Received: 12 April 2022 | Accepted: 11 July 2022

DOI: 10.1111/bjet.13262

ORIGINAL ARTICLE

British Journal of Educational Technology BERA

The role of indoor positioning analytics in assessment of simulation-based learning

Lixiang Yan[1] | Roberto Martinez-Maldonado[1] | Linxuan Zhao[1] | Samantha Dix[2] | Hollie Jaggard[2] | Rosie Wotherspoon[2] | Xinyu Li[1] | Dragan Gašević[1]

[1]Centre for Learning Analytics at Monash, Faculty of Information Technology, Monash University, Clayton, Victoria, Australia

[2]Faculty of Medicine Nursing and Health Sciences, Monash University, Clayton, Victoria, Australia

Correspondence

Lixiang Yan, Centre for Learning Analytics at Monash, Faculty of Information Technology, Monash University, Clayton, VIC 3168, Australia.
Email: lixiang.yan@monash.edu

Funding information

Australian Research Council; Jacobs Foundation

Abstract: Simulation-based learning provides students with unique opportunities to develop key procedural and teamwork skills in close-to-authentic physical learning and training environments. Yet, assessing students' performance in such situations can be challenging and mentally exhausting for teachers. Multimodal learning analytics can support the assessment of simulation-based learning by making salient aspects of students' activities visible for evaluation. Although descriptive analytics have been used to study students' motor behaviours in simulation-based learning, their validity and utility for assessing performance remain unclear. This study aims at addressing this knowledge gap by investigating how indoor positioning analytics can be used to generate meaningful insights about students' tasks and collaboration performance in simulation-based learning. We collected and analysed the positioning data of 304 healthcare students, organised in 76 teams, through correlation, predictive and epistemic network analyses. The primary findings were (1) large correlations between students' spatial-procedural behaviours and their group performances; (2) predictive learning analytics that achieved an acceptable level (0.74 AUC) in distinguishing between low-performing and high-performing teams regarding collaboration performance; and (3) epistemic networks that can be used for assessing the behavioural differences across multiple teams. We also present the teachers' qualitative evaluation of the utility of these analytics and implications for supporting formative assessment in simulation-based learning.

KEYWORDS
assessment, CSCL, learning analytics, performance, teamwork

Practitioner notes

What is currently known about this topic

- Assessing students' performance in simulation-based learning is often challenging and mentally exhausting.
- The combination of learning analytics and sensing technologies has the potential to uncover meaningful behavioural insights in physical learning spaces.
- Observational studies have suggested the potential value of analytics extracted from positioning data as indicators of highly-effective behaviour in simulation-based learning.

What this paper adds

- Indoor positioning analytics for supporting teachers' formative assessment and timely feedback on students' group/team-level performance in simulation-based learning.
- Empirical evidence supported the potential use of epistemic networks for assessing the behavioural differences between low-performing and high-performing teams.
- Teachers' positively validated the utility of indoor positioning analytics in supporting reflective practices and formative assessment in simulation-based learning.

Implications for practitioners

- Indoor positioning tracking and spatial analysis can be used to investigate students' teamwork and task performance in simulation-based learning.
- Predictive learning analytics should be developed based on features that have direct relevance to teachers' learning design.
- Epistemic networks analysis and comparison plots can be useful in identifying and assessing behavioural differences across multiple teams.

INTRODUCTION

Simulation-based learning in physical learning spaces is gaining more prominence as an educational approach that provides students with the opportunity to practice their procedural and teamwork skills in immersive and close-to-authentic settings (Chernikova et al., 2020). High-stakes disciplines, such as emergency response and healthcare education, particularly value this approach and its high ecological validity as critical experience can be gained in simulations without exposing students and victims or patients to unnecessary risks (Lateef, 2010). Teachers have also used simulations to assess and provide feedback on students' performance in various domains whereas students are meant to demonstrate technical and functional expertise; problem-solving and decision-making skills; and effective collaboration and communication skills (Sarfati et al., 2019; Theodoulou et al., 2018).

Assessing students' team performance in simulation-based learning is often challenging (Lateef, 2010; Sarfati et al., 2019; Theodoulou et al., 2018). Literature on collaboration and teamwork suggests that while individual members' performance is related to team performance, other group-level factors, such as how team members collaborate with each other and utilise the resource and space available, also play essential roles (Fransen et al., 2013; Hall et al., 2018). For such assessments to be both reliable and developmental, teachers must be able to reliably evaluate the dynamics among multiple students and between students and procedural tasks within the simulation learning space (Lateef, 2010). Teachers are also often required to complete these assessments in real-time as the results would contribute to providing students with constructive feedback immediately after the simulation, during the *debrief*, which is essential to the educational effectiveness of simulation-based learning (Barry Issenberg et al., 2005; Fanning & Gaba, 2007; Rooney & Nyström, 2018). Fulfilling these responsibilities would demand teachers to have an in-depth understanding of the simulation scenario and the science of teamwork (Lateef, 2010). Consequently, teachers without prior encounters may struggle with assessing performance and providing immediate feedback to students.

Multimodal learning analytics (MMLA) can potentially support the assessment and feedback provision process in simulation-based learning by making salient aspects of students' learning behaviours visible for objective evaluation and for provoking reflection (Crescenzi-Lanna, 2020; Sharma & Giannakos, 2020; Yan, Zhao, et al., 2022). Recent small-scale studies in MMLA have used wearable positioning sensors to automatically capture students' positioning traces in simulation-based learning (Echeverria et al., 2018; Fernandez-Nieto, Martinez-Maldonado, Echeverria, et al., 2021; Fernandez-Nieto, Martinez-Maldonado, Kitto, & Shum, 2021; Zhao et al., 2022). These fine-grained *x-y* positioning traces have been modelled to extract insights about students' spatial and collaborative behaviours, including students' movement within a team (Echeverria et al., 2018; Fernandez-Nieto, Martinez-Maldonado, Kitto, & Shum, 2021); an individual's presence in spaces of interest (Fernandez-Nieto, Martinez-Maldonado, Echeverria, et al., 2021); and students' interaction with other team members (Fernandez-Nieto, Martinez-Maldonado, Echeverria, et al., 2021; Zhao et al., 2022). Yet, this research direction is still at a preliminary stage. Most of the prior works have been limited to merely modelling an individual student's spatial or collaborative behaviours, with little connection to the teamwork or task performance. The only exception is previous work by Zhao et al. (2022) who correlated students' non-verbal features with teachers' assessment on team collaboration but focused primarily on the non-verbal features in students' speech instead of their spatial/motor behaviours. To the best of our knowledge, there has been limited previous work that has aimed at automatically modelling students' team performance in simulation-based learning using their physical positioning traces.

The current study was focused on mapping students' physical positioning traces to meaningful analytics that teachers could use to support the formative assessment of students' team performance in simulation-based learning. We collected the positioning traces of 304 students in 76 healthcare simulations performed over four weeks. Students' positioning traces were modelled into six types of spatial-procedural behaviours (summarised later in Table 2). Correlation analyses were conducted to investigate the relationships between students' spatial-procedural behaviours and teachers' assessment of task and collaboration performances. We then constructed machine learning models to automate the distinction between high-performing and low-performing teams. Moreover, epistemic network analyses were conducted to illustrate the differences in behavioural connections between high-performing and low-performing teams. Finally, we conducted a focus group session with two senior teaching staff to evaluate the validity of our findings and the potential usages of these analytics in post-simulation debriefs. In sum, the contribution of this paper is twofold: (1) we propose an innovative methodological approach to extract and represent meaningful insights about students' team performance in simulation-based learning through indoor positioning analytics and epistemic networks, and (2) explore how teachers can potentially use these analytics in their reflective debriefs to provide constructive feedback to students. The current findings could provide the methodological and empirical basis for future research to initiate the investigation and development of automated supportive educational technologies for formative assessments.

BACKGROUND

Assessment and feedback in simulation-based learning

Assessment in simulation-based learning is focused on evaluating students' performance on both technical and non-technical aspects (Miller, 1990). The technical aspects often involve assessing students' knowledge and competence of procedural skills. For example, in clinical settings, students' technical performance would be related to tasks such as patient assessment, oxygen delivery and other procedural behaviours (Michelson & Manning, 2008). Whereas, the non-technical aspects are mostly about whether the communication and collaboration among team members were effective whilst dealing with multiple tasks and distractions (Holcomb et al., 2002). Like any other formative assessments, these assessments follow five principles (with psychometrics properties [Cizek et al., 2019; Way et al., 2010]); they need to be *valid*—capturing the constructs they are designed to measure (Messick, 1995), *reliable*—producing consistent results across similar contexts (Crocker & Algina, 2009), *fair*—providing students with equal opportunities to succeed (Gipps & Stobart, 2009), *sustainable*—easy to implement and maintain (Swiecki et al., 2022), and *developmental*—facilitating learning progression through feedback (Swiecki et al., 2022; Taras, 2005) in order to have actual benefits on students. However, trade-offs often exist between different assessment principles (Swiecki et al., 2022), especially for simulation-based learning where assessments are conducted in close-to-authentic contexts (Gulikers et al., 2004). Such assessment could be more valid in measuring students' procedural and collaboration skills than standardised tests (Theodoulou et al., 2018), but with the potential consequences of being less reliable and sustainable due to the highly contextualised responses/solutions students could provide (eg, different strategies of resolving a medical emergency [Sarfati et al., 2019]) and the resource burdens in conducting these assessments (eg, training teachers to conduct systematic observations [Decker et al., 2013]). Additionally, whether an assessment in simulation-based learning is developmental highly depends on the presence and quality of immediate feedback from teachers (Michelson & Manning, 2008).

Immediate feedback, an essential part of formative assessment, has been widely recognised as the most important component of simulation-based learning (Barry Issenberg et al., 2005; Chernikova et al., 2020; Hall & Tori, 2017; Michelson & Manning, 2008). This feedback is often provided to students immediately after the simulation, during a dedicated debrief session, where teachers commonly utilise formative assessment results and observational notes as evidence to scaffold students' performance and behaviours (Hall & Tori, 2017). The educational theory of constructivism often guides the scaffolding process by focusing feedback on both students' correct and incorrect behaviours as well as directions for improvements (Decker et al., 2013). The goal of feedback or formative assessment is to elicit critical thinking and learning integration in students, which is essential for cultivating their professional practice and preparing them for workforce placements (Hall & Tori, 2017). However, such assessments and feedback could also become harmful when teachers merely focus on the technical aspects of students' performance and do not provide sufficient feedback on their communication and collaboration skills (Fanning & Gaba, 2007). Therefore, teachers must pay close attention to the task-related behaviours of individual students as well as the dynamics among different team members. This responsibility is extremely challenging to fulfil, especially in close-to-authentic and complex simulations that involve multiple tasks and students (Barry Issenberg et al., 2005).

