JOURNAL OF LEARNING ANALYTICS

Volume 11(3), 62–81. https://doi.org/10.18608/jla.2024.8487

Designing a Human-Centred Learning Analytics Dashboard In-Use

Riordan Alfredo[1], Vanessa Echeverria[2], Linxuan Zhao[3], LuEttaMae Lawrence[4], Jie Xiang Fan[5], Lixiang Yan[6], Xinyu Li[7], Zachari Swiecki[8], Dragan Gašević[9], Roberto Martinez-Maldonado[10]

Abstract

Despite growing interest in applying human-centred design methods to create learning analytics (LA) systems, most efforts have concentrated on initial design phases, with limited exploration of how LA tools and practices can co-evolve during the actual learning and teaching activities. This paper examines how a human-centred LA dashboard can be further refined and adapted by teachers while actively using it in a real-world scenario (i.e., *design-in-use*), beyond its intended design (i.e., *design-for-use*). We use *instrumental genesis* as a theoretical lens to analyze the temporary and permanent *instrumentalization* of design features and individual and collective *instrumentation* of the LA dashboard. The analysis of semi-structured individual interviews with five nursing teachers who used an LA dashboard to guide team reflections with 224 students (56 teams) revealed technical and pedagogical changes that occurred in both the system's features (*instrumentalization*) and teaching practices (*instrumentation*). We found that teachers adopted the LA dashboard beyond initially intended ways by (i) providing emotional support with the analytics, (ii) reducing details in AI-powered data visualizations for easier comprehension, (iii) creating data narratives to address data limitations, and (iv) collectively developing new practices to use the LA dashboard for co-teaching. Therefore, teachers' design-in-use of the LA dashboard highlights the ongoing need for design improvements to address challenges posed by dynamic data and complex algorithms underlying AI and analytics interfaces.

Notes for Practice

- This paper emphasizes the importance of reporting design-in-use processes in learning analytics (LA) to address evolving educational challenges and alignment of LA with actual pedagogical and practical needs.
- An instrumental genesis lens can be used to analyze the temporary and permanent instrumentalization of design features, and individual and collective instrumentation of LA systems, such as LA dashboards, providing insights into how teachers shape and personalize the system during use.
- This paper also highlights the significance of designing flexible LA dashboard features that allow for teacher-personalized adaptations, fostering a sense of ownership and agency in effectively using the LA dashboard in the wild.
- There is a need for careful consideration of data transparency and explainability when teachers design analytics and AI-powered visualizations in use.

Keywords

Design-in-use, instrumental genesis, human-centred learning analytics, AI-powered, co-design, dashboard, nursing education

Submitted: 23-05-2023 — **Accepted:** 31-08-2024 — **Published:** 23-10-2024

Corresponding author[1] Email: riordan.alfredo@monash.edu Address: Centre for Learning Analytics Monash (CoLAM), Faculty of Information Technology (FIT), Monash University, Australia. ORCID iD: https://orcid.org/0000-0001-5440-6143

[2]Email: vanessa.echeverria@monash.edu Address: CoLAM, FIT at Monash University, Australia and Escuela Superior Politécnica del Litoral, Guayaquil, Ecuador. ORCID iD: https://orcid.org/0000-0002-2022-9588

[3]Email: linxuan.zhao@monash.edu Address: CoLAM, FIT at Monash University, Australia. ORCID iD: https://orcid.org/0000-0001-5564-0185

[4]Email: lu.lawrence@usu.edu Address: Utah State University, Logan, UT 84321 USA. ORCID iD: https://orcid.org/0000-0001-6066-5096

[5]Email: jie.fan@monash.edu Address: CoLAM, FIT at Monash University, Australia. ORCID iD: https://orcid.org/0009-0000-8585-2760

[6]Email: lixiang.yan@monash.edu Address: CoLAM, FIT at Monash University, Australia. ORCID iD: https://orcid.org/0000-0003-3818-045X

[7]Email: xinyu.li1@monash.edu Address: CoLAM, FIT at Monash University, Australia. ORCID iD: https://orcid.org/0000-0003-2681-4451

[8]Email: zach.swiecki@monash.edu Address: CoLAM, FIT at Monash University, Australia. ORCID iD: https://orcid.org/0000-0002-7414-5507

[9]Email: dragan.gasevic@monash.edu Address: CoLAM, FIT at Monash University, Australia. ORCID iD: https://orcid.org/0000-0001-9265-1908

Corresponding author[10] *Email: roberto.martinezmaldonado@monash.edu, CoLAM, FIT at Monash University, Australia. ORCID iD: https://orcid.org/0000-0002-8375-1816*

1. Introduction

The rise of artificial intelligence (AI) and analytics technologies has opened new opportunities for enhancing the quality of learning and teaching. These technologies can provide data-informed insights, promising a better understanding of learning processes and improved educational outcomes (Buckingham Shum & Luckin, 2019; Williamson & Kizilcec, 2022). In this context, various human-centred learning analytics (LA) systems have emerged with the aim of helping teachers improve their teaching practice by increasing awareness of students' progress and their learning experiences (Buckingham Shum et al., 2019, 2024; Kaliisa & Dolonen, 2023). Several LA systems have included teacher-facing LA, such as LA dashboards (e.g., Williamson & Kizilcec, 2022; R. D. Alfredo et al., 2023; Lee-Cultura et al., 2023; Echeverria et al., 2024), designed to aid teachers in making informed educational decisions and interventions, as well as providing feedback. Yet, despite the beneficial intentions in designing LA systems with educational stakeholders, it is vital to examine how teachers practically implement and adapt these tools in real-world settings, taking into account the distinction between intended goals and actual outcomes of the use of new technologies in specific educational contexts (Tchounikine, 2017; Alhadad et al., 2018; Ouhaichi et al., 2023; R. Alfredo, Echeverria, Jin, Yan, et al., 2024). This becomes particularly crucial when dealing with AI and analytics interfaces, considering the challenging task of translating often complex algorithmic outputs into actions, and the need for continual adjustments by designers and users to fully understand what can and cannot be done with the data (Q. Yang et al., 2020).

A rich body of human-computer interaction (HCI) research has investigated how users tend to modify the intended usage of designs of a system to meet their personal needs when "appropriating" the system in a real-world setting (Tchounikine, 2017; Knowles et al., 2019). Within HCI and design disciplines, the term *appropriation* refers to the process by which users adapt, modify, or repurpose a designed product, system, or interface to fit their specific needs or preferences (Overdijk et al., 2006). This process is a manifestation of the broader concept of *design-in-use*, which has been an area of focus in various domains, such as education (Martinez-Maldonado et al., 2018), customer service (Folcher, 2003), manufacturing (Hara et al., 2013), and consumer software (Kim & Lim, 2023), for more than two decades. "Design-in-use" happens during actual use (in *use time*), allowing tools and practices to evolve and adapt dynamically in real time. In contrast, "design-for-use" occurs during the planning and development stages (in *design time*), focusing on creating tools and systems before they are implemented.

In educational contexts, although some efforts have been made to report the design changes to personalize educational technologies by the users (e.g., Core et al., 2017; Martinez-Maldonado et al., 2018), there is still limited empirical evidence on the co-evolution of human-centred AI-powered LA systems while teachers use them in the wild (Topali et al., 2024). Moreover, existing works in LA have focused either on systems developed primarily by researchers with limited input from educational stakeholders during the planning stages (Carvalho et al., 2019; Kim & Lim, 2023; Schmitz et al., 2022; Dollinger et al., 2019) or on systems created with more active participation from teachers and students, but still within the design time phase (R. Alfredo, Echeverria, Jin, Yan, et al., 2024).

This paper addresses this gap by illustrating the *design-in-use* of a teacher-facing LA dashboard in an authentic educational scenario (in the wild): a team-based clinical healthcare simulation in higher education. We conducted individual, in-depth, semi-structured interviews with five nursing teachers ($n = 5$), who used the designed dashboard to guide team reflections with 224 students (56 teams), to gain insights into the co-evolution of both the technology and the teaching practices. We investigated this through the lens of *instrumental genesis* (Rabardel & Bourmaud, 2003; Lonchamp, 2012; Martinez-Maldonado et al., 2018; Vágová, 2021), a theory that provides a comprehensive understanding of how technology evolves beyond its original design when put into use. We report on how nursing teachers address the challenges posed by dynamic data and complex algorithms, realizing teachers as active designers rather than passive consumers of the designed analytics system (R. Alfredo, Echeverria, Jin, Yan, et al., 2024). This positions our paper as a qualitative design initiative aimed at generating rich, context-specific knowledge that can inform future design practices and contribute to theory building. To the best of our knowledge, this is the first paper exploring *design-in-use* in the context of LA.

2. Background and Related Work

2.1 Foundations of Design-in-Use and Instrumental Genesis

Design-in-use is a concept that acknowledges the ongoing influence of users on design *artifacts* (i.e., systems or tools) even after their release (Nelson et al., 2009; Hara et al., 2013; Martinez-Maldonado et al., 2018; Kim & Lim, 2023). This concept emphasizes the active involvement of end-users in shaping and personalizing these artifacts during their use. In essence, design-in-use can be seen as an extension of the design process into practical usage, a form of technology appropriation (Folcher, 2003; Tchounikine, 2017). To further understand this process, we can look at the theory of *instrumental genesis*, which also builds on the notion of technology appropriation (Folcher, 2003). Developed by Rabardel and Bourmaud (2003), this theory proposes the construct of an instrument comprising *utilization schemes* and *artifacts*, where a utilization scheme refers to the strategies and methods employed by users to interact with and leverage a tool, while an artifact is the physical or digital tool itself, which becomes an instrument through its use and adaptation by users (Folcher, 2003). In sum, instrumental genesis provides a framework for understanding how users transform or change an artifact into a useful instrument.

Instrumental genesis explains this transformation through two processes: *instrumentalization* and *instrumentation*. These processes, which can occur simultaneously, illustrate the dynamic relationship between the user's utilization scheme and artifact (Vágová, 2021). Through instrumentalization, the user adapts the artifact to their needs, while through instrumentation, the user's activities and thought processes are shaped by the artifact. This dual process underscores the interactive and reciprocal nature of design-in-use as a co-evolving process of humans and artifacts.

2.1.1 Instrumentalization

In this artifact-oriented process, the focus is on changing the system's design and capabilities to meet the user's needs. These changes are commonly driven by the user's knowledge and preferences, aiming to enhance the system's effectiveness in achieving goals or tasks (Rabardel & Bourmaud, 2003). Vágová (2021) proposes two levels of instrumentalization, permanent and temporary. We present the following definitions.

Permanent instrumentalization involves permanently changing the artifact's original interface to perform a new function. An example of this can be seen in the work of Martinez-Maldonado and colleagues (2018), where researchers collaborated with teachers to explore how users in multi-device environments engage in permanently modifying the interfaces of artifacts to fulfill new functions. They would change the physical arrangement of devices (hardware), adjust the content displayed on different surfaces (software), and refine how team members interacted with the technology to enhance their collaborative activities (user interface features). This represents a significant shift from the artifact's original function, leading to a permanent change in its interfaces or structure.

Temporary instrumentalization refers to momentarily changing the use of the artifact for a specific action under particular circumstances, leading to a *temporary* change in the artifact's capability or goal or task without modifying interfaces, for instance, using a wrench as a hammer (Lonchamp, 2012) or using a book as a ruler (Vágová, 2021). This represents the possibility of the main goal temporarily changing in one-time events or due to situational factors without any changes in its structure or interfaces.

2.1.2 Instrumentation

Instrumentation is a subject-oriented process in which the focus shifts from the artifact's capabilities and constraints to those of the users (Rabardel & Bourmaud, 2003). This approach is about how the artifact influences the user's knowledge and ways to use it. The artifact is perceived as an extension of the users, shaping and being shaped by their actions. A utilization scheme emerges as a psychological construct that represents the user's structured approach or strategy for using the artifact to accomplish specific tasks. Lonchamp (2012) explains that instrumentation occurs at the *individual* and the *collective* levels of artifact utilization.

At the **individual instrumentation** level, users have agency and control over the artifact, enabling them to shape utilization schemes (Rabardel & Bourmaud, 2003). It refers to the process by which users individually learn to use an artifact and incorporate it into their routine (Lonchamp, 2012; Cerratto Pargman et al., 2018). The extent of user engagement with the artifact is intrinsically linked to this process. For example, a user learning to navigate a new system would undergo the instrumentation process (Carvalho et al., 2019). This process involves understanding the system's features and developing strategies for its effective use (Vágová, 2021). The extent of the user's engagement with the artifact, ranging from minimal usage to complete integration that fundamentally changes the utilization scheme, can reflect the stages of the individual instrumentation (Lagrange & Richard, 2022).

At the **collective instrumentation** level, a group of users shares an instrument, coordinating individual actions and aggregating them to achieve common goals (Rabardel & Bourmaud, 2003). For example, teachers may work as a team to use a system to monitor their classes based on the collected and analyzed students' learning data (e.g., Van Schoors et al., 2023). This collective instrumentation is usually common in a team-teaching environment (Decuyper et al., 2023), allowing teachers to collaboratively coordinate their teaching methods, share resources, and integrate their efforts toward improving students' learning.

The foundations of *design-in-use* and *instrumental genesis* presented in this section underpinned the analysis of the design of the LA dashboard *in use* presented below.

