Automating Multimodal Data Storytelling for Embodied Team Learning

by *Gloria Milena Fernández Nieto*

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Learning Analytics

under the supervision of

Dr. Roberto Martínez Maldonado
Prof. Simon Buckingham Shum
A./Prof. Kirsty Kitto

University of Technology Sydney
Connected Intelligence Centre
July 2022, Australia

Automating Multimodal Data Storytelling for Embodied Team Learning

by *Gloria Milena Fernández Nieto*

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Learning Analytics

under the supervision of

Dr. Roberto Martínez Maldonado
Prof. Simon Buckingham Shum
A./Prof. Kirsty Kitto

University of Technology Sydney
Connected Intelligence Centre
July 2022, Australia

Abstract

There is a growing interest in creating Learning Analytics (LA) interfaces that support students and teachers directly. Thus far, many of these solutions have been materialised as dashboards and visualisations. However, although a growing number of prototypes and commercial products aimed at supporting students/teachers exist, their limitations are coming under careful scrutiny. For instance, many visual LA tools are failing to provide meaningful and relevant insights that can support students reflections on their embodied teamwork activity.

Moreover, there are additional challenges in visualising and communicating the wide variety of *multimodal sensor data* captured from physical spaces, in a way that supports educational stakeholders (e.g. teachers or students), who as casual users, have limited training in data analysis and interpretation. Thus, this thesis engages research in Information Visualisation (InfoVis) and specifically the **Guidance** visualisation paradigm that aims to support casual users, or those users with low analysis expertise, to narrow the gap of data visualisation interpretations. **Data Storytelling** is one way to provide guidance, as a compression technique to help an audience effectively understand what is important in an visualisation, communicating key messages combining *data*, *visualisations*, and *narratives*. 'Telling stories' with data in these ways should enable the elicitation of deeper reflections in an effective manner.

This thesis tackles the above challenges by investigating: *"How can salient aspects of embodied team activity be automatically identified and derived insights be communicated to support timely, productive reflection?"* Four research questions were derived: (1) What **modelling techniques** can enable identification of salient aspects of multimodal embodied team activity according to the learning design (i.e. teachers' pedagogical intentions)? (2) How can insights be extracted from multimodal sensors and **communicated to students and teachers** to support teaching and reflection on embodied team activity? (3) To what extent can students and teachers **reflect** on embodied team activity using MMLA interfaces? and (4) To what extent can MMLA interfaces for students and teachers be **automatically generated**?

This research adopted a **mixed methods** approach using quantitative and qualitative analyses to provide evidence in response to these research questions. Empirical studies of teamwork were conducted in **authentic higher education settings** in the context of *healthcare* (simulations) and *science* (physics lab) education. Automated multimodal Data Stories were co-designed, evaluated, and implemented to support students and teachers to reflect on different aspects of their embodied practice (patient-care, or co-teaching).

This research makes three types of contribution: modelling, prototypes (MMLA interfaces), and implementation. Regarding **modelling** , two contributions are presented: (i) a methodology to map from teachers' pedagogical intentions to salient aspects of multimodal embodied team activity, and (ii) the exploration of a set of five multimodal data modelling techniques. In terms of **MMLA interfaces** two contributions are presented: (iii) a set of seven multimodal data interfaces were created, and (iv) the analysis and evaluation of them to understand their potential for reflection. Finally, in terms of **implementations**, this research contributes a (v) functional architecture description and documentation and (vi) a *reference implementation* to automatically generate multimodal LA data stories.

Results from this research point to the potential of creating alternative ways to communicate multimodal data insights to teachers and students, by combining visualisation, narrative and storytelling, driven by teachers' pedagogical intentions. Multimodal data stories of embodied team activity are effective tools to support: (a) teachers as they reflect on student progress; (b) students as they reflect, recall and improve their future practices; (c) teachers in providing timely feedback according to their pedagogical intentions; and (d) teachers as they reflect on their co-teaching practice. In addition, this thesis identifies that LA designers should identify representations that best fit teachers' and students' needs, by contextualising and aligning pedagogical intentions with the visual analytics.