The traditional approach of conducting formative assessment in simulation-based learning predominately relied on teachers' systematic observations and students' self and peer assessment (Fanning & Gaba, 2007; Hall & Tori, 2017). Self and peer assessment have, however, been found to be inaccurate, especially for low-performing students who are overly generous with their scoring for themselves and their peers (Davis et al., 2006). On the other hand, the reliability and validity of teachers' systematic observations depend on their prior experiences in simulated environments, and their familiarisation with the learning objectives (Hall & Tori, 2017). Decker et al. (2013) advised that teachers should have formal training and be assessed by more experienced facilitators before conducting systematic observations independently. Otherwise, they might lack the competency to provide constructive feedback about students' technical and non-technical performances as they would get overwhelmed by the high complexity of the simulation scenarios. Even for experienced teachers, the mental fatigue from repetitive and intensive observations could potentially jeopardise the quality of the constructive feedback students received (Klusmann et al., 2008). Therefore, providing teachers with supportive tools is essential for the sustainable assessment of simulation-based learning.

A common technological aid that has been used to support teachers with their formative assessment and feedback in simulation-based learning involves the use of video recordings (Fanning & Gaba, 2007; Scherer et al., 2003). Teachers could prepare exemplar recordings to illustrate examples of correct behaviours to students, which has been found to positively influence students' learning outcomes, especially for students with low prior knowledge (Chernikova et al., 2020). Yet, video recordings are impractical for the purpose of providing immediate feedback during the debrief as it is almost impossible for teachers to actually memorise and precisely identify the key moments of multiple students in the video while conducting systematic observations (Fraser et al., 2018). Additionally, playing back video recordings of the whole simulation during the debrief can also be distractive and time-consuming, reducing the direct reflection and interaction time between teachers and students (Fanning & Gaba, 2007). Consequently, the lack of evidence-based educational technologies in supporting teachers' formative assessment and immediate feedback has remained as a persistent gap in simulation-based learning (Mariani & Doolen, 2016).

Learning analytics and assessment

The field of learning analytics has a strong tie with learning assessment. Gašević, Greiff and Shaffer (Gašević et al., 2022) identified three core themes about the relationships between research in learning analytics and assessment. These themes include (1) *analytics for assessment*—using learning analytics approaches to support the assessment of learning, (2) *analytics of assessment*—developing analytic approaches to examine different principles of existing assessment practices and (3) *validity of measurement*—constructing conceptual models and practical approaches to assuring validity in the measurement of learning in learning analytics. In the context of simulation-based learning, *analytics for assessment* would be related to research on learning analytics approaches that can support evaluating students' procedural and collaboration skills and generating evidence for constructive feedback (eg, Fernandez-Nieto, Martinez-Maldonado, Kitto, & Shum, 2021; Martinez-Maldonado, Echeverria, Fernandez Nieto, & Buckingham, 2020). *Analytics of assessment* would involve research that aims to understand the reliability of teachers' assessment and scoring criteria across multiple simulations (eg, Zhao et al., 2022). Finally, studies investigating the *validity* *of measurement* would consider the external assessment validity of theory-based trace data and behavioural features in capturing students' external performance measures (eg, Milligan & Griffin, 2016). While all three of these themes are equally important, most of the existing learning analytics research on simulation-based learning focused on the first theme, *analytics for assessment* (Echeverria et al., 2018; Fernandez-Nieto, Martinez-Maldonado, Echeverria, et al., 2021; Fernandez-Nieto, Martinez-Maldonado, Kitto, & Shum, 2021; Martinez-Maldonado, Echeverria, Fernandez Nieto, & Buckingham, 2020; Zhao et al., 2022). The current study explored the role of learning analytics in assessment and feedback provision of simulation-based learning from all three aspects; further details were elaborated in the research questions presented in the following subsection.

Multimodal and indoor positioning analytics

The combination of learning analytics and sensing technologies may hold the promising future for educational technologies that can potentially support formative assessment by automating and augmenting the process of immediate feedback in simulation-based learning (Sharma & Giannakos, 2020). This subfield of MMLA endeavours to generate evidence-based insights about students' learning behaviours and cognitive and emotional states from fine-grained physical and physiological traces (Blikstein, 2013). The data collection capabilities of various combinations of sensing technologies and analysis methodologies in educational research have been well established in prior MMLA studies (Crescenzi-Lanna, 2020; Yan, Martinez-Maldonado, et al., 2022; Yan, Zhao, et al., 2022). In addition, a recent systematic literature review also illustrated the potential benefits of MMLA technologies in supporting reflective teaching and learning practices (Sharma & Giannakos, 2020). The rest of this section outlines some recent studies exploring the use of MMLA and sensing technologies in simulation-based learning. It also elaborates the rationale behind each of our research questions, in light of the three themes of learning analytics and assessment described above.

The predominant research directions for generating insights about students' behaviours in simulation-based learning involve the use of wearable physiological sensors and indoor positioning tracking (Yan, Zhao, et al., 2022). For example, in physically simulated clinical classrooms, these tracking systems have been used to investigate students' team dynamics by assigning wearable trackers to each student (Fernandez-Nieto, Martinez-Maldonado, Echeverria, et al., 2021). In particular, students' physiological responses, such as heart rate, skin temperature, and electrodermal activity, have been captured through wearable sensors and used to model individuals' stress responses while dealing with simulated emergency events (Ronda-Carracao et al., 2021). However, the large variation within and between individuals' physiological responses demands an accurate baseline (Dawson et al., 2017), which is often infeasible to measure in practice. Although analysing physiological synchrony can reveal insights about metacognitive experiences and group performance without the demand of baselines (Dindar et al., 2020), this direction remains largely for educational research instead of practical purposes.

Indoor positioning tracking and spatial analysis have been increasingly used to generate insights about students' behaviours to support reflective practices in simulation-based learning. This line of research, indoor positioning analytics, is established based on Hall's (1966) foundation work on the *theory of proxemics*, which has proposed different sets of interpersonal distances or proximity individuals would maintain during social interactions and the associations between these distances and the nature of their relationships; including intimate (0–0.45 m), personal (0.45–1.22 m), social (1.2–3.7 m) and public (more than 3.7 m) distances. This theory has evident relevance for analysing human behaviours in physical learning spaces (Sorokowska et al., 2017). According to the empirical evidence from decades of social psychology studies, physical proximity is one of the best predictors of interactions (Back et al., 2008). For example, Echeverria et al. (2018) developed a zone-based classification model to automate the mapping from students' physical positioning traces to movements within a team. Based on this model, Fernandez-Nieto, Martinez-Maldonado, Echeverria, et al. (2021) translated spatial data into the interaction among team members through social network analysis and the transition between spaces of interest using epistemic network analysis. Martinez-Maldonado, Echeverria, Fernandez Nieto, and Buckingham (2020) further combined students' spatial behaviours with their task behaviours (coded manually) to generate prototypes of a layered timeline of actions through principles of data storytelling. These modelling and visualisation results were considered meaningful and helpful by both the teachers (Fernandez-Nieto, Martinez-Maldonado, Kitto, & Shum, 2021; Martinez-Maldonado, Echeverria, Fernandez Nieto, & Buckingham, 2020) and the students (Martinez-Maldonado, Echeverria, Fernandez Nieto, & Buckingham, 2020). Although these prior studies illustrated the *descriptive analytics* (about behaviours and events that unfolded during a simulation) that can be extracted from students' positioning traces, none of them has investigated *diagnostic analytics* that can be used to reveal relationships between students' spatial behaviours and teachers' assessment of team performance. These diagnostic analytics (*analytics of assessment*) could potentially provide insights into the reliability of teachers' assessment and scoring criteria across different teams and between different simulation scenarios. This gap in the existing MMLA research motivates the first research question of this study:

- **RQ1**: How do students' spatial-procedural behaviours, extracted from indoor positioning traces, relate to teachers' assessment of team performance in terms of procedural (task) and collaboration learning objectives?

Predictive analytics (*validity of measurement*) have also been frequently used in MMLA research for the automatic identification of students' task performance (see reviews Sharma & Giannakos, 2020; Yan, Zhao, et al., 2022). A previous study has illustrated the potential of position-based predictive analytics in supporting teachers to pinpoint low-progress students (Yan, Martinez-Maldonado, et al., 2022). Likewise, predictive analytics may also provide the opportunity to support inexperienced teachers by automating the distinction between low-performing and high-performing teams in simulation-based learning. This automation could potentially reduce teachers' mental burdens and make the assessments in simulation-based learning more sustainable compared to only relying on observations. Thus, the second research question explored this opportunity:

- **RQ2**: To what extent can students' spatial-procedural behaviours be used to automate the distinction between low-performing and high-performing teams across various learning objectives?

While predictive analytics may help to distinguish between low-performing and high-performing teams, it does not provide additional insights to unpack such differences. Making these behavioural differences visible is essential for formative assessment and immediate feedback to remain developmental (Fanning & Gaba, 2007). Teachers could potentially use in-depth behavioural insights to provide examples of correct behaviours and guide low-performing teams toward the direction of improvement (Hall & Tori, 2017). This potential of indoor positioning analytics (*analytics for assessment*) was explored in the third research question:

- **RQ3**: To what extent can the differences in behavioural connections of low-performing and high-performing teams be identified from their indoor positioning traces?

Finally, the birth of human-centred learning analytics emphasised the importance of involving educational stakeholders in the design of effective learning analytics (Buckingham Shum et al., 2019). Milligan (2018) also stressed that the measurement and insight of learning in learning analytics should provide sufficient utility for teachers and are interpretable by them. Additionally, demonstrating results to relevant stakeholders, such as teachers, and incorporating their responses has also been a common approach in prior MMLA research on simulation-based learning (Fernandez-Nieto, Martinez-Maldonado, Echeverria, et al., 2021; Fernandez-Nieto, Martinez-Maldonado, Kitto, & Shum, 2021; Martinez-Maldonado, Echeverria, Fernandez Nieto, & Buckingham, 2020; Zhao et al., 2022). Therefore, the validity and interpretability of the aforementioned analytics (RQ1–3) and their potential utility in post-simulation debrief were evaluated by teachers through the final research question:

- **RQ4**: How can teachers use indoor positioning analytics to support their formative assessment and immediate feedback during simulation-based learning?