2.2 Designing a Teacher-Facing LA Dashboard

One of the key areas where LA has shown significant potential is in the design and development of teacher-facing LA systems. These are commonly designed to provide teachers with insights and actionable information about their students' learning processes in the form of a dashboard (Kaliisa et al., 2023). These systems can also be powered with AI and analytics for various applications (see LA dashboard reviews by Kaliisa et al., 2023, 2024). They can facilitate the interpretation of students' affective data (e.g., Chen et al., 2022; Fernandez Nieto et al., 2022; R. D. Alfredo et al., 2023; Kaliisa & Dolonen, 2023), enable dynamic grouping of students in a classroom (e.g., *Pair-Up by* K. B. Yang et al., 2023), and automatically notify teachers to intervene in learning activities (e.g., *Tutti* by Aleven et al., 2022). The end-user interface of these systems is commonly shaped in the form of a dashboard that is usually complemented with data visualizations (e.g., Amarasinghe et al., 2021; Lee-Cultura et al., 2023; Ouhaichi et al., 2023; Echeverria et al., 2024). Recent AI-powered LA dashboards have leveraged generative AI (GenAI) with large language models (LLMs) to enhance the LA cycle by aiding in the analysis and presentation of data (Yan, Martinez-Maldonado, & Gasevic, 2024). In other words, the GenAI LA dashboard, such as *VizChat* by Yan, Zhao, and colleagues (2024) or *GePeTo* by Becerra and colleagues (2024), can support teachers by generating insightful feedback, identifying patterns in student data, and presenting this information in an easily interpretable manner.

Yet, the design of teacher-facing LA dashboards presents several challenges. One significant challenge that has been reported is the limited involvement of teachers in the design process, leading to tools that may not align with teachers' pedagogical needs and expectations (Kaliisa & Dolonen, 2023; Williamson & Kizilcec, 2022; R. Alfredo, Echeverria, Jin, Swiecki, et al., 2024). Additionally, existing teacher-facing LA dashboard design is often focused on solving technical data aspects rather than actual usage and pedagogical needs (Susnjak et al., 2022; Nazaretsky et al., 2022), potentially hindering their practical use in the wild (Larrabee Sønderlund et al., 2019; Friend Wise et al., 2023; Kaliisa et al., 2024). This lack of alignment between the design of LA systems and teachers' actual practices can result in mistrust and adoption challenges (Slade et al., 2023).

Furthermore, the complexity of interpreting analytics or AI from the dashboard can increase teachers' cognitive load (Alhadad et al., 2018; Amarasinghe et al., 2021), potentially discouraging them from adopting them in their teaching practices. These challenges underscore the importance of prioritizing teacher involvement in the design process to maximize alignment between the LA dashboard and teachers' pedagogical intentions (Martinez-Maldonado, 2023). One way to enable teachers' participation in the LA design process is by actively involving stakeholders through design practices (R. Alfredo, Echeverria, Jin, Yan, et al., 2024).

As mentioned above, considerable attention has been paid to *design-for-use* in LA, such as generating ideas and prototyping with educators (e.g., Schmitz et al., 2022; Mohseni et al., 2023; G. M. Fernandez-Nieto et al., 2024; Wiley et al., 2024). Yet, this may inadvertently disregard the possibility of users acting as designers in practical usage situations (Nelson et al., 2009; Kim & Lim, 2023). Teachers can still shape the features in the system (instrumentalization) and change their teaching practices (instrumentation) to adapt to the designed system (Mohseni et al., 2023; Lawrence et al., 2023), which can be a crucial yet overlooked element in the design of an LA system. Specifically, this presents a gap where *design-in-use* has been underexplored in the teacher-facing LA dashboard.

2.3 Design-in-Use LA Dashboard

Some educational studies have explored the concept of design-in-use through *instrumental genesis* (e.g., Lonchamp, 2012; Martinez-Maldonado et al., 2018; Carvalho et al., 2019; Vágová, 2021). However, these studies have not explicitly addressed the intersection between design-in-use and LA use. The underexplored territory lies in understanding the design implications that can emerge when teachers use a multimodal LA dashboard in a real-world education setting. This motivates our central overarching question: **How do teachers refine and adapt an LA dashboard while it is *in use*?** Using instrumental genesis as a theoretical lens, we dissected this central research question by investigating two sub-questions covering instrumentalization and instrumentation.

From an instrumentalization perspective, while this process has been studied to design technology-rich physical and digital educational spaces (Carvalho et al., 2019; Lagrange & Richard, 2022), the design of LA dashboard artifacts presents unique challenges. From an AI and analytics perspective, designing smooth interactions between humans, who possess varied analytical expertise, and highly dynamic AI and analytics interfaces is difficult (Q. Yang et al., 2020). These user interfaces change depending on algorithm outputs and vary in complexity, making it challenging to envision, in *design time*, the interaction issues that can emerge during use. From an educational perspective, the design of LA interfaces needs to be aligned to the teacher's learning design, which often needs to be dynamically adjusted during an educational delivery according to the particular pedagogical needs of students and educational contexts (K. B. Yang et al., 2023; Martinez-Maldonado, 2023). As a result, our first research (sub)question is the following: **RQ1. How do teachers permanently and/or temporarily change the LA dashboard to support personalized teaching practices in the wild?**

From an instrumentation perspective, and as mentioned above, teaching can also be seen as a design process (Alhadad et al., 2018; Nicholson et al., 2022). From this perspective, teachers can be seen as designers of learning experiences, emphasizing the importance of deliberate struggle, iterative processes, and the use of data and analytics to inform decision-making. However, when teachers use an LA system, the dynamics of changes in teaching practices remain underexplored. Moreover, while the work of Martinez-Maldonado and colleagues (2018) provides some insights into the concept of collective instrumentation in the context of instructional design, the unintended uses and emerging collective teaching practices surrounding the instrumentation of an LA dashboard have not been previously investigated. This suggests a gap in our understanding and allows further investigation into how such unexpected appropriation could influence LA design practices. With this in mind, we pose our second research (sub)question: **RQ2. How do teachers change and adapt their teaching practices, both individually and collectively, while using an LA dashboard in the wild?**

3. Study Context

3.1 Healthcare Simulation in Nursing

This study was conducted as a part of the regular classes of a third-year nursing undergraduate course at a university in Australia in August and September 2023. In this course, a team-based healthcare simulation was a compulsory activity for these students to attend, scheduled for four weeks ($n = 56$ simulation sessions) starting in early August 2023. Each simulation session ran for an average of 20.53 min (Std Dev = 5.31 min) and involved four students, totalling 226 students. In each simulation, two students assumed the role of primary nurses (PN), a role akin to graduate nurses. The remaining two students acted as secondary nurses (SN), a role akin to ward nurses. After the simulation, students participated in a whole-classroom reflective debrief, a structured reflection process that involves reflecting on actions and on teamwork and communication performance, identifying practical knowledge gaps, and seeking solutions for improvements in the future (Cheng et al., 2014).

In debriefing, a co-teaching approach was adopted (Decuyper et al., 2023), with two teachers taking the roles of primary and secondary debriefers observing students' activity from a control room. The primary debriefer's responsibilities included annotating students' actions during the simulation, facilitating reflective discussions, and drawing attention to significant learning moments. The secondary debriefer played a supportive role, supplementing the primary debriefer's discussions with additional information and providing assistance to students who might be experiencing anxiety due to the stress of the clinical simulation. The other teachers role-played as a voice actor of a patient manikin (similar to G. M. Fernandez-Nieto et al., 2021) and a doctor during a simulation. A simulation had three main key events: the handover of the patient's information, the secondary nurse entering the room, and the doctor entering the room.

3.2 Sensor-Based Data Collection in the Wild

Per the teachers' standard protocol, every student was equipped with a position-tracking sensor and a personal microphone to capture each student's spatial and verbal data, respectively (see Figure 2.4a). The spatial data was collected using the *Pozyx Creator* toolkit. Each participant would be equipped with a waist bag containing a Pozyx positioning chip to capture their body orientation and spatial coordinates (x-y) at a frequency of around 1 Hz. In addition, each participant would also be equipped with a wireless headset microphone (*Shure PGA31*) to capture their verbal interactions. Our existing multimodal LA system (Martinez-Maldonado et al., 2023) analyzed the data collected from these devices and sent them as visualizations to the teacher-facing LA dashboard (detailed in Appendix A.1; see Figure 1). Teachers integrated the LA dashboard as a routine component of their classes. The study received ethical approval from the ethics committee of the university (MUHREC ID 39190).

3.3 LA System Design in Use

Five nursing teachers ($n = 5$) facilitated these 56 healthcare simulations with *DebriefPad Dashboard* for four weeks (see Figure 2.4b and 2.4c). We noticed that the originally designed version of the system met the main educational objective, but teachers pointed out aspects of the dashboard and their practices that could work better and required changes. The research team updated the system in the first week of use (after running $n = 16$ simulations), as a result of continued design-in-use, and reached a stable version for the remaining simulations for three weeks ($n = 40$ simulations).

As inspired by the work of Kaliisa and Dolonen (2023), feedback from teachers was collected to update the *DebriefPad Dashboard* in the first week of use. This feedback was collected through conversations with teachers at the end of each day ($n = 6$ sessions), which can often lead to more honest and spontaneous responses than formal interviews or surveys (Muller, 2002). The teachers' responses were promptly shared and discussed with the other researchers and software developers to update the system. To meet requested changes on functional requirements while maintaining the system's availability during data collection, our software developer employed development and operations practice (de Bayser et al., 2015) and deployed the updated system as soon as it was available the next day. This collaborative design approach was intended so that insights gained from the feedback were integrated into the implementation by maintaining the involvement of teachers' voices in the design, embracing their lived experience as a valid and significant form of expertise (Martinez-Maldonado, 2023).

Two researchers and a software developer discussed and reached a consensus to update the system, which was to change, for example, the *communication behaviour graph* and remove the "projector preview page." The communication behaviour graph is an AI-powered network graph visualization that uses LLMs to represent critical communication behaviours' co-occurrences by processing visually and classifying students' speech data (see Figure 4) (Zhao et al., 2024). The "projector preview page"

Figure 1. Two pages of *DebriefPad Dashboard*. The actions-tagging page (top) was used by the teacher when observing nursing students during a simulation to record the simulations' key events and students' clinical and teamwork actions. The projector-controller page (bottom) was used by the teacher during a debriefing to display analytics (i.e., visualizations) on a projector screen.

displays how the set of selected visualizations is arranged and looks before being presented on the co-located public screen (i.e., projector; see Figure 5). To ensure replicability, the LA system and dashboard features are detailed in Appendix A.2.

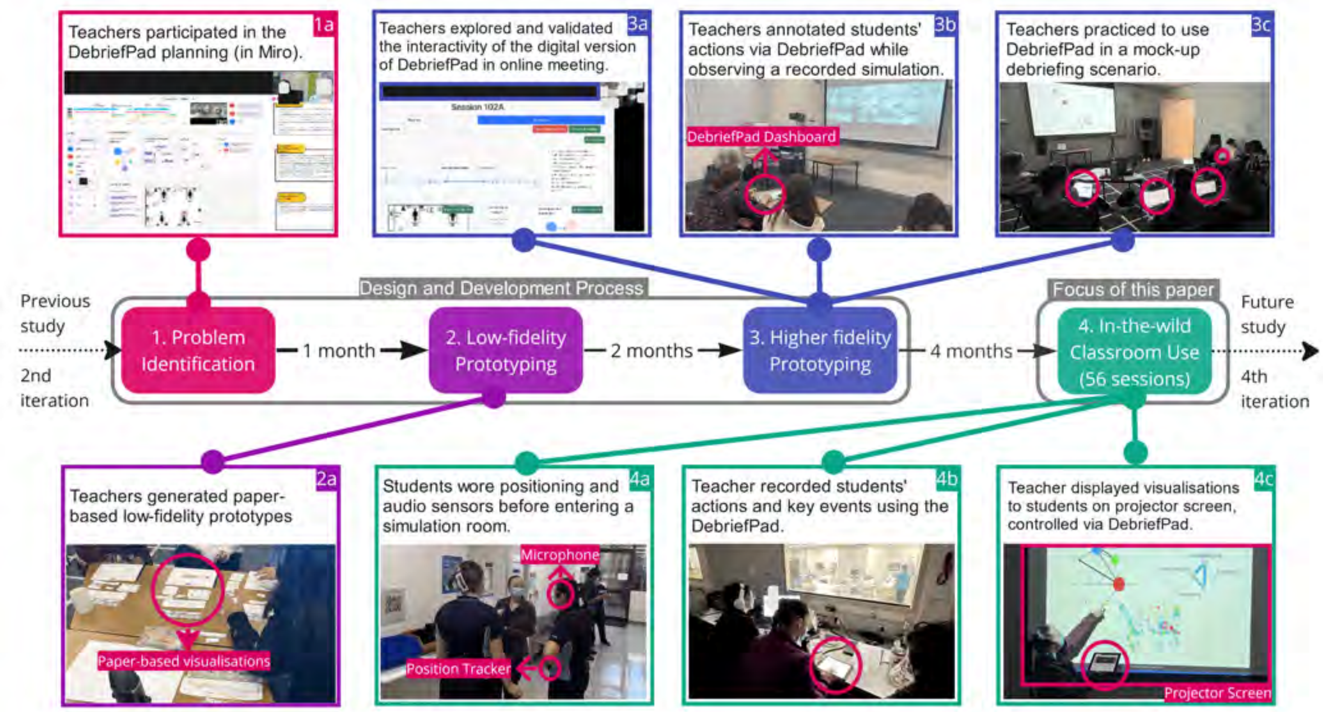

Figure 2. Design process and in use of the *DebriefPad Dashboard* with nursing teachers, adapted from the LATUX framework (Martinez-Maldonado et al., 2016).