Keywords: guidance, data storytelling, embodied team learning, learning by reflection, multimodal data, pedagogical intentions, learning design.

Author's declaration

I, *Gloria Milena Fernández-Nieto* declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy in Learning Analytics, in the *Connected Intelligence Centre* at the University of Technology Sydney, Australia, is wholly my own work, unless otherwise referenced or acknowledged.

In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution. In addition, I received direct authorisation from the main authors to include their images to document previous research in Chapter 2. This research is supported by the Australian Government Research Training Program under the International Research Training Program Scholarship (IRTP).

Production Note:

SIGNATURE: Signature removed prior to publication.

[Gloria Milena Fernández Nieto]

DATE: 21 July, 2022

PLACE: Sydney, Australia

Dedication

To my beloved husband and my Colombian and Aussie families which have always been there giving me support and motivation. My heart is with you all and will always be with you ...

Acknowledgments

I gratefully acknowledge the University of Technology Sydney for support of my Ph.D. scholarship. Generally, scholarship motivates many international students who, like me, are achieving goals. My sincere thanks for becoming dreams in realities.

Sincere admiration and thanks to my supervisory team that has always been there to support my journey. Professor Simon Buckingham Shum, who brought me the option to be part of the Connected Intelligence Centre (CIC) as a Ph.D. researcher, and his fabulous team. Professor Kirsty Kitto for her great contribution with ideas and improvements throughout the journey. Particularly, Dr. Roberto Martinez-Maldonado for his great support and for encouraging me to work hard while enjoying the path. I was very lucky to have a great supervisory team; the sessions were always pleasant and full of challenges, motivation, and enthusiasm to support others with our research.

One of my lessons learnt in my Ph.D. was about collaboration and teamwork. I have worked with great academics around the world, and I want to acknowledge your great work and commitment to my research. Special thanks to the Nursing team at UTS and Monash University. Teachers and students were always willing to participate and contribute to the research. It is grateful to see how participants (teachers and students) envisage this thesis as a tool that support their daily practices. A big thanks to Dr. Carmen Axisa and Professor Doug Elliott, your hard work and commitment was of great help for this research.

Deep thanks to the CIC team, I found in you great colleagues and friends. When I arrived at CIC most of you were in your last Ph.D. year, yet you welcomed me and made me a part of the team immediately. You all have made my time in Australia as if I were home. Special thanks to Vanessa, Carlos, Evelyn, and Sophie for the great, funny, and supportive conversations about never giving up. Gabrielle, thanks for your patience and great attitude to support and solve all my questions, you saved my life many times.

It has been said that "it is the friends you can call up at 4 a.m. that matter"; considering that Australia is far away from Colombia, in that case, all my friends matter. Special thanks to my Colombian and Mexican friends Angela, Adri, Eliana, Angie, Andrea, and Itzel.

[Spanish alert!] Principalmente, gracias a mi familia en Colombia. Los amo y aprecio todo el apoyo y acompañamiento durante estos años lejos de casa. Durante este tiempo aprendí grandes lecciones no solo como profesional sino como ser humano. En primera instancia, aprendí a no dar las cosas por hecho, la distancia me demostró que no estar cerca de las personas que amas puede ser muy doloroso, aprovecharlas al máximo es ahora parte de mi motivación de vida. Además, aprendí que la familia esta siempre para ti, en los buenas y en los malos momentos. A mi madre bella (Gloria Stella), por supuesto, sin su apoyo, su motivación y sus sacrificios, yo no estaría disfrutando de este nuevo logro, ¡gracias por todo madre! A mi padre (Ramón Fernández), porque su berraquera, su empuje y su constante trabajo me permitieron estudiar y cumplir sueños. A mi hermanito (Camilo Fernánadez), porque, el ser humano que soy hoy, es en parte por su influencia en mi, siempre has sido mi ejemplo.