This study goes beyond prior works that used students' physical positioning traces to model their spatial behaviours. In particular, while previous work on indoor positioning analytics has investigated students' movement within a team (Echeverria et al., 2018; Fernandez-Nieto, Martinez-Maldonado, Kitto, & Shum, 2021), presence in spaces of interest (Fernandez-Nieto, Martinez-Maldonado, Echeverria, et al., 2021), and interaction with other team members (Fernandez-Nieto, Martinez-Maldonado, Echeverria, et al., 2021; Zhao et al., 2022), the current study focused primarily on formative assessment and group/team-level feedback. This work may be similar to our previous works that used indoor positioning analytics to model students' progression of math and reading abilities in open learning spaces (Yan et al., 2021; Yan, Martinez-Maldonado, et al., 2022), but we focus specifically on the assessment of task and teamwork performance during simulation-based learning in this study. To our knowledge, this study is the first that explored the potential of indoor positioning analytics in supporting teachers' formative assessment and immediate feedback during simulation-based learning. The findings of this exploratory study would contribute empirical evidence to guide the development of educational technologies that can provide practical benefits to both teachers and students as a supportive tool for formative assessment and post-simulation debriefs.

METHODS

Educational context

This study was conducted in a fully immersive healthcare simulation unit as a part of the authentic assessment for third-year undergraduate nursing and midwifery students. The simulations took place in a physically simulated clinical classroom with four authentic patient beds, medical equipment (eg, oxygen masks and vital signs monitors), and technological hybrid patient manikins with variable heart rates and measurable pulses. This learning unit aimed to assess and provide opportunities for students to practice their collaboration, communication and prioritisation skills in the setting of a deteriorating patient. A total of 304 students participated in 76 healthcare simulations over four weeks. Each simulation consisted of a team of four students. Three experienced teachers took turns in facilitating the simulations and evaluating students' team performance from a one-way mirror control room. Another seven teaching staff supported the simulations by role-playing as confederates or leading the reflective debriefs after each simulation.

Students were randomly chosen to participate in one of the two different simulation scenarios, with 152 students in each scenario. In both scenarios, students were responsible for taking care of four different beds (see Figure 1). The *primary tasks* involved taking care of a deteriorating patient (Bed 4). Scenario A involved a covertly deteriorating patient that required students to actively identify her conditions and prioritise care accordingly. In contrast, Scenario B involved an overtly deteriorating patient whose relative (role played by teaching staff) actively approached the students for help. Apart from the primary task, students were also assigned with several *secondary tasks* (related to Beds 1, 2 and 3) to complete. At the start of each scenario, two students (graduate nurses) entered the simulation room first, with the other two (ward nurses) waiting outside. The ward nurses would enter the simulation room after the two graduate nurses called for assistance. Teachers monitored the students' progress and determined the time to end the simulation. On average, the duration of Scenario A (M = 20.8 minutes, SD = 5.3 minutes) was longer than Scenario B (M = 12.1 minutes, SD = 2.3 minutes).

Apparatus and data collection

The Pozyx (2022) creator toolkit (approximately 1200 USD) was used as the positioning tracking system for this study. A total of five anchors were affixed at various positions on the walls of the simulation room to record the signals transmitted by the UWB (Ultra-Wideband) tags worn by students. The positioning tag transmits signals at 60 Hz. The signals were automatically computed into real-time *x-y* coordinates (with a 100 mm accuracy) through Pozyx's proprietary positioning engine using the wireless Two-Way Ranging algorithms. Students' positioning traces were only recorded when they were positioned inside the simulation room.

Teachers evaluated students' team performance through six assessment items. These items were directly drawn from the learning objectives of the simulation unit. As shown in Table 1, the first three items assessed students' team performance on task-related learning objectives (T1–T3) and followed by three items on collaboration-related learning objectives

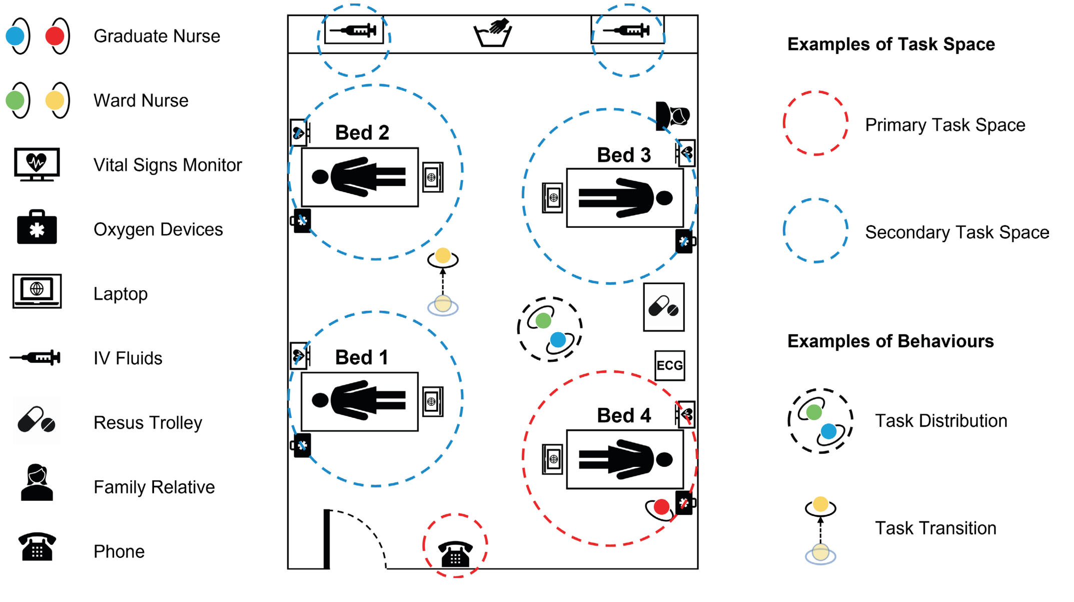

FIGURE 1 Floor plan of the simulation room and examples of the different task spaces and students' spatial-procedural behaviours.

TABLE 1 Assessment items on students' team performance

Item	Description
T1	The students have demonstrated a structured approach to patient assessment and management
T2	The students have recognised and responded to early signs of patient deterioration
T3	The students have appropriately activated MET calls
T4	The students have contributed to effective teamwork
T5	The students have communicated information effectively to the health care team
T6	The students have demonstrated an understanding of the roles of the multidisciplinary team

(T4–T6). A seven-point bipolar Likert scale, ranging from *strongly disagree* (1) to *strongly agree* (7), was used to evaluate each item. These assessment items were not a part of students' summative assessment.

Ethics approval was obtained from Monash University (Project ID: 28026). Informed written consent was also obtained from both students and teaching staff before the data collection. Among the 76 healthcare simulations, we analysed 52 sessions (208 students) where all four students consented for the data collection and for their data to be used in future publications. Students' identities were protected using de-identified tracker labels (eg, nurse-RED and nurse-BLUE).

Data processing and modelling

Students' fine-grained positioning data were prepared for the extraction of spatial-procedural behavioural features through two data cleaning procedures. First, positioning trace data were normalised to one data point per second. Linear interpolation was then applied to fill in any missing value that were lost from environmental signal occlusion (Gløersen & Federolf, 2016), limited to a maximum of 10 consecutive missing values. Otherwise, students were considered as outside of the simulation room.

A total of six features were extracted from students' positioning traces to generate insights about their spatial-procedural behaviours (see Table 2). These behavioural features were essential in covering the underlying behavioural connections among different teams, further elaborated in the Analysis section (Section 3.4). They were also developed based on the learning objectives (see Table 1) that teachers used to assess and provide feedback on students' task performances, captured by the spatial proximity between students, and collaboration performances, captured by the procedural meaning of different task spaces (see Figure 1). For each second in a simulation, individual students' behaviours were coded as one of these features. We calculated the proportion of time (percentage) a student spent performing each behaviour during a simulation to avoid any potential systematic biases from different simulation duration. The rationale and extraction procedures for these features are elaborated below.

Identifying different task spaces

Task prioritisation skills are critical for healthcare students to succeed in clinical practice (Tiwari et al., 2005). The simulation room was segmented into different task spaces based on the learning design of the scenarios to capture behavioural features related to task prioritisation (see Figure 1). As explained in Section 3.1, students primarily worked around the tasks associated with the patient in each bed. Therefore, a centroid was created for each

TABLE 2 Spatial-procedural behavioural features (percentage)

Features	Spatial-procedural behaviours
Collaborate_Primary (CP)	Students working on the primary tasks collaboratively
Independent_Primary (IP)	Students working on the primary tasks individually
Collaborate_Secondary (CS)	Students working on the secondary task collaboratively
Independent_Secondary (IS)	Students working on the secondary task individually
Task_Distribution (TD)	Students distributing the responsibility of different tasks
Task_Transition (TT)	Students transiting from one task to another task

bed. The primary task space was identified as within 1.5 m of the centroid for Bed 4. The secondary task spaces were identified as within 1.5 m of the centroid for Bed 1, 2 and 3. The 1.5 m radius was used because this distance criterion also covered the medical equipment directly related to each bed's tasks, including the vital sign monitors, oxygen devices and laptops for accessing health records. In addition, the primary task spaces also included the space around the phone (within 1.0 m radius), which directly related to the essential actions for resolving the primary task, including calling for the assistance of ward nurses and requesting a MET (Medical Emergency Team) call. The space around the IV (Intravenous) fluids stations (within 1.0 m radius) were identified as secondary task spaces. This process of segmenting the simulation room into primary and secondary task spaces was co-designed by the chief coordinator who developed the simulations and validated by two experienced teachers who have facilitated numerous simulations.