4. Methods

This study has two parts of data collection and analysis: observation and interview (see Figure 3). In the first part, we captured evidence through researchers' observations on teachers' *in use* of the *DebriefPad Dashboard* in 56 debriefing sessions, facilitated by five teachers. In the second part, the evidence was triangulated and summarized to inform the post hoc interviews with the same teachers to address the RQs. Each part is detailed below.

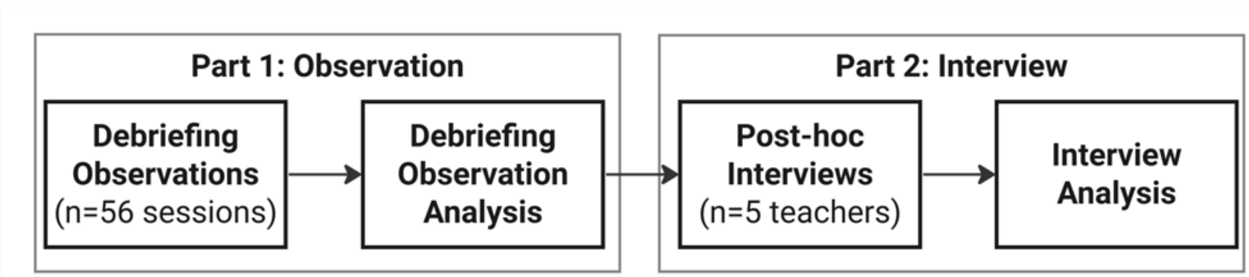

Figure 3. The diagram of this study's research methods. **Part 1:** Researchers observed teachers' 56 debriefing sessions while capturing the primary debriefer's audio recording and screens of the LA system (i.e., dashboard and projector) in "use time." **Part 2:** Addressing the RQs on the LA dashboard *in use* via post hoc individual interviews with five teachers.

4.1 Part 1: Debriefing Observation Data and Analysis

4.1.1 Debriefing Observation

We captured evidence from 56 debriefing sessions. The primary debriefer was equipped with an individual audio recording microphone, and the projector screen was recorded. In adherence to ethics requirements, we only had access to screens (i.e., projector and tablet) and teachers' voices. Hence, two researchers observed teachers' screens remotely and noted how teachers used the *DebriefPad Dashboard* in debriefing sessions (see Section 3.3, and Figure 2.4c), specifically in their interactions with the system.

4.1.2 Debriefing Observation Analysis: Triangulation for In-Use Analysis and Feature Evolution

Inspired by the work of Tessier (2012), this study triangulated evidence that combined three researchers' observation notes with screen and audio recordings from each teacher during the debriefing session, aiming to minimize researchers' interpretation bias and ensure the study's completeness and validity. The outcome of this analysis was to identify changes in use of the *DebriefPad Dashboard* that teachers would be asked about in the post hoc interview.

After completing four weeks of in-the-wild classroom use, one researcher transcribed audio recordings from each teacher's debriefing ($n = 5$ simulations, one session per teacher) using an AI-based automated transcription tool (Whisper[1]). A second researcher cross-checked the transcriptions to ensure their validity. Screenshots of the visualizations on the recorded projector screens were also taken to explore how teachers personalized how they shared visualizations with students on a projector screen. This approach allowed us to capture and analyze the teachers' verbal interactions with the system to provide an understanding of the systems used in debriefing.

Two other researchers helped triangulate and summarize all observed changes (Table 1), describing each feature's goal from intended use (*design-for-use*) and unexpected outcome (*design-in-use*). Hence, this approach enabled us to identify changes in the features and goals of the LA dashboard over time in the wild. This list of changes was used as input in the follow-up interview to validate the changes, as explained below.

Table 1. A list of LA features, each with its intended use (*design-for-use*) within the *DebriefPad Dashboard*, that had been shaped into new, unexpected design outcomes (*design-in-use*).

No.	Goal	Design-for-use (design time)	Design-in-use (use time)
1	To identify the team's critical communication behaviours	Using six nodes of *Communication Behaviour Graph* (see Figure 4, left image)	The visualization was requested to be simplified from six to four nodes, with the "Handover" and "Sharing Information" labels being removed.
2	To preview visualizations before sharing them on a projector screen	The "projector preview page" was designed to allow teachers to review the combination of visualizations on a handheld device before sharing them to the projector screen.	The "projector preview page" was requested to be removed in the final version to improve its ease of use.
3	To provide evidence about prioritization, teamwork, and communication	Supported by AI-powered LA visualizations, teachers could provide evidence of task prioritization, teamwork, and communication from multimodal data (i.e., positioning and audio).	The visualizations were used to regulate students' emotions and motivate them.
4	To gain insights from visualizations to support students with complete visualizations for reflection	The system was used only after all data was fully processed, ensuring a comprehensive view. This strategy anticipated using the system primarily toward the end of the debriefing.	A practice was developed to share with students in the middle or at the start of debriefing, as soon as data was available based on the order of data visualization being processed.
5	To be used only by the primary debriefer during simulation and debriefing	The primary debriefer held the device the whole time by tagging actions while observing a simulation and controlling the visualizations on a projector during debriefing.	Initially, the primary debriefer tagged students' actions while observing a simulation and then passed the device to the secondary debriefer for debriefing. Later, the device was reassigned for individual use (primary debriefer) for the rest of the study.

4.2 Part 2: Interview Data and Analysis

4.2.1 Post Hoc Interviews

Five teachers ($n = 5$) were invited to the post hoc individual semi-structured interviews (avg. 65 mins) (Brinkmann, 2020). All teachers were female and had more than seven years of experience teaching simulation-based learning activities. They had been

[1] https://openai.com/research/whisper

engaged in a three-year partnership with the research team to develop innovations to support reflection in healthcare simulation. Each teacher was compensated with an A$60 gift card for their participation in the interview.

These interviews aimed to draw inferences from the experiences of teachers and researchers, thereby providing evidence to answer the research questions. One researcher conducted all interviews virtually via a video conferencing tool (i.e., Zoom). The interview began by asking teachers about their general experience with the *DebriefPad Dashboard*, including the benefits and challenges of using the system. To address our research questions, semi-structured interview questions were developed focusing on the design-in-use listed in Table 1, as previously analyzed. For **RQ1**, the interview questions focused on permanent instrumentalization, asking about the interface and interaction design, such as the reasons to request change of labels in the communication behaviour graph and removal of the preview screen page. Additionally, questions on temporary instrumentalization explored teachers' reasons for occasionally using the system differently than intended. For **RQ2**, the questions addressed individual instrumentation by investigating how teachers develop their practice using visualizations whenever data is available. Furthermore, questions related to collective instrumentation examined the collaborative use of the *DebriefPad Dashboard*, specifically the reasons for delegating control from primary to secondary debriefers in the first week of use, but later reverted to individual use. Specific interview questions are detailed in Appendix A.3 to ensure replicability.

4.2.2 Interview Analysis

We conducted the five phases of **deductive-inductive qualitative analysis** on the post hoc interview data (Bingham, 2023). In phase 1, the second author started organizing the transcripts of all individual interviews by tagging quotes related to questions asked in the interviews. In phase 2, the first author developed topic codes aligned to research questions (i.e., *instrumentalization* and *instrumentation*) and deductively tagged organized quotes. In phase 3, the first author conducted inductive analysis through open coding to identify emerging ideas in the data (e.g., "maintain a conversation with students," "comments on handover label," and "evolution in teaching practice").

In phase 4, themes were inductively developed on each topic based on these open codes to answer each research question (e.g., "change system's goal to motivate students," "changes in communication behaviour graph," and "tool was more integrated to teaching practice"). In phase 5, these themes were tagged and organized according to theory codes on instrumental genesis (such as "temporary instrumentalization" and "individual instrumentation"; see Section 2.1) and presented in the results section to answer research questions. The results and codebook were cross-checked by a second researcher and jointly discussed openly to improve the reliability of results (McDonald et al., 2019). Any discrepancies were extensively consulted on with the third researcher to reach a consensus.

5. Results

In this section we present the resulting themes organized according to instrumental genesis theory, illustrating temporary and permanent instrumentalization (**RQ1**) and individual and collective instrumentation (**RQ2**).

5.1 Evolution of the LA Dashboard (RQ1)

5.1.1 Simplifying the Communication Behaviour Graph to Facilitate Students' Comprehension

This feature is an AI-powered visualization displayed in the dashboard intended to identify the team's closed-loop communication via an epistemic network graph (see Figure 4). All teachers said that simplifying the labels from six nodes to four was necessary to enhance the system's effectiveness in helping teachers cover the team's closed-loop communication. During *design-in-use*, all teachers found that the two labels, "Handover" and "Sharing Information," could cause "confusion" (P2, P4), "distraction" (P3, P5), and "too much to talk about"(P1), which could be beneficial to remove. This resulted in *permanent instrumentalization* due to a modification to the system.

Three teachers (P2, P4, and P5) found the "Handover" label to be redundant information in the visualization. For example, two teachers said that it was already part of the simulation scenario, noting: "Providing handover to the students may not be beneficial as they already know it will happen" (P2, P5). Similarly, the label "Sharing Information" was discovered to overlap with other labels during in-use. For instance, P5 explained it as another piece of information that may add complexity and distraction. In sum, teachers' inputs highlight their role as designers *in-use*, capable of refining the way the algorithmic outputs are presented to better align with their pedagogical intentions.

All teachers also expressed concerns about students who may experience information overload. After we applied permanent changes to the design of this visualization, P1 and P2 expressed appreciation for the clarity and digestibility of the in-use version for students, which contained only four labels selected by the teachers. P5 further differentiated between the perspectives of teachers and students, stating: "Six nodes were about what we [teachers] wanted to see and the four-node version is about what the students want to see." This change became more visible to teachers only after they started to use the LA dashboard to conduct actual debriefs. Hence, this highlights the need to balance providing comprehensive information and maintaining clarity and digestibility for both students and teachers.

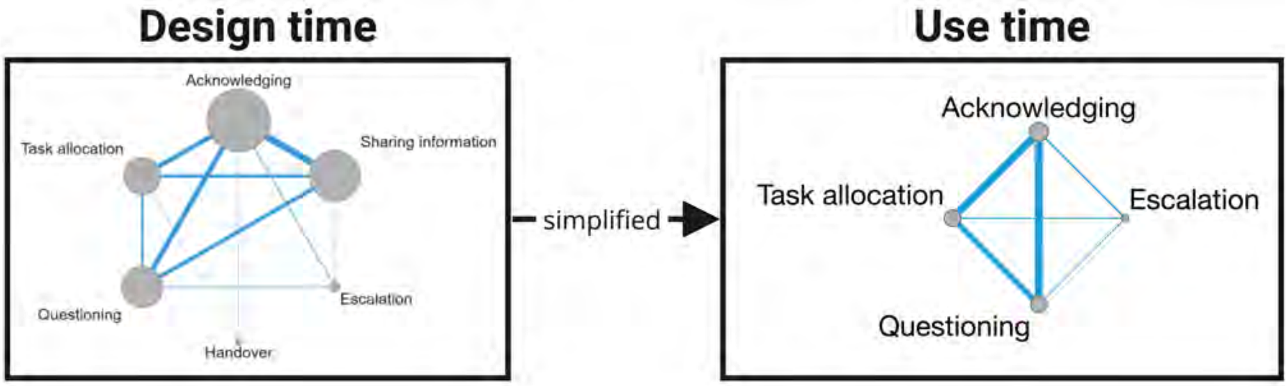

Figure 4. Left: A complete six-node communication behaviour graph as conceptualized in design time. Right: "Handover" and "Sharing Information" labels were removed from the communication behaviour graph in "use time."

5.1.2 Removing the Projector Preview Page to Decrease the Complexity of Interaction

This overlay window page, which provided a preview of their selected visualizations, was initially designed by teachers to "prevent errors during presentations to students" (P1, P3, P5) and "provide reassurance before displaying content" (P2, P4, P5) (see Figure 5). For instance, P5 elaborated that having this feature could give them a sense of psychological safety, reducing uncertainty and risk before presenting to students, stating:

> P5: *As someone who was a part of the co-design but had never used it before [in actual debriefing], I felt uncertainty. Having a preview page allows me to gain some confidence and certainty before I put myself on the line to put it on the projector screen for students to see.*

However, after scrutinizing the necessity and effectiveness of this feature in-use, teachers requested the permanent removal of the preview feature. This modification resulted in *permanent instrumentalization* of the system's interface and its use. This was realized after teachers used the LA system in the wild. Three teachers (P2, P3, and P5) explicitly emphasized that the projector preview page, initially requested during design, was unnecessary. These three teachers said its removal was necessary during in-use because visualizations were already available at the bottom of the screen, which served as a preview for teachers. As P2 stated: "I would go through and look at different key events within the simulation at the bottom of the screen before I put it up. So, I kind of use that as a preview feature in itself."

