231184991.
- Mahendiran, M., Yeung, H., Rossi, S., et al. (2023). Evaluating the effectiveness of the SPIKES model to break bad news - A systematic review. *The American Journal of Hospice & Palliative Care, 40*(11), 1231–1260. <https://doi.org/10.1177/10499091221146296>
- Maran, N. J., & Glavin, R. J. (2003). Low-to high-fidelity simulation—a continuum of medical education?. *Medical Education, 37*, 22-28.
- Martis, L., & Westhues, A. (2013). A synthesis of the literature on breaking bad news or truth telling: Potential for research in India. *Indian Journal of Palliative Care, 19*(1), 2–11. <https://doi.org/10.4103/0973-1075.110215>
- Mathieu, J. E., Gallagher, P. T., Domingo, M. A., & Klock, E. A. (2019). Embracing complexity: Reviewing the past decade of team effectiveness research. *Annual Rev. of Organizational Psychology & Organizational Behavior, 6*(1), 17-46.
- Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence, 2*, 100033.
- Pedersen, I., Solevåg, A., & Solberg, M. T. (2019). Simulation-based training promotes higher levels of cognitive control in acute situations. *Clinical Simulation in Nursing, 34*, 6–15. <https://doi.org/10.1016/j.ecns.2019.05.003>

- Sanders, C. W. (2010). A brief overview of medical education in the United States. *Ochsner Journal*, 10(1), 38–40.
- Schön, D.A. (1992). *The Reflective Practitioner: How Professionals Think in Action* (1st ed.). Routledge.
- Shrestha, S., & Das, S. (2022). Exploring gender biases in ML and AI academic research through systematic literature review. *Frontiers in Artificial Intelligence*, 5. <https://doi.org/10.3389/frai.2022.976838>
- Van Es, E. A., & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, 10(4), 571-596.
- Verkuy1, M. et al. (2020). Combining self-debriefing and group debriefing in simulation. *Clinical Simulation Nursing*, 39, 41-44.
- Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist*, 52(1), 17–37. <https://doi.org/10.1080/00461520.2016.1207538>
- GPT-4 is OpenAI's most advanced system, producing safer and more useful responses. <https://openai.com/index/gpt-4/>. Accessed: May 24, 2024.