Distinguishing collaborative behaviours

Collaboration is an essential element in healthcare simulations, especially in emergency situations (Fernandez-Nieto, Martinez-Maldonado, Echeverria, et al., 2021). The collaborative behaviours among students were distinguished based on their physical proximity. A potential instance of collaboration was identified when two students were within one meter proximity for more than 10 seconds. This method of modelling collaboration from positioning data has been used in several prior studies (Martinez-Maldonado, Echeverria, Schulte, et al., 2020; Yan et al., 2021; Yan, Martinez-Maldonado, et al., 2022) and grounded in the theoretical foundations of proxemics (Hall, 1966). The 10 second threshold was adopted to reduce the likelihood of misidentifying unintended collocation as meaningful collaboration, for example, when students were merely passing by each other (Greenberg et al., 2014). The first four behavioural features (Table 2) were developed based on the combination of collaborative behaviours and task spaces, including *Collaborate_Primary (CP)*, *Independent_Primary (IP)*, *Collaborate_Secondary (CS)* and *Independent_Secondary (IS)*.

Characterising task distribution

Outside of the identified task spaces, students' spatial behaviours could also contain certain task-related meanings. For example, Tuckman (1965) theorised that an essential stage of a collaborative team would involve members distributing the responsibility of different tasks. Students' behaviours related to task distribution were identified when two or more students had prolonged collocations (eg, more than ten seconds) outside of the identified task spaces (see an example in Figure 1), where they were more likely to engage in conversations about task responsibilities instead of working collaboratively on a task. The chief coordinator and the two experienced teachers further validated this approach for identifying *Task_Distribution (TD)* behaviours.

Classifying transition behaviours

The final behavioural feature concerns the transition behaviours from one task to another. High frequency of transition behaviours could potentially be a behavioural indicator of students feeling confused by the overwhelmed clinical environment and unable to prioritise their attention among different tasks (Papp et al., 2003). Transition behaviours, *Task_Transition (TT)*, were identified when an individual student was positioned outside of the task spaces and with no other students in their proximity. No time constraints were imposed for classifying transition behaviours as students who positioned themselves away from task spaces and other students for long periods could also indicate confusion.

Analysis

Four analyses were conducted to investigate the role of MMLA, extracted from students' positioning traces, in supporting the assessment of simulation-based learning. These analyses explored (1) diagnostic analytics about the relationships between students' spatial-procedural behaviours and teachers' assessment of team performances, (2) predictive analytics on students' collaboration and task performances and (3) learning analytics that could provide behavioural insights to support student reflection. We also conducted thematic content analysis on teachers' responses during a focus group session to investigate the validity and implications of these different analytics.

Correlation analysis—RQ1

The linear relationships between students' spatial-procedural behaviours and teachers' assessment of team performance were analysed to generate diagnostic analytics about the underlying behavioural basis of teachers' scoring criteria. The proportion of time spent performing each spatial-procedural behaviour (Table 2) were calculated on a team basis for each session. The Pearson's correlation between these behavioural features and the assessment items (in Table 1) were analysed separately for Scenario A and B to gain a better understanding of the variation in teachers' scoring criteria across different learning designs (Gašević et al., 2016). Multiple comparisons were accounted for through the Bonferroni correction method, which adjusted the significance threshold (initially 0.05) based on the number of comparisons.

Predictive analytics—RQ2

Machine learning (ML) algorithms were used to develop predictive models for automatically differentiating between high-performing and low-performing teams in nursing simulations. These models were constructed through a four-step procedure. First, for each assessment item, we labelled a team as *low-performing* if they received a score of four or below. This binary classification rule was validated by the teachers.

In the second step, the minority class was oversampled to avoid class imbalance. We implemented the synthetic minority oversampling technique (SMOTE) (Chawla et al., 2002) to balance the number of positive (low-performing) and negative cases (high-performing) to a ratio of 1:1, which originally was unbalanced at a ratio of around 1:2.47 for task performance items (T1, T2 and T3) and around 1:7.61 for collaboration performance items (T4, T5 and T6). The SMOTE method was applied to oversampling only the training sets within the cross-validation loop to avoid data contamination in the validation sets (Farrow et al., 2019). A three-fold cross-validation (repeated 20 times) was chosen to evaluate the models due to the relatively small sample size (31 teams for Scenario A and 21 teams for Scenario B). This segmentation approach would ensure sufficient data for the validation sets and the grid search process.

Feature selection was conducted during the third step to select the relevant features for predicting each assessment item. The contribution of the features to the predictive models was assessed using the Shapley values (Štrumbelj & Kononenko, 2014). This approach was chosen over other permutation-based methods because the shapely additive explanation (SHAP) provides the most granular results regarding the feature-level influence on prediction (Štrumbelj & Kononenko, 2014). Features with low impact on a model's predictions were removed.

Predictive models were then constructed based on the selected features. Five different ML classifiers from the Python Scikit-Learn library (Pedregosa et al., 2011) were used to develop the predictive models, including logistic regression (LR), support vector machine (SVM), random forests (RF), k-nearest neighbours (KNN) and multi-layer perceptron artificial neural networks (ANN). These classifiers were chosen as they were commonly used in learning analytics research to model binary classification problems (Hellas et al., 2018).

The performance of the models was evaluated through four model evaluation metrics, including area under the ROC curve (AUC), precision, recall and Cohen's kappa (k) for inter-rater reliability. Accuracy was also reported but was not used to evaluate the models as it can be misleading for highly imbalanced data-sets (Galar et al., 2011), especially for collaboration performance items. Therefore, we used grid search for hyperparameter tuning to optimise AUC. The aim was to achieve a high level of differentiation between low-performing and high-performing students as incorrect identifications could conflict with teachers' judgement, thus, potentially causing cognitive dissonance and increasing their mental burdens (Kaaronen, 2018). The performance of the models was compared with a dummy classifier from the Python Scikit-Learn library (Pedregosa et al., 2011) that always return the most frequent class label (high-performing), with a baseline AUC of 0.50.

Epistemic network analysis—RQ3

The differences in behavioural connections between low-performing and high-performing teams were analysed through epistemic network analysis (ENA). This analytic approach has commonly been used in learning analytics research to model the structure of connections in cognitive, social, and interaction data (Shaffer et al., 2016). We used the ENA Web Tool (version 1.7.0) (Marquart et al., 2018) to model the connections between students' spatial-procedural behaviours by quantifying the co-occurrence of different behaviours within a simulation session and producing a weighted network of co-occurrences. For every ten seconds (an action), a team was coded as conducting one or more of the six spatial-procedural behaviours (in Table 2). This ten-second interval was chosen to ensure that students were actually performing the spatial-procedural behaviours instead of being misclassified due to unintended collocation (Greenberg et al., 2014). We then used a moving window size of 20 actions, as recommended by Siebert-Evenstone et al. (2017), to construct a network model that shows the connection between a team's current behaviours and their behaviours that occurred within the recent temporal context (current behaviour plus behaviours in the previous 19 actions). The resulting network was constructed by aggregating the network model for each second in a given team. This process was then repeated for each team. The resulting 52 networks were normalised through dimensional reduction techniques (see Shaffer et al., 2016 for details) and plotted as points on a two-dimension epistemic network. Comparison plots were constructed by subtracting the weighted networks of low-performing from high-performing teams to illustrate the structural differences between these teams. This comparison was made for both students' task performance (sum of item T1, T2, and T3) and collaboration performance (sum of item T4, T5, and T6). Mann–Whitney *U* tests were also conducted to examine the statistical differences between low-performing and high-performing teams along the *x*-axis (MR1) and the *y*-axis (SVD2) of the comparison plots for each scenario.

Thematic content analysis—RQ4

Two of the experienced teachers (Teacher_01 and Teacher_02) were invited to a post-hoc focus group session to discuss the validity of the learning analytics generated from students' positioning traces and the role of these analytics in assessment and feedback for simulation-based learning. A retrospective reflection technique was used for the focus group session (Hassenzahl & Ullrich, 2007). The teachers were first presented with different types of learning analytics (findings of RQ1–RQ3). For the correlation results, the teachers were asked whether the findings were consistent with their understanding and knowledge of the simulations. For the prediction and epistemic network results, the teachers were asked to discuss the potential roles of these learning analytics in supporting the assessment and facilitation of the simulations, as well as express their main concerns regarding presenting these learning analytics to support student reflection in the post-simulation debrief. We conducted thematic content analysis to identify the core themes presented in teachers' responses regarding the potential adoption of predictive analytics and epistemic networks as supportive debriefing tools.

RESULTS

Correlation results—RQ1

Students' spatial-procedural behaviours exhibited significant correlations with teachers' assessment of their task and collaboration performances. In terms of *task performance*, two large positive correlations were found (Cohen, 2013). In particular, the longer students worked collaboratively on the primary tasks (*CP*), the higher teachers rated those students in terms of having effectively recognised and responded to early signs of patient deterioration (T2; $r = 0.50$, $p = 0.004$), and appropriately activated MET calls (T3; $r = 0.57$, $p < 0.001$). For *collaboration performance*, two large positive correlations and two large negative correlations were identified (Cohen, 2013). The longer students worked collaboratively on the *primary tasks* (*CP*), the higher teachers rated that the students had contributed to effective teamwork (T4; $r = 0.62$, $p < 0.001$), and the students had communicated effectively with the health care team (T5; $r = 0.59$, $p < 0.001$). In contrast, the longer students worked collaboratively on the *secondary tasks* (*CS*), the lower the teachers rated that the students had communicated effectively with the health care team (T5; $r = -0.60$, $p < 0.001$). Likewise, the longer students spent transitioning between different tasks (*TT*), the lower the teachers rated that the students had contributed to effective teamwork (T4; $r = -0.48$, $p = 0.007$).

The aforementioned correlations only exist in Scenario A but not in Scenario B. A possible explanation is that the learning design of Scenario A required the students to actively identify the primary tasks (the covertly deteriorating patient in Bed 1) and prioritise their attention accordingly. In this particular learning design, the teachers would base their scoring criteria on any self-initiated behaviour related to the primary and secondary tasks. Whereas in Scenario B, the students were passively drawn to the overtly deteriorating patient by her relative. This aspect of the learning design of Scenario B could limit the value of the students' spatial behaviours as it contained little information about their autonomous actions, thus, less valued by the teachers when assessing team performance.