Moreover, related to this, four teachers (P2–P5) further elaborated on the need for efficiency and the desire to maintain the flow of conversation with students, maximizing direct communication with them while minimizing screen time. For example, they stated: "I did not want to face down in the iPad all the time while talking with students" (P3), and "As I am getting familiar with the tool, I want to get it up there [display to the projector screen] as quickly as possible while I can keep talking [with students]" (P5). As a result, teachers found a better way to use the LA dashboard that they had not fully envisaged before in design time. This led to a re-evaluation of the "projector preview page," ultimately leading to permanent changes for more efficient and streamlined data sharing with students.

5.1.3 Emergence of an Unintended Use of the Dashboard: Helping to Provide Emotional Support to Students

All teachers ($n = 5$) adjusted the system's original capabilities, which were designed to provide behavioural evidence only regarding task prioritization, teamwork, and communication dynamics, to instead help teachers provide emotional support to students. This change was motivated by teachers noticing students' self-criticism after completing the healthcare simulation. Teachers found that students "felt down and harsh on themselves" (P1, P2, P5), were "constantly judging themselves" (P3, P4), and "felt the experience quite confronting" (P2). Despite no structural changes in the system being made as a result, this change of goal resulted in modifying the system's capabilities, and it can be considered as *temporary instrumentalization*.

This can be illustrated as follows. All teachers ($n = 5$) wanted to give their students a positive experience for completing the simulation. They believed that this simulation was a crucial part of their students' careers in nursing, and they used the system to prioritize giving them a positive experience. While acknowledging the importance of constructive criticism through data, two teachers (P1 and P2) focused on the emotional aspect of learning. They emphasized that the visualization graphs and analytics can boost students' morale and deliver a positive takeaway from the simulation, for example: "From an education point of view, I probably should have pointed out what they could have done better, but I was just trying to provide emotional support or motivate them" (P1) and "When they [the students] can actually see it up there, in terms of graphs and analytics, it can make them feel good because it was real. I want to give them a positive experience, too" (P2).

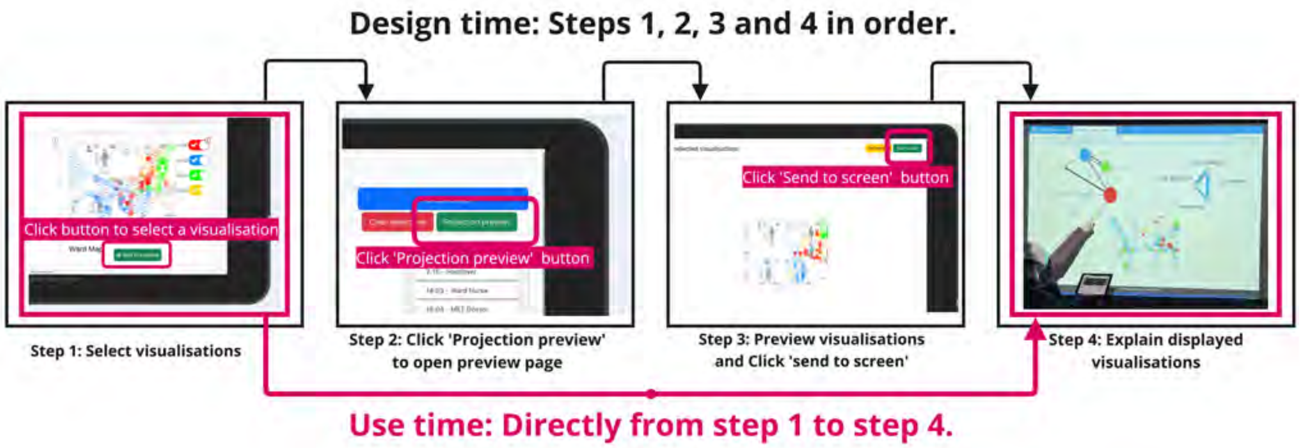

Figure 5. In step 1, the teacher (debriefer) selects one or more visualizations. In step 2, the teacher clicks the preview button. In step 3, the interface shows how the students will look at visualizations on the public screen (projector) and click "send to screen" to display a set of visualizations on the projector. In step 4, the teacher debriefs students with the assistance of LA visualizations. In *use time*, steps 1 and 4 remained while steps 2 and 3 were removed.

Moreover, P4 highlighted the importance of controlling what data to present with the purpose of regulating students' emotions. If a team's performance was poor, P4 would focus on the positive aspects and switch to a different visualization if the discussion was causing distress. In other words, P4 was willing to question the perceived reliability of the LA dashboard if it meant reducing distress among the students. As P4 described:

> P4: *If an issue arose during our debriefing conversation that indicated a poor team's performance, I would focus on the well-handled aspects instead [...]. If I sensed that the discussion was causing distress, I might blame the technology, suggesting it [the AI-powered system] could have misinterpreted the situation.*

As a result, the temporary instrumentalization of the LA dashboard meant that the visualizations they decided to display and withhold from students were used for the unenvisaged purpose of supporting students' emotions by emphasizing the positive aspects of students' learning activity. This capability of the system emerged in-use and overrode some of the original design intents.

5.2 Co-evolution of Teaching Practices (RQ2)

5.2.1 Changes in Individual Debriefing Practices

In our previous study iteration, all teachers had experience using analytics to support debriefing toward the end of the session. In this study, the data processing module was optimized so that visualizations could be available as early as the start of debriefing. Through interviews, we identified three examples of *individual instrumentation*, each reflecting a different way that teachers engaged with the LA dashboard to prepare and share the information with the class in debriefing.

Example 1: Teachers' use of the LA dashboard to validate discussions. This way of using the LA dashboard was marked by minimal system usage, where the teacher did *not* integrate the visualizations from the start of the debrief. Visualizations were only shared with students toward the end of the debriefing session. Only one teacher (P1) demonstrated this instrumentation level, and its usage was somewhat "reactionary." As P1 explained: "I would use the visualizations to back up what we [as teachers] had already said, and that happened at the end. I just went through all the visualizations one by one. If it was inaccurate [did not reflect on what happened], I just said there was an issue with the analytics and moved on." This statement suggests that P1 used the visualizations more as a validation tool for discussions that had already taken place, suggesting that P1's usage of the system might have had less integration and impact on existing debriefing practice.

Example 2: Teachers' use of the LA dashboard to prepare debriefing and share information. In this case, teachers actively monitored visualizations on the handheld LA dashboard as the data was processed and readily available. Once visualizations were available, they did not immediately share them with their students through the projector screen. Instead, they mainly used it to prepare and inform their debriefing. This could be exemplified by three teachers (P2, P3, and P4), and two practice examples emerged, which were to "preview visualizations to prepare the debriefing" (P2, P3) and "verify that the LA system was processing the data" (P4). For instance, P2 developed a practice of looking at the visualizations before entering the debriefing room to help understand the summary of the team's performance. Similarly, P3 exemplified the use where the teaching practice would be informed by the analytics to further enrich their debriefing narrative without necessarily changing the usual debriefing procedure, stating: "As soon as visualizations were available on my screen [or iPad], I made connections between analytics and what I observed [...]. Then, I would consider how to put it up in the middle or near the end of my debriefing."

In contrast, P4's approach was slightly different—they closely monitored the status of the LA dashboard to ensure that the data was being processed and served as a reassurance that the data would eventually be available for review. As P4 explained: "I think possibly we could see something was happening. They'll get it eventually rather than going to debriefing without knowing if it was working or not. If it is not coming up, I will continue debriefing without it." This statement highlights how the system's capabilities provided P4 with the flexibility to dynamically decide on the debriefing approach for students, with or without data. Interestingly, P4 exhibited little reliance on the LA system when unavailable. These uses of the LA dashboard exemplified partial integration and could signify moderate changes in their debriefing.

Example 3: Teachers' use of the LA dashboard as visual cues of student-led discussion. In this case, only one teacher (P5) referred to the visualizations throughout the full facilitation of student-led discussions. This resulted in the teacher's reliance on the system in their debriefing practice as their use became fully integrated into the debrief. As P5 explained:

> *There would probably be an evolution in how we use the visualizations because initially, we would get halfway through the debrief and then try and use them all. Then, I found out I could start to talk through them almost in the order that visualizations came through [sequentially in order of teamwork prioritization, ward map, communication network, and communication behaviour graph], right from the start. I could use it as a visual cue for students' conversations. I think it organically aligned and shaped my narrative in debriefing.*

This suggests that P5 strategically used the visualizations to develop and change their debriefing narratives right from the start of the debriefing session. P5 also explained the reliance on the data for personal debriefing practice, in which the extent of integration of the LA dashboard can also suggest the risk of over-reliance on the analytics once they become a critical part of people's practices. This use of the LA dashboard exemplified full integration and could indicate significant changes in the debriefing practice.

5.2.2 Changes from Individual to Collective Debriefing Practices and Revert to Individual Use

Although the system was originally designed to be used by teachers individually (by the primary debriefer), they changed their practice from individual to collective use, resulting in *collective instrumentation*. All teachers reflected on the in-use challenges that they encountered for using it individually, such as "familiarisation with the system and its visualisations" (P1, P2, P3, P5) and "difficulties in balancing the learning of the system with the debriefing procedure" (P3, P4, P5). Teachers agreed to use the system in pairs to address these challenges and did it on their first week of use. The secondary debriefer often controlled the visualizations, and the primary debriefer led the debriefing and focused on the conversation with students.

Two teachers, P2 and P5, highlighted the benefits of collective instrumentation. P5 shared their experience that collective practice can be beneficial for individual learning, suggesting that collective use of the system enhanced the proficiency of teachers in using the LA dashboard. Similarly, P2 explained how they learned together in the first week, suggesting an emergent structured approach to collective use and mutual training, where each member had a specific role contributing to the overall co-teaching effort: "The primary debriefer took notes and watched the secondary debriefer utilize the tool during debriefing. So, I think the main focus of that primary debriefer was just to become familiar with the tool and how that would be incorporated into the debrief later."

The dynamics between the primary and secondary debriefers in the system's collective instrumentation resulted in several challenges, as described by four teachers (P1, P3, P4, P5). P4 described "a bit of back and forth conversation," indicating a need for constant communication and adjustment between two teachers. On the other hand, P1 raised an issue about diverting from the intended secondary debriefer's role by observing "a divided attention of the secondary debriefer who should have monitored students for a sign of discomfort or stress." P3 and P5 shared experiences where the visualizations chosen by the secondary debriefers differed from what they had in mind, forcing them to find alternative explanations. In the end, all teachers concurred that they regained confidence in using the system individually with sufficient time. P5's statement, "As soon as our familiarity and trust in using the tool grew quite rapidly, it did not take very long for us to bring it back to primary debriefers to use the tool," underscored a shift to individual use after collective instrumentation. This shift suggests that teachers were growing confident in their ability to interpret and apply data individually in their debriefing.

In summary, these insights underscored the dynamic nature of instrumentation and the critical importance of collective instrumentation, especially in co-teaching practices. Once a collective utilization scheme emerged, teachers could tap into it to individually use the LA dashboard with more confidence.

6. Discussion

6.1 Instrumentalization (RQ1)

Responding to RQ1, we found that permanent and temporary changes in the LA dashboard in-use were critical for the artifact to be more effectively used by nursing teachers. These changes emerged as a result of actually using the LA dashboard. The communication behaviour graph, created using LLM algorithms, underwent **permanent instrumentalization** to simplify some of the details that were included in the original design, such as potentially overlapping domain-specific information (i.e., "Handover" and "Sharing Information"). These details were only identified by teachers during use and after enough debrief repetitions. This suggests that certain features of the system's design can be overlooked in *design-time* (Knowles et al., 2019).

In interpreting the algorithm's output, teachers had multiple interpretations of the labels and selected those they deemed important. Teachers preferred to use partial or incomplete information in the communication behaviour graph (four nodes in the in-use graph) compared to its full version (six nodes) designed before the study. This selection process, however, may be problematic because it reduces the context. The simplification of the labels resulted in very similar networks across teams, which was ultimately a problem. This issue underscores the complexity and challenges of interpreting algorithmic outputs in an AI-powered LA system by teachers, who commonly have limited analytics expertise (Fernandez Nieto et al., 2022; R. D. Alfredo et al., 2023). This example also aligns with the ongoing discussions on data literacy in LA (Dollinger & Lodge, 2018; Martinez-Maldonado, 2023; Kaliisa & Dolonen, 2023; Echeverria et al., 2024), suggesting a need for careful consideration of data transparency and explainability in the use of analytics and AI-powered visualizations.

Moreover, during the *design-for-use* phase, teachers deemed the projector preview page necessary. The need for this feature could be primarily driven by the teachers' mistrust of the data in the LA system, which is a common issue found in other LA projects (Slade et al., 2023). However, teachers faced time constraints during the debrief that they had not initially anticipated (Kaliisa & Dolonen, 2023). As the teachers continued to use the system, they began to trust the data more. This led to the understanding that advanced features were unnecessary and became a hurdle in delivering effective teaching practices. This led to permanent changes in the interface, similar to the study of Martinez-Maldonado and colleagues (2018), when teachers changed the content displayed on different surfaces to support better collaboration. Therefore, this permanent instrumentalization highlights the importance of understanding the potential *time constraints* and *trust issues* teachers may face in designing AI-powered LA systems during in-use time.