Prediction performances—RQ2

The state-of-the-art ML algorithms performed differently across the assessment items and scenarios. In Scenario A (Table 3), the model performed indistinguishably from the baseline model (AUC = 0.50) in differentiating high-performing and low-performing teams regarding their structured approach to patient assessment and management (T1). While the model performance was slightly better than the baseline model for T2, recognising and responding to early signs of patient deterioration, the large standard deviation (0.18) indicates that the model was subject to a high level of variability and, thus, potentially unreliable. The model performance was significantly better than the baseline model for T3, appropriately activating MET calls. A potential explanation for the poor performances regarding T1 and T2 is that these task-related items assessed complex behaviours (that involve both verbal and non-verbal communication) that could not be reflected in students' spatial behaviours alone. This reason could also potentially explain the better performances for T3 as this item directly related to students' spatial behaviours within one of the primary task spaces (the small red circle around the phone which is used to make the MET calls, see Figure 1). In contrast, the models outperformed the baseline model regarding each collaboration assessment item (T4–T6) with an acceptable level of differentiation (Hosmer Jr et al., 2013) (AUC of around 0.74). This finding was expected as students' spatial behaviours has been frequently found to contain meaningful insights about their interaction and collaboration in physical learning spaces (Riquelme et al., 2020; Saquib et al., 2018; Yan et al., 2021). Whereas in Scenario B (Table 4), the models performed poorly for every assessment item with low AUC and large variability. A possible explanation for these poor performances is that students' spatial behaviours could become irrelevant and even misleading for the predictive models when students' spatial behaviours were less autonomous. The learning design of Scenario B was already structured to actively draw students' attention toward the primary task space around Bed 4 (in Figure 1).

TABLE 3 The best performing model for each assessment item (scenario a): Mean and SD (in parenthesis)

Item	Dimension	Classifier	AUC	Precision	Recall	Cohen's *k*	Accuracy
T1	Task	RF	0.55 (0.12)	0.34 (0.12)	0.39 (0.12)	0.10 (0.12)	0.62 (0.12)
T2	Task	LR	0.69 (0.18)	0.22 (0.18)	0.73 (0.18)	0.19 (0.18)	0.67 (0.18)
T3	Task	KNN	0.74 (0.10)	0.59 (0.10)	0.75 (0.10)	0.44 (0.10)	0.73 (0.10)
T4	Collaboration	RF	0.73 (0.08)	0.49 (0.08)	0.52 (0.08)	0.42 (0.08)	0.89 (0.08)
T5	Collaboration	RF	0.73 (0.07)	0.32 (0.07)	0.53 (0.07)	0.35 (0.07)	0.89 (0.07)
T6	Collaboration	RF	0.74 (0.07)	0.48 (0.07)	0.52 (0.07)	0.42 (0.07)	0.89 (0.07)

TABLE 4 The best performing model for each assessment item (scenario B): Mean and SD (in parenthesis)

Item	Dimension	Classifier	AUC	Precision	Recall	Cohen's *k*	Accuracy
T1	Task	RF	0.44 (0.17)	0.08 (0.17)	0.17 (0.17)	−0.09 (0.17)	0.61 (0.17)
T2	Task	SVM	0.39 (0.17)	0.03 (0.17)	0.10 (0.17)	−0.16 (0.17)	0.56 (0.17)
T3	Task	SVM	0.47 (0.15)	0.22 (0.15)	0.25 (0.15)	−0.06 (0.15)	0.57 (0.15)
T4	Collaboration	ANN	0.56 (0.18)	0.21 (0.18)	0.42 (0.18)	0.08 (0.18)	0.64 (0.18)
T5	Collaboration	RF	0.49 (0.16)	0.14 (0.16)	0.32 (0.16)	−0.04 (0.16)	0.60 (0.16)
T6	Collaboration	SVM	0.54 (0.18)	0.29 (0.18)	0.38 (0.18)	0.08 (0.18)	0.62 (0.18)

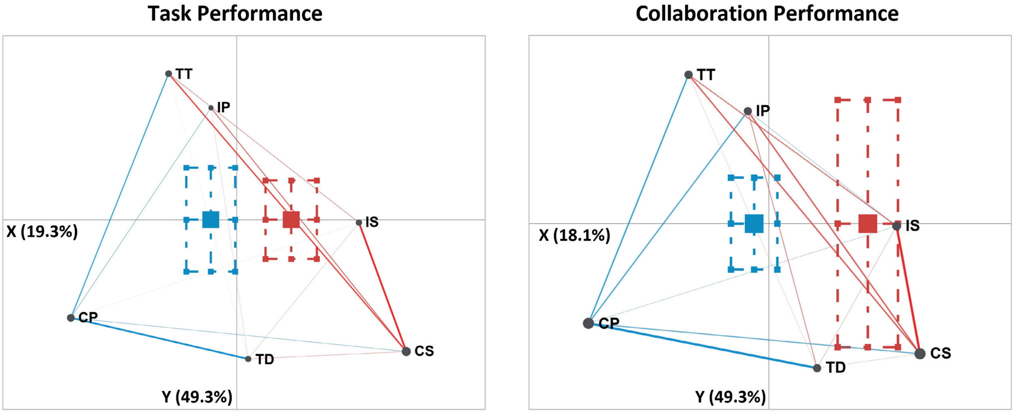

FIGURE 2 Comparison plots showing the epistemic networks of high-performing teams (blue) and low-performing teams (red) regarding task (left) and collaboration (right) performance in Scenario A, with the corresponding means (squares) and 95% confidence intervals (dashed boxes). The percentages in parentheses represent the total variance in the model accounted for by each dimension (axis).

Epistemic networks—RQ3

Epistemic networks analysis revealed significant differences between the behavioural connections of high-performing and low-performing teams in Scenario A. Along the *x*-axis (MR1) of task performance, a Mann–Whitney *U* test showed that the spatial-procedural behaviours of high-performing teams ($N = 21$) were significantly different from low-performing teams ($N = 10$; $U = 16$, $p < 0.001$, $r = 0.85$). Likewise, there was also a significant difference between high-performing ($N = 27$) and low-performing teams ($N = 4$) regarding their collaboration performance ($U = 0$, $p < 0.001$, $r = 1.00$). As shown in the comparison plots (Figure 2), for both task and collaboration performance, high-performing teams' worked predominantly on the primary tasks. In particular, the behaviours of high-performing teams were strongly characterised by the connections between *Collaborate_Primary* (CP) and *Task_Transition* (TT) as well as between *Collaborate_Primary* (CP) and *Task_Distribution* (TD). They also made a stronger connection between *Collaborate_Primary* (CP) and *Independent_Primary* (IP). In contrast, low-performing teams focused mainly on the secondary tasks. Their behaviours were strongly characterised by the connections between *Collaborate_Secondary* (CS) and *Task_Transition* (TT) as well as between *Collaborate_Secondary* (CS) and *Independent_Primary* (IP). The strong triadic connection among *Collaborate_Secondary* (CS), *Independent_Secondary* (IS), and *Task_Transition* (TT) further highlighted that low-performing teams were transiting between the secondary tasks, working collaboratively or independently.

In Scenario B, the behavioural differences varied in terms of task and collaboration performance (see Figure 3). A significant difference was found between high-performing (N = 17) and low-performing teams (N = 4) among the x-axis regarding their task performance (U = 61, p = 0.01, r = −0.79). The behaviours of high-performing teams were strongly characterised by the connections between *Collaborate_Primary* (CP) and *Task_Distribution* (TD), whereas, the dyadic connections between *Collaborate_Primary* (CP) and *Independent_Secondary* (IS) was the strongest for low-performing teams (Figure 3, left). On the other hand, for collaboration performance, the behavioural differences between high-performing (N = 18) and low-performing teams (N = 3) were not statistically significant among both the x-axis and the y-axis. Yet, visual analysis of the comparison plot (Figure 3, right) revealed that high-performing teams were strongly characterised by the triadic connections among *Independent_Primary* (IP), *Independent_Secondary* (IS) and *Task_Transition* (TT), as well as the triadic connections among *Independent_Secondary* (IS), *Task_Transition* (TT) and *Task_Distribution* (TD). Whereas, the behaviours of low-performing teams were strongly characterised by the dyadic connections between *Independent_Primary* (IP) and *Collaborate_Primary* (CP). These findings supported our explanations for previous results regarding RQ1 and RQ2, which emphasised that the learning design of Scenario B has actively drawn students to the primary task spaces.

Insights from teachers—RQ4

Consistency between correlational findings and teachers' assessment criteria

Teachers reported that the correlations were consistent with their understanding and knowledge of the simulations. In particular, the positive correlations between *Collaborate_Primary* (CP) and assessment items T2 to T5 resonated with teachers' assessment focus on collaboration and task prioritisation. One of the teachers mentioned that “Definitely, the value was placed on them [the students] working collaboratively and where the priority of care was, which was in Bed 4. The more we saw them working together around Bed 4, we definitely thought that they had achieved the learning outcomes” (Teacher_01). We observed the same reaction for the negative correlations between *Collaborate_Secondary* (CS) and item T5 and between *Task_Transition* (TT) and item T4. For example, Teacher_01 mentioned

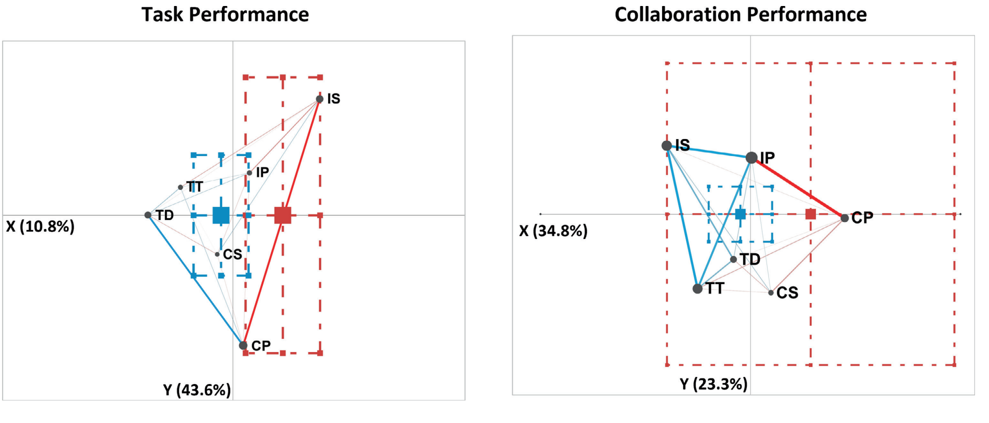

FIGURE 3 Comparison plots showing the epistemic networks of high-performing teams (blue) and low-performing teams (red) regarding task (left) and collaboration (right) performance in Scenario B.

that “The transition behaviours, generally speaking, to me that is students not really knowing what they are doing and moving around a bit panicky, and not really been part of the team that are trying to solve the problem around deterioration in Bed 4”.

As for the strength of these correlations, Teacher_02 provided an interesting explanation: “There may be some impacts on the correlations where it looks like students, based on their wristbands [positioning tags] or their location, are not at Bed 4, but we have rated them as good communication and been effective, this may potentially be because they have communicated well but doing tasks that are related to Bed 4 and sometimes require them to step away”. This explanation identifies a limitation of indoor positioning analytics in capturing students' task and collaboration behaviours for assessment purposes.

Predictive analytics as confirmatory tools for inexperienced teachers

The core theme that emerged from teachers' responses regarding the potential use case of predictive analytics is that these prediction models can serve as the basis for developing *confirmatory tools* that help inexperienced teachers to be more confident with their assessment results. For example, one of the teachers suggested that “If I am an inexperienced teacher and got a message [from predictive tools] that said the positioning analytics indicate poor collaboration, that would probably confirm what I was thinking and observing. Then I could go into the debrief and confidently talk about collaboration as the main point of the debrief” (Teacher_01). Likewise, Teacher_2 concurred with the statement “This is what I saw, and this is what the data says, and so I feel comfortable to say it”.

The teachers also suggested that the accuracy of these prediction models, especially for task-related learning objectives (T1–T3), could potentially be improved by including students' audio traces and timeline data from the simulation scenario script. Teacher_01 mentioned that “I think if there is something in the audio, you will be able to hear them if they are talking about a MET call. If we use the scenario data, we know when deterioration occurs. So around that time, is there audio around MET call?”

Epistemic networks as supportive evidence to motivate reflection in the debrief

Teachers demonstrated profound interest toward showing the epistemic networks (Figures 2 and 3) as *supportive evidence* in the debrief to initiate students' self-reflection process. One of the teachers suggested that “This [epistemic network] would allow us to have a little bit of data to support what happened so that we can have a really good discussion in the debrief” (Teacher_02). Teacher_01 supported this usage scenario with the statement “I agree completely. Taking the students through something like you just did with us [explaining the epistemic networks] is a really good visual to then open up the discussion about, okay, tell me what was happening”. The visual comparisons of high-performing and low-performing teams could encourage self-reflection in both types of teams, not limiting to low-performing ones. Teacher_01 explained that “Obviously, it would be more important when they [teams] are low-performing, but equally when they are high-performing, saying that this shows you did a good job, what were some of the elements that helped you collaborate?”

Team-based evidence over individual-level insights for the debrief

Another interesting theme that emerged from teachers' responses is their preference for demonstrating evidence on team behaviours over individual actions in the debrief. This preference was also found when teachers discussed the visual illustrations of the epistemic networks. In particular, Teacher_02 mentioned that “I actually think this would be one of the best things that we could use [for debrief]. Because it does not show an individual's movement, it shows the team, and it lets you talk about them as a team”. She (Teacher_02) further elaborated that “So there is no feeling of, oh no, they can see where I moved. This [epistemic network] is just very much about that team dynamic that occurred within the room”. Teacher_01 also expressed a similar preference in the statement “I agree. It takes the individual out”. Additionally, Teacher_01 also thought the team-based epistemic networks were more useful than individual-level ones as “It also takes the decision making out because we could show them a video of the scenario. But actually, they still got to interpret that as being not the right thing. Whereas this [the comparison plot] shows, it is already interpreted. It is actually not the right thing. Let us talk about why it is not the right thing”. Whereas, such comparisons are more difficult to illustrate on an individual level as the definition of the right thing depends on the student's role.

DISCUSSION

This study explored the potential of indoor positioning analytics in supporting the assessment and feedback on students' team performances in simulation-based learning. The main findings of this research are discussed below.

Main findings

For the first research question (RQ1), students' spatial-procedural behaviours extracted from indoor positioning traces demonstrated significant and large correlations with teachers' assessment on task and collaboration performances in Scenario A but not in Scenario B of the simulation. The significant findings in Scenario A resonate with prior research on the relationships between spatial behaviours and student performances (Yan, Martinez-Maldonado, et al., 2022). This finding illustrates the potential of indoor positioning analytics in understanding assessment reliability by uncovering the behavioural consistency in teachers' assessment and scoring criteria. Additionally, the non-significant results in Scenario B provided empirical evidence to further emphasise the importance of considering the learning design in educational research, as well as the need to tailor learning analytics based on teachers' learning design intentions (Gašević et al., 2016; Mangaroska & Giannakos, 2018).

The findings of the second research question (RQ2) indicate that students' spatial-procedural behaviours can achieve an acceptable level of distinction between low-performing and high-performing teams regarding their collaboration performances but not their task performances in simulation-based learning. This finding illustrated a potential limitation of indoor positioning analytics, that is, its inability to distinguish different procedural behaviours that occurred within the same physical spaces, which has been mentioned in the limitation sections of prior MMLA studies (Martinez-Maldonado, Echeverria, Schulte, et al., 2020; Yan et al., 2021). The current finding also stressed the importance of assessing the external validity of any trace data and behavioural features used to generate learning analytics as such validity could vary across different learning contexts (Gašević et al., 2022). Likewise, the alignment between machine learning features (eg, behavioural features) and high-order educational constructs (eg, procedural performances) is essential when developing predictive analytics. This need resonates with the call for context-specific learning analytics and the issues of generalised predictive models (Gašević et al., 2016).

In terms of the third research question (RQ3), the differences in behavioural connections of low-performing and high-performing teams can be identified and visualised from their indoor positioning traces through epistemic networks analysis and comparison plots. The success of this modelling approach with indoor positioning traces resonates with the illustrations in Fernandez-Nieto et al.'s studies (Fernandez-Nieto, Martinez-Maldonado, Echeverria, et al., 2021; Fernandez-Nieto, Martinez-Maldonado, Kitto, & Shum, 2021) but also extended the application scenarios of indoor positioning analytics beyond merely showing the transition between spaces of interest in one or two teams. The visible differences between correct and incorrect behaviours among multiple teams in simulation-based learning can also be illustrated through indoor positioning analytics using comparison plots (eg, Figure 2) without the need for further interpretation (the additional process of students interpreting behaviours as being not the right thing). This finding also provided empirical evidence to support the use of learning analytic approaches, particularly epistemic network analysis, as supportive tools to promote developmental assessment in simulation-based learning.

For the final research question (RQ4), teachers demonstrated overall positive feelings toward using indoor positioning analytics, especially epistemic networks, to support their formative assessment and immediate feedback during simulation-based learning. The positive attitude and willingness to embrace educational technologies as supportive tools in the debrief are consistent with the findings of the previous interview and focus group studies (Fernandez-Nieto, Martinez-Maldonado, Kitto, & Shum, 2021; Martinez-Maldonado, Echeverria, Fernandez Nieto, & Buckingham, 2020). Teachers' profound interests in the epistemic networks and comparison plots also supported using these figures as visualisation tools to guide student reflection in the debrief. This finding could potentially resolve Fernandez-Nieto, Martinez-Maldonado, Kitto, and Shum (2021) concerns about the potential confusions that might be associated with directly exposing teachers and students to epistemic networks. Their ability to comprehend the epistemic networks should not be a problem with network nodes that are clearly defined, easily understandable and directly related to learning behaviours (eg, Table 2). Thus, the findings of RQ4 provided qualitative evidence to support the methodological foundations (Milligan, 2018) of indoor positioning analytics as its interpretability and utility for supporting assessment and feedback were endorsed by teachers.

Implications for research and practice

The current findings have several implications regarding the principles of assessments (elaborated in Section 2.1). The diagnostic analytics (RQ1) can be used to understand the consistency and reliability of assessment by uncovering the behavioural basis of teachers' assessment criteria (the relationships between learning behaviours and assessment outcomes). This implication is particularly important in complex learning activities (eg, in simulation-based learning), where teachers' assessment criteria may be subject to a high level of variability, resulting from the constantly changing dynamics between students and procedural tasks (Lateef, 2010). Sustainable assessment could be promoted through predictive analytics (RQ2) and epistemic networks (RQ3). Both of these analytics can support teachers to implement and maintain the assessment of simulation-based learning, for example, by reducing the evidence teachers have to collect manually for the debrief (Hall & Tori, 2017). Additionally, epistemic networks (RQ3) can also be used to promote developmental assessment by strengthening teachers' feedback with visual evidence and exemplars (Chernikova et al., 2020), as well as encouraging self-reflection in both low-performing and high-performing teams.

The current findings also have several implications for future learning analytics and educational research and practice. Researchers interested in understanding the decision-making process of teachers' assessment (eg, why teachers rated a team as high-performing) in simulation-based learning could benefit from using indoor positioning analytics. The combination of spatial analysis and indoor positioning tracking could overcome the shortcomings of traditional research methodologies, such as the observer bias in systematic observations (Hall & Tori, 2017) and the recall bias in post-hoc survey questionnaires (Davis et al., 2006), and increase the reliability and validity of data on students' spatial-procedural behaviours. Additionally, the automatic and minimum intrusive nature of this approach also enables researchers to collect more data over multiple simulations without additional resource burdens (other than the one-time cost of the sensors). This approach could benefit the sustainability of research projects where multiple data collections occur in the same learning space over long periods (Yan, Zhao, et al., 2022). For example, researchers can use indoor positioning analytics to understand the temporal changes in students' procedural and collaboration behaviours when investigating the development of students' learning behaviours in simulation-based learning (eg, Scherer et al., 2003).

The practical implications of this research for supporting assessment and feedback in simulation-based learning are well documented in teachers' responses during the focus group session. Although teachers intended to use predictive analytics as confirmatory tools to support inexperienced teachers to be more confident with their assessment results, the level of differentiation (AUC) needs to be improved as the current model performances, even for predicting collaboration performances (Table 3), would still result in false-positives and false-negatives, which could cause cognitive dissonance and potential reducing inexperienced teachers' confidence (Kaaronen, 2018). Future studies should improve these model performances by incorporating other data sources, such as audio (Zhao et al., 2022), before adopting predictive analytics in practice. Whereas epistemic networks in the form of comparison plots can be directly used in the debrief as visual evidence to support teachers' immediate feedback on students' team performance and initiate reflective learning. Future research can develop a structured approach for teachers to adopt when guiding students through epistemic networks.