In contrast, *temporary instrumentalization* was exemplified when our teachers used the LA dashboard to motivate students and shared its control with the secondary debriefer in their first week of use. These two examples not only illustrate the need to add flexibility to the LA system features but also show that when teachers are given agency (Kaliisa & Dolonen, 2023), they can have more control to shape the original intended system's goal. By empowering teachers with agency and control, the LA dashboard can become a dynamic tool that can be adapted to meet unique pedagogical needs, where the flexibility not only enhances the system's effectiveness but also can foster a sense of ownership among teachers, thereby promoting meaningful and impactful use of AI and analytics in education (Buckingham Shum & Luckin, 2019; Holstein et al., 2019; R. Alfredo, Echeverria, Jin, Yan, et al., 2024). This finding underscores the importance of designing LA systems that are adaptable, that can be designed in-use, and that can cater to the dynamic nature of teaching practices (Moltudal et al., 2022).

6.2 Instrumentation (RQ2)

Responding to RQ2, we observed that teachers adapted their teaching practices individually and collectively while using the AI-powered LA system in the wild. These adaptations are illustrated through described changes in the ways teachers created and adapted their utilization schemes through co-evolving teaching practices around the use of the LA dashboard.

Individual instrumentation can uncover how teachers adapt their teaching practices to the system's capabilities whenever the data is available. The examples provided in the results section illustrate how instrumentation happened differently for different teachers when appropriating the analytics. What is particularly noteworthy is that teachers were not merely passive recipients of the system's analytics (Kim & Lim, 2023); they were active designers who used the analytics to inform and shape their teaching practices, both individually and as a teaching team. This is evident in *instrumentation* related to using the LA dashboard informing teachers' debriefing only. The emerging new practice of previewing data in preparing for debriefing did not involve showing the data directly to students. In contrast, the deeper appropriation by one of the teachers who fully integrated the use of the LA dashboard in the debriefing can completely change students' learning experience by transforming those debriefs into data-intensive reflection sessions. However, a significant reliance on the system can pose challenges, especially in the latter case. Although LA dashboards can provide valuable insights, over-reliance on these systems can limit teachers' professional judgment and autonomy (Kaliisa & Dolonen, 2023; Lawrence et al., 2023; Echeverria et al., 2024). This underscores the importance of balancing algorithmic automation and agency when designing and using an AI-powered LA system (Martinez-Maldonado, 2023).

Additionally, when primary debriefers intuitively decided to share control of the system with the secondary debriefer to reduce the workload, it demonstrated *collective instrumentation*. This could be seen as one of the pieces of evidence of peer

ISSN 1929-7750 (online).

training that was suggested by teachers in the work of Kaliisa and Dolonen (2023). However, the potential of these strategies to improve data literacy and their integration into teachers' professional development needs further study (Decuyper et al., 2023). Moreover, teachers discovered that learning to use the system together built trust in data usage. However, designing trustworthy human-centred LA may have overlooked this collective practice as a viable solution, which requires further investigation. This underlines the need to consider how groups of users, not just individuals, will use human-centred LA collectively in authentic educational settings (Buckingham Shum et al., 2024).

6.3 Implications for Research and Practice

6.3.1 Implications for LA Researchers and Designers

The results from this study illustrate the importance of continuing the design of LA systems in-use. Although the *DebriefPad Dashboard* was designed with the active participation and decision-making of nursing teachers, which has been highlighted as an important practice in current LA research and development (Dollinger & Lodge, 2018; Martinez-Maldonado, 2023; Buckingham Shum et al., 2019, 2024), both the artifact and the teaching practices needed to co-evolve in-use. The design-in-use of our LA system revealed how design can address contemporary educational issues, such as the need to ensure teachers' agency in data-intensive educational systems (Kloos et al., 2022; Kaliisa & Dolonen, 2023; R. Alfredo, Echeverria, Jin, Swiecki, et al., 2024).

Additionally, we suggest that LA researchers and designers report the design process in "use time" and not only what happens in "design time," especially from direct stakeholders who have contributed to the system's design. Through our illustration using an *instrumental genesis* lens in the LA dashboard, we highlighted that this approach could be a supplementary method to the iterative design of an interactive LA system. Other LA researchers and designers may build on this study's method to identify technological integration challenges (*instrumentalization*) and co-evolution of pedagogical practices (*instrumentation*) when closing the analytics loop providing end-user LA interfaces, particularly when working with complex data and AI-driven analytics. This approach can also enable the systematic reporting of in-the-wild studies, specifically regarding the evolution of the artifact and the subjects.

6.3.2 Implications for Educational Practitioners

One of the key aspects of supporting a teacher's agency is understanding the dynamics of power relationships in the educational institution (Lee & Gargroetzi, 2023; Dollinger & Lodge, 2018). Typically, teachers do not have the authority to make final decisions in the system, which is a crucial consideration for educational decision-makers (e.g., educational directors, instructional designers, school administrators, and IT academic leaders) (Friend Wise et al., 2023). Our findings highlight that the design process should aim to introduce tools that can be continually used and improved, not merely imposed on teachers. In this context, teachers can still play a pivotal role in decision-making within educational institutions as designers (Paniagua & Istance, 2018). Even though they may not be able to make permanent changes to the LA software system (*instrumentalization*), they can still be part of the changes in pedagogical practices when using the system (*instrumentation*) and make suggestions on the configuration of the LA tools. It is vital to raise awareness that while teachers may not always accept an initial design, they can still participate in the design process during "use time" and potentially improve the chances of adopting a new LA innovation. Therefore, this paper contributes with a theoretical and methodological lens to the calls from the recent two systematic literature reviews in LA and AI in education calling for the active involvement of educational stakeholders, including teachers, as active participants in the design process (R. Alfredo, Echeverria, Jin, Yan, et al., 2024; Topali et al., 2024).

6.4 Limitations and Future Work

This paper has several limitations. First, we acknowledge that our study is contextualized within a nursing simulation context in higher education. As such, the insights derived from this research may not be directly generalizable to other teaching contexts (e.g., K–12). As this paper focuses on qualitative design research and illustrates the concept of design-in-use within a nursing simulation, the findings from the design research can still be transferable, offering valuable principles or frameworks that other researchers and practitioners can adapt to similar contexts. Future studies could explore how these insights apply to different contexts beyond simulation in higher education. For instance, future work may explore how the practical application of *design-in-use* impacts teacher-student relationships, considering the unique student data available in other educational contexts.

The next limitation relates to the sample size. Five participants were involved in the post hoc interviews since they were the main teachers of the simulations considered in the study. Yet, this reflected the authenticity of our study. We prioritized depth of qualitative insights over data quantity, in which a smaller and more engaged group often yields more valuable insights than a larger random group. Besides, these participants were not randomly selected since they participated and were integral to our design process, as was done in the work of Knowles and colleagues (2019). Another limitation is this study did not delve into the varying levels of LA literacy among the teachers involved. The ability to engage with and modify such systems may be influenced by one's understanding and proficiency in LA (Alhadad et al., 2018). However, given the specific focus of our study, a comprehensive exploration of this aspect was beyond its scope. This limitation underscores the need for future research to investigate to what extent teachers' LA literacy may influence their ability to contribute to the design of LA systems.

Moreover, while our research has provided insights into the design process and the usage of the LA system from the nursing teachers' perspective, it is important to acknowledge that students' voices were not included in this study. This focus was intentional, as the LA system was a teacher-facing LA dashboard, and our primary aim was to understand the design-in-use from the teachers' perspective. However, designing LA systems in higher education without students might be counterproductive or harmful (Lee & Gargroetzi, 2023). We acknowledge that the *collective instrumentation* process, particularly the dialogues between students and teachers about visualizations that were being displayed on the shared projector screen, may hold significant potential for enriching our understanding of the dynamics of teachers' and students' interactions through LA. The analysis of such conversations could have sparked further design ideas for co-creating human-centred LA with teachers and students (Dollinger & Lodge, 2018). Due to ethical considerations, and as requested by the nursing teaching staff, our protocol did not permit the recording of students' conversations during the debrief sessions to protect their privacy and ensure that their participation was voluntary and free from surveillance concerns. As a result, we see this as an area for potential exploration in future studies.

7. Conclusion

This study illustrated that LA design is not a static process but a dynamic endeavour that evolves with use. Using the lens of instrumental genesis, encompassing both *instrumentalization* and *instrumentation* dimensions, we delved into the design-in-use on the co-evolution of the artifact (an LA dashboard) and the subject (teachers' teaching practices). Therefore, the examination of design-in-use is crucial in the design process of an LA system. It not only ensures that the AI LA technology effectively fulfills its intended purpose but also facilitates systematic reporting of the reflection on the continuous improvement of both the system and users. This study provides evidence on the exploration of design-in-use in the utilization of a teacher's LA system, thereby highlighting the need for further research in this area.

8. Declaration of Conflicting Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

9. Funding

Riordan Alfredo gratefully acknowledges Monash University for his PhD scholarship. This research was funded partially by the Australian Government through the Australian Research Council (project numbers DP210100060, DP220101209, DP240100069), Digital Health Cooperative Research Centre, and Jacobs Foundation (Research Fellowship).

References

Aleven, V., Blankestijn, J., Lawrence, L., Nagashima, T., & Taatgen, N. (2022). A dashboard to support teachers during students' self-paced AI-supported problem-solving practice. In I. Hilliger, P. Muñoz-Merino, T. De Laet, A. Ortega-Arranz, & T. Farrell (Eds.), *Educating for a New Future: Making Sense of Technology-Enhanced Learning Adoption* (pp. 16–30). Springer. https://doi.org/10.1007/978-3-031-16290-9_2

Alfredo, R., Echeverria, V., Jin, Y., Swiecki, Z., Gasevic, D., & Martinez-Maldonado, R. (2024). SLADE: A method for designing human-centred learning analytics systems. In *Proceedings of the 14th International Conference on Learning Analytics and Knowledge* (LAK 2024), 18–22 March 2024, Kyoto, Japan (pp. 24–34). ACM. https://doi.org/10.1145/3636555.3636847

Alfredo, R., Echeverria, V., Jin, Y., Yan, L., Swiecki, Z., Gašević, D., & Martinez-Maldonado, R. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers and Education: Artificial Intelligence*, *6*, 100215. https://doi.org/10.1016/j.caeai.2024.100215

Alfredo, R. D., Nie, L., Kennedy, P., Power, T., Hayes, C., Chen, H., McGregor, C., Swiecki, Z., Gašević, D., & Martinez-Maldonado, R. (2023). "That student should be a lion tamer!" StressViz: Designing a stress analytics dashboard for teachers. In *Proceedings of the 13th International Conference on Learning Analytics and Knowledge* (LAK 2023), 13–17 March 2023, Arlington, Texas, USA (pp. 57–67). ACM. https://doi.org/10.1145/3576050.3576058

Alhadad, S. S. J., Thompson, K., Knight, S., Lewis, M., & Lodge, J. M. (2018). Analytics-enabled teaching as design: Reconceptualisation and call for research. In *Proceedings of the Eighth International Conference on Learning Analytics and Knowledge* (LAK 2018), 7–9 March 2018, Sydney, Australia (pp. 427–435). ACM. https://doi.org/10.1145/3170358.3170390

Amarasinghe, I., Hernández-Leo, D., & Ulrich Hoppe, H. (2021). Deconstructing orchestration load: Comparing teacher support through mirroring and guiding. *International Journal of Computer-Supported Collaborative Learning*, *16*(3), 307–338. https://doi.org/10.1007/s11412-021-09351-9

Baldauf, M., Adegeye, F., Alt, F., & Harms, J. (2016). Your browser is the controller: Advanced web-based smartphone remote controls for public screens. In *Proceedings of the Fifth ACM International Symposium on Pervasive Displays* (PerDis 2016), 20–26 June 2016, Oulu, Finland (pp. 175–181). ACM. https://doi.org/10.1145/2914920.2915026

Becerra, Á., Mohseni, Z., Sanz, J., & Cobos, R. (2024). A generative AI-based personalized guidance tool for enhancing the feedback to MOOC learners. In *Proceedings of the 2024 IEEE Global Engineering Education Conference* (EDUCON 2024), 8–11 May 2024, Kos, Greece (pp. 1–8). IEEE. https://doi.org/10.1109/EDUCON60312.2024.10578809

Bingham, A. J. (2023). From data management to actionable findings: A five-phase process of qualitative data analysis. *International Journal of Qualitative Methods*, *22*, 16094069231183620. https://doi.org/10.1177/16094069231183620

Brinkmann, S. (2020). Unstructured and semistructured interviewing. In P. Leavy (Ed.), *Oxford handbook of qualitative research*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190847388.013.22

Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-centred learning analytics. *Journal of Learning Analytics*, *6*(2), 1–9. https://doi.org/10.18608/jla.2019.62.1

Buckingham Shum, S., Martínez-Maldonado, R., Dimitriadis, Y., & Santos, P. (2024). Human-centred learning analytics: 2019–24. *British Journal of Educational Technology*, *55*(3), 755–768. https://doi.org/10.1111/bjet.13442

Buckingham Shum, S. J., & Luckin, R. (2019). Learning analytics and AI: Politics, pedagogy and practices. *British Journal of Educational Technology*, *50*(6), 2785–2793. https://doi.org/10.1111/bjet.12880

Carvalho, L., Martinez-Maldonado, R., & Goodyear, P. (2019). Instrumental genesis in the design studio. *International Journal of Computer-Supported Collaborative Learning*, *14*(1), 77–107. https://doi.org/10.1007/s11412-019-09294-2

Cerratto Pargman, T., Nouri, J., & Milrad, M. (2018). Taking an instrumental genesis lens: New insights into collaborative mobile learning. *British Journal of Educational Technology*, *49*(2), 219–234. https://doi.org/10.1111/bjet.12585

Chen, S., Liu, Y., Lu, R., Zhou, Y., Lee, Y.-C., & Huang, Y. (2022). "Mirror, mirror, on the wall"—Promoting self-regulated learning using affective states recognition via facial movements. In *Proceedings of the 2022 Designing Interactive Systems Conference* (DIS 2022), 13–17 June 2022, online (pp. 1300–1314). ACM. https://doi.org/10.1145/3532106.3533500

Cheng, A., Eppich, W., Grant, V., Sherbino, J., Zendejas, B., & Cook, D. A. (2014). Debriefing for technology-enhanced simulation: A systematic review and meta-analysis. *Medical Education*, *48*(7), 657–666. https://doi.org/10.1111/medu.12432

Core, C., Conci, A., De Angeli, A., Masu, R., & Morreale, F. (2017). Designing a musical playground in the kindergarten. In *Proceedings of the 2017 Conference on Electronic Visualisation and the Arts* (EVA 2017), 13–17 July 2017, London, UK (pp. 1–12). BCS Learning & Development Ltd. https://doi.org/10.14236/ewic/HCI2017.41

de Bayser, M., Azevedo, L. G., & Cerqueira, R. (2015, May). ResearchOps: The case for devops in scientific applications. In R. Badonnel, J. Xiao, S. Ata, F. De Turk, V. Groza, & C. Raniery P. dos Santos (Eds.), *Proceedings of the 2015 IFIP/IEEE International Symposium on Integrated Network Management* (IM 2015), 11–15 May 2015, Ottawa, Ontario, Canada (pp. 1398–1404). https://doi.org/10.1109/INM.2015.7140503

Decuyper, A., Tack, H., Vanblaere, B., Simons, M., & Vanderlinde, R. (2023). Collaboration and shared responsibility in team teaching: A large-scale survey study. *Education Sciences*, *13*(9), 896. https://doi.org/10.3390/educsci13090896

Dollinger, M., Liu, D., Arthars, N., & Lodge, J. (2019). Working together in learning analytics towards the co-creation of value. *Journal of Learning Analytics*, *6*(2), 10–26. https://doi.org/10.18608/jla.2019.62.2

Dollinger, M., & Lodge, J. M. (2018). Co-creation strategies for learning analytics. In *Proceedings of the Eighth International Conference on Learning Analytics and Knowledge* (LAK 2018), 7–9 March 2018, Sydney, Australia (pp. 97–101). ACM. https://doi.org/10.1145/3170358.3170372

Echeverria, V., Yan, L., Zhao, L., Abel, S., Alfredo, R., Dix, S., Jaggard, H., Wotherspoon, R., Osborne, A., Buckingham Shum, S., Gašević, D., & Martinez-Maldonado, R. (2024). TeamSlides: A multimodal teamwork analytics dashboard for teacher-guided reflection in a physical learning space. In *Proceedings of the 14th International Conference on Learning Analytics and Knowledge* (LAK 2024), 18–22 March 2024, Kyoto, Japan. ACM. https://doi.org/10.1145/3636555.3636857

Fernandez Nieto, G. M., Kitto, K., Buckingham Shum, S., & Martinez-Maldonado, R. (2022). Beyond the learning analytics dashboard: Alternative ways to communicate student data insights combining visualisation, narrative and storytelling. In *Proceedings of the 12th International Conference on Learning Analytics and Knowledge* (LAK 2022), 21–25 March 2022, online (pp. 219–229). ACM. https://doi.org/10.1145/3506860.3506895

Fernandez-Nieto, G., An, P., Zhao, J., Buckingham Shum, S., & Martinez-Maldonado, R. (2022). Classroom Dandelions: Visualising Participant Position, Trajectory and Body Orientation Augments Teachers' Sensemaking. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 1–17. https://doi.org/10.1145/3491102.3517736

Fernandez-Nieto, G. M., Martinez-Maldonado, R., Echeverria, V., Kitto, K., Gašević, D., & Buckingham Shum, S. (2024). Data storytelling editor: A teacher-centred tool for customising learning analytics dashboard narratives. In *Proceedings of the 14th International Conference on Learning Analytics and Knowledge* (LAK 2024), 18–22 March 2024, Kyoto, Japan (pp. 678–689). ACM. https://doi.org/10.1145/3636555.3636930

Fernandez-Nieto, G. M., Martinez-Maldonado, R., Kitto, K., & Buckingham Shum, S. (2021). Modelling spatial behaviours in clinical team simulations using epistemic network analysis: Methodology and teacher evaluation. In *Proceedings of the 11th International Conference on Learning Analytics and Knowledge* (LAK 2021), 12–16 April 2021, Irvine, California, USA (pp. 386–396). ACM. https://doi.org/10.1145/3448139.3448176

Folcher, V. (2003). Appropriating artifacts as instruments: When design-for-use meets design-in-use. *Interacting with Computers*, *15*(5), 647–663. https://doi.org/10.1016/S0953-5438(03)00057-2

Friend Wise, A., Brackett, A., & Maddox, B. (2023, March 12). Flexible coupling of learning analytics research and practice in the university: A collective strengths approach. In *Companion Proceedings of the 13th International Conference on Learning Analytics and Knowledge* (LAK 2023), 13–17 March 2023, Arlington, Texas, USA (pp. 25–28). https://www.solaresearch.org/wp-content/uploads/2023/03/LAK23_CompanionProceedings.pdf

Hara, T., Shimada, S., & Arai, T. (2013). Design-of-use and design-in-use by customers in differentiating value creation. *CIRP Annals*, *62*(1), 103–106. https://doi.org/10.1016/j.cirp.2013.03.080

Holstein, K., Harpstead, E., Gulotta, R., & Forlizzi, J. (2020). Replay enactments: Exploring possible futures through historical data. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (DIS 2020), 6–10 July 2020, Eindhoven, Netherlands (pp. 1607–1618). ACM. https://doi.org/10.1145/3357236.3395427

Holstein, K., McLaren, B. M., & Aleven, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher–AI complementarity. *Journal of Learning Analytics*, *6*(2), 27–52. https://doi.org/10.18608/jla.2019.62.3

Kaliisa, R., & Dolonen, J. A. (2023). CADA: A teacher-facing learning analytics dashboard to foster teachers' awareness of students' participation and discourse patterns in online discussions. *Technology, Knowledge and Learning*, *28*(3), 937–958. https://doi.org/10.1007/s10758-022-09598-7

Kaliisa, R., Jivet, I., & Prinsloo, P. (2023). A checklist to guide the planning, designing, implementation, and evaluation of learning analytics dashboards. *International Journal of Educational Technology in Higher Education*, *20*(1), 28. https://doi.org/10.1186/s41239-023-00394-6

Kaliisa, R., Misiejuk, K., López-Pernas, S., Khalil, M., & Saqr, M. (2024, March). Have learning analytics dashboards lived up to the hype? A systematic review of impact on students' achievement, motivation, participation and attitude. In *Proceedings of the 14th International Conference on Learning Analytics and Knowledge* (LAK 2024), 18–22 March 2024, Kyoto, Japan (pp. 295–304). ACM. https://doi.org/10.1145/3636555.3636884

Kim, H., & Lim, Y.-K. (2023, July). Investigating how users design everyday intelligent systems in use. In D. Byrne, N. Martelaro, A. Boucher, D. Chatting, S. F. Alaoui, S. Fox, I. Nicenboim, & C. MacArthur (Eds.), *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (DIS 2023), 10–14 July 2023, Pittsburgh, Pennsylvania, USA (pp. 702–711). ACM. https://doi.org/10.1145/3563657.3596039

Kloos, C. D., Dimitriadis, Y., Hernández-Leo, D., Alario-Hoyos, C., Martínez-Monés, A., Santos, P., Muñoz-Merino, P. J., Asensio-Pérez, J. I., & Safont, L. V. (2022). H2O Learn—Hybrid and human-oriented learning: Trustworthy and human-centered learning analytics (TaHCLA) for hybrid education. In I. Kallel, H. M. Kammoun, & L. Hsairi (Eds.), *Proceedings of the 2022 IEEE Global Engineering Education Conference* (EDUCON 2022), 28–31 March 2022, Tunis, Tunisia (pp. 94–101). IEEE. https://doi.org/10.1109/EDUCON52537.2022.9766770

Knowles, B., Bull, C. N., Davies, N., Simm, W., Bates, O., & Hayes, N. (2019). Examining interdependencies and constraints in co-creation. In *Proceedings of the 2019 Conference on Designing Interactive Systems* (DIS 2019), 23–28 June 2019, San Diego, California, USA (pp. 291–302). ACM. https://doi.org/10.1145/3322276.3322317

Lagrange, J.-B., & Richard, P. R. (2022). Instrumental genesis in the theory of MWS: Insight from didactic research on digital artifacts. In A. Kuzniak, E. Montoya-Delgadillo, & P. R. Richard (Eds.), *Mathematical work in educational context* (pp. 211–228, Vol. 18). Springer International Publishing. https://doi.org/10.1007/978-3-030-90850-8_9

Larrabee Sønderlund, A., Hughes, E., & Smith, J. (2019). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology*, *50*(5), 2594–2618. https://doi.org/10.1111/bjet.12720

Lawrence, L., Echeverria, V., Yang, K., Aleven, V., & Rummel, N. (2023). How teachers conceptualise shared control with an AI co-orchestration tool: A multiyear teacher-centred design process. *British Journal of Educational Technology*, *55*(3), 823–844. https://doi.org/10.1111/bjet.13372

Lee, H. H., & Gargroetzi, E. (2023). "It's like a double-edged sword": Mentor perspectives on ethics and responsibility in a learning analytics–supported virtual mentoring program. *Journal of Learning Analytics*, *10*(1), 85–100. https://doi.org/10.18608/jla.2023.7787

Lee-Cultura, S., Sharma, K., & Giannakos, M. (2023). MultiModal teacher dashboards: Challenges and opportunities of enhancing teacher insights through a case study. *IEEE Transactions on Learning Technologies*, *17*, 181–201. https://doi.org/10.1109/TLT.2023.3276848

Lonchamp, J. (2012). An instrumental perspective on CSCL systems. *International Journal of Computer-Supported Collaborative Learning*, *7*(2), 211–237. https://doi.org/10.1007/s11412-012-9141-4

Martinez-Maldonado, R. (2023). Human-centred learning analytics: Four challenges in realising the potential. *Learning Letters*, *1*, 6. https://doi.org/10.59453/FIZJ7007

Martinez-Maldonado, R., Carvalho, L., & Goodyear, P. (2018). Collaborative design-in-use: An instrumental genesis lens in multi-device environments. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 118:1–118:24. https://doi.org/10.1145/3274387

Martinez-Maldonado, R., Echeverria, V., Fernandez-Nieto, G., Yan, L., Zhao, L., Alfredo, R., Li, X., Dix, S., Jaggard, H., Wotherspoon, R., Osborne, A., Shum, S. B., & Gašević, D. (2023). Lessons learnt from a multimodal learning analytics deployment in-the-wild. *ACM Transactions on Computer-Human Interaction*, *31*(1), 8:1–8:41. https://doi.org/10.1145/3622784

Martinez-Maldonado, R., Pardo, A., Mirriahi, N., Yacef, K., Kay, J., & Clayphan, A. (2016). LATUX: An iterative workflow for designing, validating and deploying learning analytics visualisations. *Journal of Learning Analytics*, *2*(3), 9–39. https://doi.org/10.18608/jla.2015.23.3

McDonald, N., Schoenebeck, S., & Forte, A. (2019). Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 72:1–72:23. https://doi.org/10.1145/3359174

Miller, K., Riley, W., & Davis, S. (2009). Identifying key nursing and team behaviours to achieve high reliability. *Journal of Nursing Management*, *17*(2), 247–255. https://doi.org/https://doi.org/10.1111/j.1365-2834.2009.00978.x

Mohseni, Z. A., Masiello, I., & Martins, R. M. (2023). Codeveloping an easy-to-use learning analytics dashboard for teachers in primary–secondary education: A human-centered design approach. *Education Sciences*, *13*(12), 1190. https://doi.org/10.3390/educsci13121190

Moltudal, S. H., Krumsvik, R. J., & Høydal, K. L. (2022). Adaptive learning technology in primary education: Implications for professional teacher knowledge and classroom management. *Frontiers in Education*, *7*. https://doi.org/10.3389/feduc.2022.830536

Muller, M. J. (2002). Participatory design: the third space in HCI. In J. A. Jacko & A. Sears (Eds.), *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (pp. 1051–1068). L. Erlbaum Associates Inc. https://dl.acm.org/doi/10.5555/772072.772138

Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology*, *53*(4), 914–931. https://doi.org/10.1111/bjet.13232

Nelson, J., Buisine, S., & Aoussat, A. (2009). Design in use: Some methodological considerations. In *CIRP MS'09, 42nd CIRP Conference on Manufacturing Systems* (pp. 3–5). https://api.semanticscholar.org/CorpusID:167215398

Nicholson, R., Bartindale, T., Kharrufa, A., Kirk, D., & Walker-Gleaves, C. (2022). Participatory design goes to school: Co-teaching as a form of co-design for educational technology. In S. Barbosa, C. Lampe, C. Appert, D. A. Shamma, S. Drucker, J. Williamson, & K. Yatani (Eds.), *CHI Conference on Human Factors in Computing Systems* (CHI 2022), 29 April–5 May 2022, New Orleans, Louisiana, USA (pp. 1–17). ACM. https://doi.org/10.1145/3491102.3517667

Ouhaichi, H., Spikol, D., & Vogel, B. (2023). Research trends in multimodal learning analytics: A systematic mapping study. *Computers and Education: Artificial Intelligence*, *4*, 100136. https://doi.org/10.1016/j.caeai.2023.100136

Overdijk, M., Van Diggelen, W., Tomadaki, E., & Scott, P. (2006). Technology appropriation in face-to-face collaborative learning. In *Innovative Approaches for Learning and Knowledge Sharing, EC-TEL 2006 Workshops Proceedings,* 1 October 2006, Crete, Greece (pp. 89–96). https://www.researchgate.net/publication/221549900_Technology_Appropriation_in_Face-to-Face_Collaborative_Learning

Paniagua, A., & Istance, D. (2018). *Teachers as designers of learning environments: The importance of innovative pedagogies*. OECD. https://doi.org/10.1787/9789264085374-en

Rabardel, P., & Bourmaud, G. (2003). From computer to instrument system: A developmental perspective. *Interacting with Computers*, *15*(5), 665–691. https://doi.org/10.1016/S0953-5438(03)00058-4

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust speech recognition via large-scale weak supervision. *arXiv preprint arXiv:2212.04356*.