Ethical considerations

The data-driven approach of this study is sensitive to several ethical issues. For example, although the positioning tracker's intrusiveness is minimal, the physical presence of the sensors could still increase students' feeling of being surveilled by third parties other than the teachers, which is reported to be a major concern of students in higher education (Whitelock-Wainwright et al., 2020). This feeling could potentially have unforeseen influences on their performances (Dawson, 2006), and thus, obtaining informed consent from students is essential before using sensor-based learning analytics. Additionally, issues regarding individual privacy and data security also need to be addressed (Alwahaby et al., 2021; Crescenzi-Lanna, 2020; Yan, Zhao, et al., 2022). The current study used data anonymisation to remove individual identities from the dataset, thus protecting students' privacy. This approach has also been applied in several recent MMLA studies (Bassiou et al., 2016; Praharaj et al., 2018; Yan, Martinez-Maldonado, et al., 2022; Yan, Martinez-Maldonado, et al., 2022). Likewise, focusing on group-level analytics also removes individual elements from any results used in future debriefs. While this approach cannot support personalised feedback for each student, responses from teachers reflected its potential value in supporting group reflections.

Ethical considerations regarding analytics-based assessments also need to be addressed. In particular, the current sensor-based approach can be subject to the ethical issues of assessing students when they are unaware of being assessed (Mangaroska et al., 2021). Therefore, it is essential to emphasise the potential usage of the data to students and provide them with ways to access and control their own data (Timmis et al., 2016). We encourage researchers and practitioners to communicate these information to students both in the consent form and verbally to prevent students from missing any important information, such as their traces data could be used for assessment purposes.

Limitations

The current findings and implications are subject to two limitations. The indoor positioning analytics may only be suitable for supporting the assessment and feedback of learning activities that satisfy two essential conditions of learning design. First, students' performances need to be strongly associated with their collaboration on and prioritisation of different tasks (eg, collaborating on primary instead of secondary tasks). Second, these behaviours can be clearly distinguished from their positioning traces in the learning spaces (eg, through physical proximity and task spaces in Figure 1). Using indoor positioning analytics to research or support learning activities that contain little meaningful spatial information by design would be ineffective and inappropriate, just like the current findings on Scenario B. Additionally, we only conducted the focus group session with senior teachers with extensive simulation experiences. Inexperienced teachers may have different thoughts about using indoor positioning analytics for the debrief. For example, the epistemic networks (eg, Figure 2) might be less intuitive to them.

CONCLUSION

This study goes beyond prior works that used indoor positioning analytics to model students' movement within a team (Echeverria et al., 2018; Fernandez-Nieto, Martinez-Maldonado, Kitto, & Shum, 2021), presence in spaces of interest (Fernandez-Nieto, Martinez-Maldonado, Echeverria, et al., 2021), and interaction with other team members (Fernandez-Nieto, Martinez-Maldonado, Echeverria, et al., 2021; Zhao et al., 2022) by focusing on formative assessment and group/team-level feedback. Specifically, this paper illustrated how students' physical positioning traces could be mapped into meaningful learning analytics to support teachers' formative assessment and timely feedback on students' team performances in simulation-based learning. The current findings have potential implications for further research on educational technologies that can make students' learning behaviours visible for teachers to use as evidence in the debrief. This research may also pave the way for a sensor-based and technology-driven approach to assessing students' team performances in simulation-based learning.

ACKNOWLEDGEMENTS

This research was funded partially by the Australian Government through the Australian Research Council (project number DP210100060). Roberto Martinez-Maldonado's research is partly funded by Jacobs Foundation. Open access publishing facilitated by Monash University, as part of the Wiley – Monash University agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST

The authors have declared no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request. The data are not publicly available due to privacy or ethical restrictions.

ETHICS STATEMENT

Ethics approval was obtained from Monash University (Project ID: 28026).

REFERENCES

Alwahaby, H., Cukurova, M., Papamitsiou, Z., & Giannakos, M. (2021). The evidence of impact and ethical considerations of multimodal learning analytics: A systematic literature review. *EdArXiv* edarxiv.org/sd23y

Back, M. D., Schmukle, S. C., & Egloff, B. (2008). Becoming friends by chance. *Psychological Science*, *19*(5), 439–440.

Barry Issenberg, S., Mcgaghie, W. C., Petrusa, E. R., Lee Gordon, D., & Scalese, R. J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: A BEME systematic review. *Medical Teacher*, *27*(1), 10–28.

Bassiou, N., Tsiartas, A., Smith, J., Bratt, H., Richey, C., Shriberg, E., D'Angelo, C., & Alozie, N. (2016). Privacy-preserving speech analytics for automatic assessment of student collaboration. In *Interspeech* (pp. 888–892). International Speech Communication Association.

Blikstein, P. (2013). Multimodal learning analytics. In *LAK' 13* (pp. 102–106). Association for Computing Machinery.

Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-centred learning analytics. *Journal of Learning Analytics*, *6*(2), 1–9.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357.

Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., & Fischer, F. (2020). Simulation-based learning in higher education: A meta-analysis. *Review of Educational Research*, *90*(4), 499–541.

Cizek, G. J., Andrade, H. L., & Bennett, R. E. (2019). Formative assessment. In *Handbook of formative assessment in the disciplines* (pp. 3–19). Routledge.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Crescenzi-Lanna, L. (2020). Multimodal learning analytics research with young children: A systematic review. *BJET*, *51*(5), 1485–1504.

Crocker, L., & Algina, J. (2009). *Introduction to classical and modern test theory*. Cengage Learning.

Davis, D. A., Mazmanian, P. E., Fordis, M., Van Harrison, R., Thorpe, K. E., & Perrier, L. (2006). Accuracy of physician self-assessment compared with observed measures of competence: A systematic review. *JAMA*, *296*(9), 1094–1102.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2017). *The electrodermal system*. Cambridge University Press.

Dawson, S. (2006). The impact of institutional surveillance technologies on student behaviour. *Surveillance & Society*, *4*(1/2), 69–84.

Decker, S., Fey, M., Sideras, S., Caballero, S., Boese, T., Franklin, A. E., Gloe, D., Lioce, L., Sando, C. R., Meakim, C., & Borum, J. C. (2013). Standards of best practice: Simulation standard VI: The debriefing process. *Clinical Simulation in Nursing*, *9*(6), S26–S29.

Dindar, M., Järvelä, S., & Haataja, E. (2020). What does physiological synchrony reveal about metacognitive experiences and group performance? *British Journal of Educational Technology*, *51*(5), 1577–1596.

Echeverria, V., Martinez-Maldonado, R., Power, T., Hayes, C., & Shum, S. B. (2018). Where is the nurse? Towards automatically visualising meaningful team movement in healthcare education. In *AIED* (pp. 74–78). Springer.

Fanning, R. M., & Gaba, D. M. (2007). The role of debriefing in simulation-based learning. *Simulation in Healthcare*, *2*(2), 115–125.

Farrow, E., Moore, J., & Gašević, D. (2019). Analysing discussion forum data: A replication study avoiding data contamination. In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 170–179). Association for Computing Machinery.

Fernandez-Nieto, G., Martinez-Maldonado, R., Echeverria, V., Kitto, K., An, P., & Buckingham, S. S. (2021). What can analytics for teamwork proxemics reveal about positioning dynamics in clinical simulations? *PACM HCI*, *5*(CSCW1), 1–24.

Fernandez-Nieto, G. M., Martinez-Maldonado, R., Kitto, K., & Shum, S. B. (2021). Modelling spatial behaviours in clinical team simulations using epistemic network analysis: methodology and teacher evaluation. In *LAK21: 11th International Learning Analytics and Knowledge Conference* (pp. 386–396). Association for Computing Machinery.
Fransen, J., Weinberger, A., & Kirschner, P. A. (2013). Team effectiveness and team development in CSCL. *Educational Psychologist*, *48*(1), 9–24.
Fraser, K. L., Meguerdichian, M. J., Haws, J. T., Grant, V. J., Bajaj, K., & Cheng, A. (2018). Cognitive load theory for debriefing simulations: Implications for faculty development. *Advances in Simulation*, *3*(1), 1–8.
Galar, M., Fernandez, A., Barrenechea, E., Bustince, H., & Herrera, F. (2011). A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *42*(4), 463–484.
Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education*, *28*, 68–84.
Gašević, D., Greiff, S., & Shaffer, D. W. (2022). Towards strengthening links between learning analytics and assessment: Challenges and potentials of a promising new bond. *Computers in Human Behavior*, *134*, 107304. https://www.sciencedirect.com/science/article/pii/S0747563222001261
Gipps, C., & Stobart, G. (2009). *Fairness in assessment. In: Educational assessment in the 21st century* (pp. 105–118). Springer.
Gløersen, Ø., & Federolf, P. (2016). Predicting missing marker trajectories in human motion data using marker intercorrelations. *PLoS One*, *11*(3), e0152616.
Greenberg, S., Boring, S., Vermeulen, J., & Dostal, J. (2014). Dark patterns in proxemic interactions: A critical perspective. In *DIS'14* (pp. 523–532). Association for Computing Machinery.
Gulikers, J., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development*, *52*(3), 67–86.
Hall, E. T. (1966). *The hidden dimension, vol. 609*. Doubleday.
Hall, K., & Tori, K. (2017). Best practice recommendations for debriefing in simulation-based education for Australian undergraduate nursing students: An integrative review. *Clinical Simulation in Nursing*, *13*(1), 39–50.
Hall, K. L., Vogel, A. L., Huang, G. C., Serrano, K. J., Rice, E. L., Tsakraklides, S. P., & Fiore, S. M. (2018). The science of team science: A review of the empirical evidence and research gaps on collaboration in science. *American Psychologist*, *73*(4), 532–548.
Hassenzahl, M., & Ullrich, D. (2007). To do or not to do: Differences in user experience and retrospective judgments depending on the presence or absence of instrumental goals. *Interacting with Computers*, *19*(4), 429–437.
Hellas, A., Ihantola, P., Petersen, A., Ajanovski, V. V., Gutica, M., Hynninen, T., Knutas, A., Leinonen, J., Messom, C., & Liao, S. N. (2018). Predicting academic performance: A systematic literature review. In *Proceedings companion of the 23rd annual ACM conference on innovation and technology in computer science education* (pp. 175–199). Association for Computing Machinery.
Holcomb, J. B., Dumire, R. D., Crommett, J. W., Stamateris, C. E., Fagert, M. A., Cleveland, J. A., Dorlac, G. R., Dorlac, W. C., Bonar, J. P., Hira, K., Aoki, N., & Mattox, K. L. (2002). Evaluation of trauma team performance using an advanced human patient simulator for resuscitation training. *Journal of Trauma and Acute Care Surgery*, *52*(6), 1078–1086.
Hosmer, D. W., Jr., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression, vol. 398*. John Wiley & Sons.
Kaaronen, R. O. (2018). A theory of predictive dissonance: Predictive processing presents a new take on cognitive dissonance. *Frontiers in Psychology*, *9*, 2218.
Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Teachers' occupational well-being and quality of instruction: The important role of self-regulatory patterns. *Journal of Educational Psychology*, *100*(3), 702–715.
Lateef, F. (2010). Simulation-based learning: Just like the real thing. *Journal of Emergencies, Trauma and Shock*, *3*(4), 348–352.
Mangaroska, K., & Giannakos, M. (2018). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies*, *12*(4), 516–534.
Mangaroska, K., Martinez-Maldonado, R., Vesin, B., & Gašević, D. (2021). Challenges and opportunities of multimodal data in human learning: The computer science students' perspective. *Journal of Computer Assisted Learning*, *37*, 1030–1047.
Mariani, B., & Doolen, J. (2016). Nursing simulation research: What are the perceived gaps? *Clinical Simulation in Nursing*, *12*(1), 30–36.
Marquart C, Hinojosa C, Swiecki Z, Eagan B, Shaffer D. *Epistemic Network Analysis (Version 1.7. 0)[Software]*. epistemicnetwork.org 2018.
Martinez-Maldonado, R., Echeverria, V., Fernandez Nieto, G., & Buckingham, S. S. (2020). From data to insights: A layered storytelling approach for multimodal learning analytics. In *Proceedings of the 2020 chi conference on human factors in computing systems* (pp. 1–15). Association for Computing Machinery.

Martinez-Maldonado, R., Echeverria, V., Schulte, J., Shibani, A., Mangaroska, K., & Buckingham Shum, S. (2020, July). Moodoo: Indoor positioning analytics for characterising classroom teaching. In *International Conference on Artificial Intelligence in Education* (pp. 360–373). Springer.

Messick, S. (1995). Standards of validity and the validity of standards in performance asessment. *Educational Measurement: Issues and Practice*, *14*(4), 5–8.

Michelson, J. D., & Manning, L. (2008). Competency assessment in simulation-based procedural education. *The American Journal of Surgery*, *196*(4), 609–615.

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine*, *65*(9), S63–S67.

Milligan, S. K. (2018). Methodological foundations for the measurement of learning in learning analytics. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 466–470). Association for Computing Machinery.

Milligan, S. K., & Griffin, P. (2016). Understanding learning and learning design in MOOCs: A measurement-based interpretation. *Journal of Learning Analytics*, *3*(2), 88–115.

Papp, I., Markkanen, M., & von Bonsdorff, M. (2003). Clinical environment as a learning environment: Student nurses' perceptions concerning clinical learning experiences. *Nurse Education Today*, *23*(4), 262–268.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., & Vanderplas, J. (2011). Scikit-learn: Machine learning in python. *The Journal of Machine Learning Research*, *12*, 2825–2830.

Pozyx, *Unleash the power of real-time location solutions*; 2022. https://www.pozyx.io/.

Praharaj, S., Scheffel, M., Drachsler, H., & Specht, M. (2018). Multimodal analytics for real-time feedback in co-located collaboration. In *European Conference on Technology Enhanced Learning* (pp. 187–201). Springer.

Riquelme, F., Noel, R., Cornide-Reyes, H., Geldes, G., Cechinel, C., Miranda, D., Villarroel, R., & Munoz, R. (2020). Where are you? Exploring micro-location in indoor learning environments. *IEEE Access*, *8*, 125776–125785.

Ronda-Carracao, M. A., Santos, O. C., Fernandez-Nieto, G., & Martinez-Maldonado, R. (2021). Towards exploring stress reactions in teamwork using multimodal physiological data. In *In First International Workshop on Multimodal Artificial Intelligence in Education (MAIED 2021) – Held at International Conference on Artificial Intelligence in Education (AIED 2021)* (pp. 49–60). CEUR Workshop Proceedings.

Rooney, D., & Nyström, S. (2018). Simulation: A complex pedagogical space. *Australasian Journal of Educational Technology*, *34*(6).

Saquib, N., Bose, A., George, D., & Kamvar, S. (2018). Sensei: Sensing educational interaction. *PACM IMWUT*, *1*(4), 1–27.

Sarfati, L., Ranchon, F., Vantard, N., Schwiertz, V., Larbre, V., Parat, S., Faudel, A., & Rioufol, C. (2019). Human-simulation-based learning to prevent medication error: A systematic review. *Journal of Evaluation in Clinical Practice*, *25*(1), 11–20.

Scherer, L. A., Chang, M. C., Meredith, J. W., & Battistella, F. D. (2003). Videotape review leads to rapid and sustained learning. *The American Journal of Surgery*, *185*(6), 516–520.

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics*, *3*(3), 9–45.

Sharma, K., & Giannakos, M. (2020). Multimodal data capabilities for learning: What can multimodal data tell us about learning? *British Journal of Educational Technology*, *51*(5), 1450–1484.

Siebert-Evenstone, A. L., Irgens, G. A., Collier, W., Swiecki, Z., Ruis, A. R., & Shaffer, D. W. (2017). In search of conversational grain size: Modeling semantic structure using moving stanza windows. *Journal of Learning Analytics*, *4*(3), 123–139.

Sorokowska, A., Sorokowski, P., Hilpert, P., Cantarero, K., Frackowiak, T., Ahmadi, K., Alghraibeh, A. M., Aryeetey, R., Bertoni, A., Bettache, K., Blumen, S., Błażejewska, M., Bortolini, T., Butovskaya, M., Castro, F. N., Cetinkaya, H., Cunha, D., David, D., David, O. A., … Pierce, J. D., Jr. (2017). Preferred interpersonal distances: A global comparison. *Journal of Cross-Cultural Psychology*, *48*(4), 577–592.

Štrumbelj, E., & Kononenko, I. (2014). Explaining prediction models and individual predictions with feature contributions. *Knowledge and Information Systems*, *41*(3), 647–665.

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., Selwyn, N., & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, *3*, 100075.

Taras, M. (2005). Assessment–summative and formative–some theoretical reflections. *British Journal of Educational Studies*, *53*(4), 466–478.

Theodoulou, I., Nicolaides, M., Athanasiou, T., Papalois, A., & Sideris, M. (2018). Simulation-based learning strategies to teach undergraduate students basic surgical skills: A systematic review. *Journal of Surgical Education*, *75*(5), 1374–1388.

Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. (2016). Rethinking assessment in a digital age: Opportunities, challenges and risks. *British Educational Research Journal*, *42*(3), 454–476.

Tiwari, A., Lam, D., Yuen, K., Chan, R., Fung, T., & Chan, S. (2005). Student learning in clinical nursing education: Perceptions of the relationship between assessment and learning. *Nurse Education Today*, *25*(4), 299–308.

Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin*, *63*(6), 384–399.

Way, W. D., Dolan, R. P., & Nichols, P. (2010). Psychometric challenges and opportunities in implementing formative assessment. In *Handbook of formative assessment* (pp. 297–315). Routledge.

Whitelock-Wainwright, A., Gašević, D., Tsai, Y. S., Drachsler, H., Scheffel, M., Muñoz-Merino, P. J., Tammets, K., & Delgado Kloos, C. (2020). Assessing the validity of a learning analytics expectation instrument: A multinational study. *Journal of Computer Assisted Learning*, *36*(2), 209–240.

Yan, L., Martinez-Maldonado, R., Cordoba, B. G., Deppeler, J., Corrigan, D., Nieto, G. F., & Gasevic, D. (2021). Footprints at school: Modelling in-class social dynamics from students' physical positioning traces. In *LAK21: 11th international learning analytics and knowledge conference* (pp. 43–54). Association for Computing Machinery.

Yan, L., Martinez-Maldonado, R., Gallo Cordoba, B., Deppeler, J., Corrigan, D., & Gašević, D. (2022). Mapping from proximity traces to socio-spatial behaviours and student progression at the school. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13203

Yan, L., Martinez-Maldonado, R., Zhao, L., Deppeler, J., Corrigan, D., & Gasevic, D. (2022). How do teachers use open learning spaces? Mapping from teachers' socio-spatial data to spatial pedagogy. In *LAK22: 12th International Learning Analytics and Knowledge Conference LAK22* (pp. 87–97). Association for Computing Machinery. https://doi.org/10.1145/3506860.3506872

Yan, L., Zhao, L., Gasevic, D., & Martinez-Maldonado, R. (2022). Scalability, sustainability, and ethicality of multimodal learning analytics. In *LAK22: 12th International Learning Analytics and Knowledge Conference LAK22* (pp. 13–23). Association for Computing Machinery. https://doi.org/10.1145/3506860.3506862

Zhao, L., Yan, L., Gasevic, D., Dix, S., Jaggard, H., Wotherspoon, R., Alfredo, R., Li, X., & Martinez-Maldonado, R. (2022). Modelling co-located team communication from voice detection and positioning data in healthcare simulation. In *LAK22: 12th International Learning Analytics and Knowledge Conference* (pp. 370–380). Association for Computing Machinery.

How to cite this article: Yan, L., Martinez-Maldonado, R., Zhao, L., Dix, S., Jaggard, H., Wotherspoon, R., Li, X., & Gašević, D. (2023). The role of indoor positioning analytics in assessment of simulation-based learning. *British Journal of Educational Technology*, *54*, 267–292. https://doi.org/10.1111/bjet.13262