Riley, W., Hansen, H., Gürses, A. P., Davis, S., Miller, K., & Priester, R. (2008). The nature, characteristics and patterns of perinatal critical events teams. In K. Henriksen, J. B. Battles, M. A. Keyes, & M. L. Grady (Eds.), *Advances in patient safety: New directions and alternative approaches (Vol. 3: Performance and tools)*. Agency for Healthcare Research; Quality (US). https://www.ncbi.nlm.nih.gov/books/NBK43664

Schmitz, M., Scheffel, M., Bemelmans, R., & Drachsler, H. (2022). FoLA2—A method for co-creating learning analytics–supported learning design. *Journal of Learning Analytics*, *9*(2), 265–281. https://doi.org/10.18608/jla.2022.7643

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. https://doi.org/10.18608/jla.2016.33.3

Slade, S., Prinsloo, P., & Khalil, M. (2023). "Trust us," they said. Mapping the contours of trustworthiness in learning analytics. *Information and Learning Sciences*, *124*(9/10), 306–325. https://doi.org/10.1108/ILS-04-2023-0042

Susnjak, T., Ramaswami, G. S., & Mathrani, A. (2022). Learning analytics dashboard: A tool for providing actionable insights to learners. *International Journal of Educational Technology in Higher Education*, *19*(1), 1–23. https://doi.org/10.1186/s41239-021-00313-7

Tchounikine, P. (2017). Designing for appropriation: A theoretical account. *Human–Computer Interaction*, *32*(4), 155–195. https://doi.org/10.1080/07370024.2016.1203263

Tessier, S. (2012). From field notes, to transcripts, to tape recordings: Evolution or combination? *International Journal of Qualitative Methods*, *11*(4), 446–460. https://doi.org/10.1177/160940691201100410

Topali, P., Ortega-Arranz, A., Rodríguez-Triana, M. J., Er, E., Khalil, M., & Açkapınar, G. (2024). Designing human-centered learning analytics and artificial intelligence in education solutions: A systematic literature review. *Behaviour & Information Technology*. https://doi.org/10.1080/0144929X.2024.2345295

Vágová, R. (2021). Designing combinations of physical and digital manipulatives to develop students' visualisation. *Open Education Studies*, *3*(1), 56–75. https://doi.org/10.1515/edu-2020-0140

Van Schoors, R., Elen, J., Raes, A., & Depaepe, F. (2023). Tinkering the teacher–technology nexus: The case of teacher- and technology-driven personalisation. *Education Sciences*, *13*(4), 349. https://doi.org/10.3390/educsci13040349

Wiley, K., Dimitriadis, Y., & Linn, M. (2024). A human-centred learning analytics approach for developing contextually scalable K-12 teacher dashboards. *British Journal of Educational Technology*, *55*(3), 845–885. https://doi.org/10.1111/bjet.13383

Williamson, K., & Kizilcec, R. (2022). A review of learning analytics dashboard research in higher education: Implications for justice, equity, diversity, and inclusion. In *Proceedings of the 12th International Conference on Learning Analytics and Knowledge* (LAK 2022), 21–25 March 2022, online (pp. 260–270). ACM. https://doi.org/10.1145/3506860.3506900

Yan, L., Martinez-Maldonado, R., & Gasevic, D. (2024). Generative artificial intelligence in learning analytics: Contextualising opportunities and challenges through the learning analytics cycle. In *Proceedings of the 14th International Conference on Learning Analytics and Knowledge* (LAK 2024), 18–22 March 2024, Kyoto, Japan (pp. 101–111). ACM. https://doi.org/10.1145/3636555.3636856

Yan, L., Zhao, L., Echeverria, V., Jin, Y., Alfredo, R., Li, X., Gaševi'c, D., & Martinez-Maldonado, R. (2024). VizChat: Enhancing learning analytics dashboards with contextualised explanations using multimodal generative AI chatbots. In A. M. Olney, I.-A. Chounta, Z. Liu, O. C. Santos, & I. I. Bittencourt (Eds.), *Artificial intelligence in education* (pp. 180–193). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-64299-9_13

Yang, K. B., Echeverria, V., Lu, Z., Mao, H., Holstein, K., Rummel, N., & Aleven, V. (2023). Pair-Up: Prototyping human-AI co-orchestration of dynamic transitions between individual and collaborative learning in the classroom. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (CHI 2023), 23–28 April 2023, Hamburg, Germany (pp. 1–17). ACM. https://doi.org/10.1145/3544548.3581398

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). Re-examining whether, why, and how human-AI interaction is uniquely difficult to design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (CHI 2020), 25–30 April 2020, Honolulu, Hawaii, USA (pp. 1–13). ACM. https://doi.org/10.1145/3313831.3376301

Zhao, L., Echeverria, V., Swiecki, Z., Yan, L., Alfredo, R., Li, X., Gasevic, D., & Martinez-Maldonado, R. (2024). Epistemic network analysis for end-users: Closing the loop in the context of multimodal analytics for collaborative team learning. In *Proceedings of the 14th International Conference on Learning Analytics and Knowledge* (LAK 2024), 18–22 March 2024, Kyoto, Japan (pp. 90–100). ACM. https://doi.org/10.1145/3636555.3636855

Zhao, L., Swiecki, Z., Gasevic, D., Yan, L., Dix, S., Jaggard, H., Wotherspoon, R., Osborne, A., Li, X., Alfredo, R., & Martinez-Maldonado, R. (2023). METS: Multimodal learning analytics of embodied teamwork learning. In *Proceedings of the 13th International Conference on Learning Analytics and Knowledge* (LAK 2023), 13–17 March 2023, Arlington, Texas, USA (pp. 186–196). ACM. https://doi.org/10.1145/3576050.3576076

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 *Proceedings of the 12th International Conference on Learning Analytics and Knowledge* (LAK 2022), 21–25 March 2022, online (pp. 370–380). ACM. https://doi.org/10.1145/3506860.3506935

1. Appendix A: Supplementary Material

A.1 The Development of *DebriefPad Dashboard* with Teachers

Between February 2023 and the deployment of the simulation classes (in August 2023), a co-creation team was consolidated with the purpose of creating the LA dashboard called *DebriefPad Dashboard*. This comprised two human-centred design researchers, two LA data visualization researchers, and two software developers. Complementing this, five nursing teachers and two senior lecturers who brought a deep understanding of nursing education were also part of the team. Our co-creation process involved several (bi-weekly) focus-group meetings and design workshops. Drawing inspiration from the LATUX framework (Martinez-Maldonado et al., 2016), we summarize our design and development stages (stages 1 to 3) in Figure 6, while stage 4 is detailed in the main manuscript.

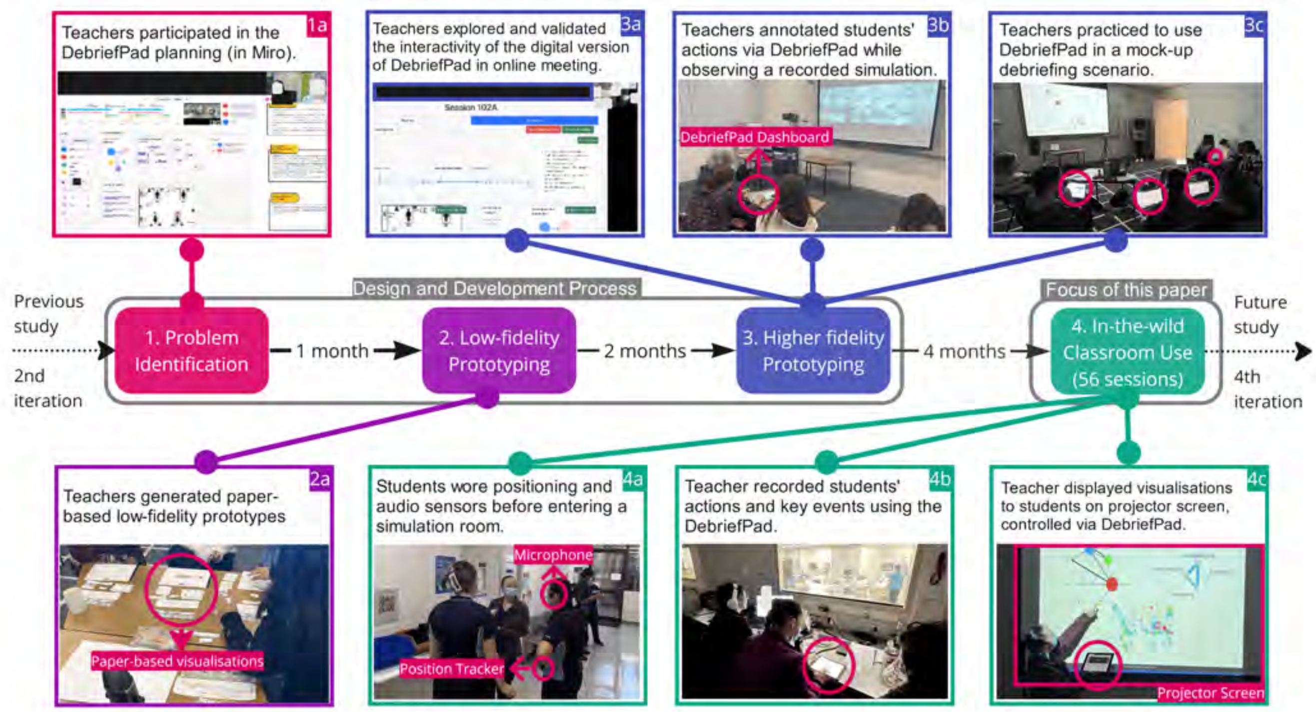

Figure 6. Design process and in-use of *DebriefPad Dashboard* with nursing teachers, adapted from LATUX framework (Martinez-Maldonado et al., 2016).

Stage 1: Problem identification. This stage aimed to identify problems from the existing LA system in our previous study (Echeverria et al., 2024) and establish improvements in the upcoming LA dashboard deployment. To facilitate this process, we organized three focus-group meetings ($n = 3$; Figure 6.1a). These three meetings were attended by our internal research team and a group of four educators of this study, including the two senior lecturers and two teachers. Their insights and expertise were invaluable in identifying the areas of the existing LA system that needed enhancement. During these discussions, the idea of annotating students' actions while observing a simulation and controlling visualizations on a handheld tablet computer emerged. This tablet was intended to improve portability from an existing simpler LA dashboard that was accessed on a desktop computer. In addition, teachers requested functionalities to filter data about students' simulation based on the annotated actions. This phase's outcome allowed the research team to explore the technological feasibility of the proposed improvements.

Stage 2: Low-fidelity prototyping. Once the research team agreed on the functionalities of the LA dashboard, a paper-based design workshop was conducted with three nursing teachers, one senior lecturer, and two teachers (Figure 6.2a). Inspired by the work of Mohseni and colleagues (2023), the primary objective was to explore and identify optimal ways to organize various visualizations on a handheld tablet computer. The teachers were provided pictures of initial versions of data visualizations, co-created with other teachers a year before (Echeverria et al., 2024), as well as other tablet material design elements (buttons, labels, tablet screen, etc.). They were also tasked with designing a layout to best accommodate these visualizations on a tablet interface. The teachers engaged in a collaborative discussion, sharing their insights and perspectives on the proposed designs. They evaluated each design's usability and relevance to their knowledge of debriefing use. After a comprehensive review, they collectively voted to select the most effective wireframe design. This chosen design was the blueprint for developing a higher-fidelity prototype in the next stage. The outcome of this stage was not just a wireframe but a shared vision for an LA system.

Stage 3: Higher-fidelity prototyping. During this stage, teachers compiled a list of actions for a clinical simulation and evaluated the system's data visualizations and interactions through two user-testing activities. This was a four-month-long stage and included four online meetings, two in-person focus-group meetings, and one training workshop ($n = 7$ sessions; see Figure 6.3a–c). Two senior lecturers introduced three more teachers with roles similar to the existing ones. These teachers participated in at least three co-creation sessions including training, two months before the classroom deployment. Moreover, the developers were actively engaged in developing and updating the system iteratively. Their work was primarily based on the features that were co-created and agreed upon during the meetings. An integral part of this stage was testing these high-fidelity prototypes and the initial versions of the *DebriefPad Dashboard*. The main objective of this testing phase was to validate the system's interactivity and identify any design issues that might hinder its functionality. The testing consisted of a replay enactment (Holstein et al., 2020), allowing the teachers to use the system while they observed a recorded simulation of students' activities. These recordings were sourced from previous iterations of the study (Echeverria et al., 2024; Zhao et al., 2023; Martinez-Maldonado et al., 2023). This approach provided the teachers with practical and hands-on experience, enabling them to understand and evaluate the tool's effectiveness in a close to real world setting.

Stage 4: In-the-wild classroom use. In this final stage, teachers used *DebriefPad Dashboard* in the wild ($n = 56$ sessions). Students wore sensors that streamed positions and audio data to the system during the simulation (Figure 6.4a). The system processed the information and produced four visualizations, detailed in the next section. While the teacher observed students completing the simulation scenario, they tagged students' actions and three critical moments in the scenario, namely "End of Handover," "Secondary Nurse Enters Sim Room," and "Doctor Enters Sim Room" (Figure 6.4b). After the simulation scenario ended, these students convened in a debriefing room with two teachers to debrief and discuss students' clinical, teamwork, and communication performance, with the aid of *DebriefPad Dashboard* via its projector view feature (Figure 6.4c).

A.2 *DebriefPad Dashboard* System Overview

The result of the co-creation process was the *DebriefPad Dashboard* system. This system's primary goal was to allow teachers to provide task prioritization, teamwork, and communication evidence for supporting students' reflection on their clinical simulation in the debriefing. The *DebriefPad Dashboard* was an interactive web application developed with the JavaScript framework (*ReactJS* and *NodeJS*) and deployed on a local web server as per teachers' design requirements. This web application could be accessed on a handheld tablet computer (i.e., iPad; see Figure 7), with a goal of improving its portability. This application consisted of two main pages, the "actions-tagging page" and the "projector-controller page." The actions-tagging page was used to annotate students' actions while teachers were observing the simulation, and the projector-controller page was used during the teacher-guided debriefing activity after a simulation using a series of data visualizations.

A.2.1 Data Visualizations and Processing

After the simulation, four visualizations were generated using the collected multimodal data from individual students' positions and audio recordings. Three visualizations (*Teamwork Barchart*, *Ward Map*, and *Communication Network*) had been co-created by and evaluated with other teachers in a previous year (Echeverria et al., 2024). The current paper thus emphasizes the description of the communication behaviour graph because teachers co-created it for this deployment and heavily refined it through design-in-use. Given the focus of this study on interface design-in-use, a comprehensive account of technical details of the communication behaviour graph (i.e., data training, labelling, etc.) has been intentionally simplified.

The collected spatial and verbal data was processed and sent to the *DebriefPad Dashboard* in a sequence based on processing speed. The spatial coordinates and body orientation were the fastest to process, as they were interpolated to obtain one position of each student in the learning space per second. The spatial coordinates were processed to generate the *Teamwork Barchart*. Second, a voice activity detector (py-WebtcVAD[2]) was used on the audio recordings to extract the time frames when students talked. This time frame data was combined with spatial coordinates to generate the *Ward Map*. Third, the interpolated body orientation, spatial coordinates, and time frames were used to determine who was talking to whom, employing an f-formation-based algorithm (Zhao et al., 2022), and then generating the *Communication Network*. Finally, we used the OpenAI Whisper medium model (Radford et al., 2022) to automatically transcribe the audio recordings. These transcriptions were coded using a fine-tuned BERT-based classification model to classify the four critical behaviours (Zhao et al., 2024). This coded data was used to generate the communication behaviour graph, an AI-powered LA visualization. This approach ensured that each visualization was dispatched as soon as it was ready, optimizing the overall data processing workflow. The visual layout of each visualization is as follows.

[2]https://github.com/wiseman/py-webrtcvad

1. **Teamwork Barchart.** The barchart aimed to visualize five key student behaviours during a simulation: working together/individually on tasks for *Ruth* (the patient in an emergency), working together/individually on other patients (the stable patients), and transiting among the beds (Echeverria et al., 2024). The behaviours were tracked using students' spatial coordinates. The duration of these behaviours was accumulated and categorized accordingly. The outcome was a barchart representing the proportion of time students spent on each of the five behaviours during the simulation.

2. **Ward Map.** The simulation room map, inspired by G. Fernandez-Nieto and colleagues (2022), aimed to illustrate students' spatial and verbal behaviours during a simulation. This map was created by integrating and customizing NBAShotTracker[3]. Four colours were used to represent each student, with red and blue for primary nurses (PN) 1 and 2, and green and yellow for secondary nurses (SN) 1 and 2. Coloured hexagons on the map indicated where a student stood, with deeper border colours indicating longer standing times. The depth of a hexagon's fill represented the duration of a student's speech at that location. The outcome was a diagrammatic map that provided a visual representation of both the spatial and verbal behaviours of students during the simulation.

3. **Communication Network.** The network aimed to illustrate the communication volume during a simulation, with each node representing a participant. Nodes for students would increase in size based on the presence of verbal communication captured by the microphones, while nodes for actors (teachers role-playing as doctor, relative, and primary patient) remained static. The thickness and direction of the edges represented the amount and recipient of the communication, respectively. The outcome is a network that visually represents the communication dynamics among participants during the simulation.

4. **Communication Behaviour Graph.** Inspired by epistemic networks that are commonly used in LA (Shaffer et al., 2016), the aim was to illustrate the co-occurrences of critical communication behaviours in a team, such as closed-loop communication (Zhao et al., 2024), derived from existing healthcare teamwork studies (Zhao et al., 2024; Miller et al., 2009; Riley et al., 2008). These behaviours included (1) *task allocation*—the verbal delegation of tasks (Riley et al., 2008); (2) *handover*—the structured communication of medical situations to the new team members (Zhao et al., 2024); (3) *escalation*—the announcement of an emergency on patient's health conditions (Zhao et al., 2024); (4) *sharing information*—the proactive provision of information to others (Miller et al., 2009); (5) *questioning*—the verbal asking of questions (Miller et al., 2009); and (6) *acknowledging*—explicit acknowledging and responding with requested information (Zhao et al., 2024). The network graph was developed with nodes representing these behaviours. Edges represented co-occurrences, with thickness indicating frequency. Node size represented the co-occurrence of a behaviour with others. The outcome is a network graph visually representing critical communication behaviours' co-occurrences.

A.2.2 The Actions-Tagging Page

The teacher with the role of primary debriefer was meant to use this page (Figure 7, top image) to annotate key actions in the simulation, such as "the handover of patient's information," "the entry of secondary nurses," and "the arrival of the doctor," while observing the students' activity during the simulation from the control room (see Figure 6.4b). Teachers could annotate the team's actions using the preset list of actions tailored to the simulation scenario and write custom personal notes for their reference in the debriefing session. This action list was created with teachers in the co-created workshops and meetings as described above. The timestamps from these annotations were intended to be used to filter the data of the visualizations using the projector-controller page during the actual debriefing.

A.2.3 Projector-Controller Page

Inspired by Baldauf and colleagues (2016), this page (Figure 7, bottom image) allowed teachers to share generated visualizations from their tablet to the classroom projector screen in a debrief room after students' simulation. By default, these visualizations represented a whole simulation scenario from the beginning to the end. The teachers could filter these visualizations by interacting with the draggable timeline or selecting one of four preset key event buttons, according to the timing they tagged in the *actions-tagging page*. This allowed teachers to narrow down the data displayed, making the visualizations more manageable and relevant to what the teachers wanted to focus on (i.e., a specific key event, a particular time frame, or a certain action) before sending these visualizations to the public classroom display. Teachers co-created an overlay window screen called the "projector preview page" to allow the primary debriefer to review a set of selected visualizations on their tablet before showing them to the students (see Figure 6.4c). The projector screen could automatically reorganize its layout to fit selected visualizations and update visualizations when the teacher filtered them on a visualization control panel.

[3] https://github.com/ebaek/NBAShotTracker

Figure 7. Two pages of *Debrief Pad Dashboard*. The actions-tagging page (top) was used by the teacher when observing nursing students during a simulation to record simulations' key events and students' clinical and teamwork actions. The projector-controller page (bottom) was used by the teacher during a debriefing to display LA visualizations on a projector screen.

A.3 Post Hoc Interview Questions

One researcher conducted all five interviews virtually using an online video conferencing tool (Zoom). The interview began by asking teachers about their general experience with the *DebriefPad Dashboard*, including the benefits and challenges of using the system. To address our research questions, semi-structured interview questions were developed focusing on the design-in-use listed in Table 2, as previously analyzed. These deviations from the *design-for-use* prompted us to explore the underlying reasons for shaping the system and uncovering any changes in their teaching practices.

Table 2. A list of co-created features, each with its intended use (*design-for-use*) within the *DebriefPad Dashboard*, that had been shaped into unexpected outcomes (*design-in-use*).

No.	Goal	Design-for-use	Design-in-use
1	To identify the team's critical communication behaviours	Using six nodes of *Communication Behaviour Graph* (see Section A.2.1)	The visualization was requested to be simplified from six to four nodes, with the "Handover" and "Sharing Information" labels being removed.
2	To preview visualizations before sharing them on a projector screen	The "projector preview page" was designed, allowing teachers to review the combination of visualizations on a handheld device before sharing them to a projector screen.	The "projector preview page" was requested to be removed in the final version to improve its ease of use.
3	To provide evidence about prioritization, teamwork, and communication.	Supported by LA data visualizations, teachers could provide evidence of task prioritization, teamwork, and communication.	The visualizations were used to regulate students' emotion and motivate them.
4	To guide students with complete visualizations for reflection.	The system was used only after all data was fully processed, ensuring a comprehensive view. This strategy anticipated using the system primarily toward the end of the debriefing.	A practice was developed to share with students in the middle or at the start of debriefing, as soon as data was available based on the order of data visualization being processed (see Section A.2.1).
5	To be used only by the primary debriefer during simulation and debriefing.	The primary debriefer held the device the whole time by tagging actions while observing a simulation and controlling the visualizations on a projector during debriefing.	Initially, the primary debriefer tagged students' actions while observing a simulation and then passed the device to the secondary debriefer for debriefing. Later, the device was reassigned for individual use by the primary debriefer for the rest of the study.

To address **RQ1**, we began with *permanent instrumentalization*, which pertains to permanent modifications to the feature in a system. This included the simplification of the communication behaviour graph and the removal of the projector preview page (refer to Table 2, rows 1 and 2).

1. We asked teachers about the original design decisions: "During the design, we agreed to have [design-for-use feature]; do you remember why we took this design decision?"

2. We probed into decisions that led to changes in the original intention during use: "After you used the system, you gave us feedback about [the communication behaviour graph and the project preview page], and we changed it in the first week. Can you remember why we decided to change it?"

3. We asked teachers to reflect on the final features as they were implemented in practice: "What were your thoughts on the [updated feature] when you used it in your debriefing practice?"

In response to *temporary instrumentalization*, which can be exemplified by an occasional shift in the system's use to motivate and regulate students' emotions without modifying the system (Table 2, row 3), we asked teachers:

1. "The primary objective of the system is to make the team's performance visible, particularly in terms of task prioritization, teamwork, and communication strategies. Do you think you have used it in a way that deviates from this main goal? For example, to motivate students or regulate their emotions?"

To address **RQ2**, we began with *individual instrumentation*, which pertains to changes in individual practice affected by the system's use. An example was when teachers developed expertise in sharing visualizations with students whenever data was available, which may show if the system could change their debriefing practice (refer to Table 2, row 4). We began our investigation by asking participants about their personal experiences with the *DebriefPad Dashboard* and any potential challenges they encountered. We used screenshots of how they individually projected the visualizations on a screen to formulate three questions, as follows:

1. "We observed that you displayed these visualizations [independently or in combination] on the projector screen. Could you explain why you chose to use it this way?"
2. "We noticed that you used it [*at the start, in the middle, or at the end*]. Could you elaborate on your reasoning for this timing?"
3. "Compared to your usual debriefing practice, such as making observations using pen and paper and fully verbal feedback, to what extent do you think the system has changed your debriefing practice?"

Regarding *collective instrumentation*, this could be exemplified by changes in practice where the system was used collectively, even though it was originally designed for individual use (see Table 2, row 5). In this scenario, primary debriefers delegated control to secondary debriefers in the first week ($n = 8$ sessions). The primary debriefer would concentrate on the conversation with students, while the secondary debriefer prepared visual cues to facilitate the discussions. However, they decided to revert to individual use for the remaining sessions, restoring the original role of the primary debriefer in controlling the system. We posed three questions:

1. "During the first week, we observed that you decided to have the secondary debriefer operate the tool while the primary debriefer focused on the conversation during the debriefing. Could you explain why you adopted this approach?"
2. "What was the strategy you decided to implement when conducting the debrief in pairs to ensure a smooth debriefing process?"
3. "After some time, you decided to use the tool individually again. Could you explain why you decided to revert to individual use?"