Durante estos años mi familia creció, ahora no solo soy Fernández Nieto también soy Garcés Takegami y Carmona Suárez. Mi esposo (Oscar K. Garcés) fué, es, y será siempre el mejor apoyo y motivador de mi vida, ¡te amo! Por tantas cosas maravillosas y por su amor estoy muy agradecida con la vida, los amo familia y este nuevo logro se los dedico a ustedes.

Finally, I could not be luckier because now I also have an Aussie family. Dianne and Glenn Holdstock welcomed me a couple of years ago as their guest, and since then they have treated me as part of their family. My Colombian family and I are sincerely grateful for the home you offer me here in Australia. I also acknowledge Dr. Chandranath Adak for providing this thesis template.

List of Publications

Journal Papers:

1. **Gloria Milena Fernandez-Nieto** et al., "Storytelling With Learner Data: Guiding Student Reflection on Multimodal Team Data", in IEEE Transactions on Learning Technologies (2021), doi: 10.1109/TLT.2021.3131842

2. **Gloria Milena Fernandez-Nieto**, Roberto Martinez-Maldonado, Vanessa Echeverria, Kirsty Kitto, Pengcheng An,and Simon Buckingham Shum. 2021. What Can Analytics for Teamwork Proxemics Reveal About Positioning Dynamics In Clinical Simulations?. Proc. ACM Hum. Comput. Interact.5, CSCW1, Article 185 (April 2021),24 pages. doi: 10.1145/3449284

3. Martinez-Maldonado, R., Gasevic, D., Echeverria, V., **Fernandez-Nieto, G.**, Swiecki, Z., and Buckingham Shum, S. (2021). What do you mean by collaboration analytics? a conceptual model. Journal of Learning Analytics, 8(1), 126–153. doi: 10.18608/jla.2021.7227

4. Martinez-Maldonado, R., Echeverria, V., Mangaroska, K., Shibani, A., **Fernandez-Nieto, G.**, Schulte, J., and Buckingham Shum, S. (2022). Moodoo the tracker: Spatial classroom analytics for characterising teachers' pedagogical approaches. International Journal of Artificial Intelligence in Education, 8(1), 1–27. 10.1007/s40593-021-00276-w

Conference proceedings:

1. **Gloria Milena Fernandez-Nieto**, Pengcheng An, Jian Zhao, Simon Buckingham Shum, and Roberto Martinez-Maldonado. 2022. Classroom Dandelions: Visualising Participant Position, Trajectory and Body Orientation Augments Teachers' Sensemaking. In CHI Conference on Human Factors in Computing Systems (CHI'22), April 29-May 5, 2022, New Orleans, LA, USA. ACM, New York, NY, USA, 17 pages, doi: 10.1145/3491102.3517736. 2022.

2. **Gloria Milena Fernandez-Nieto**,Simon Buckingham Shum, Kirsty Kitto, and Roberto Martinez-Maldonado. 2022. Beyond the Learning Analytics Dashboard: Alternative Ways to Communicate Student Data Insights Combining Visualisation, Narrative and Storytelling. In LAK22: 12 th International Learning Analytics and Knowledge Conference (LAK22), March 21 –25, 2022, Online, USA.ACM, New York,NY, USA, 16 pages. doi: 10.1145/3506860.3506895

3. **Gloria Milena Fernandez-Nieto**, Roberto Martinez-Maldonado, Kirsty Kitto,and Simon Buckingham Shum. 2021. Modelling Spatial Behaviours in Clinical Team Simulations using Epistemic Network Analysis: Methodology and Teacher Evaluation. In LAK21: 11th International Learning Analytics and Knowledge Conference (LAK21), April 12–16, 2021, Irvine, CA, USA.ACM,New York, NY, USA, 11 pages. doi: 10.1145/3448139.3448176

4. Lixiang Yan, Roberto Martinez-Maldonado, Beatriz Gallo Cordoba, Joanne Deppeler, Deborah Corrigan, **Gloria Fernandez Nieto**, and Dragan Gasevic. 2021. Footprints at School: Modelling In-class Social Dynamics from Students' Physical Positioning Traces. In LAK21: 11th International Learning Analytics and Knowledge Conference (LAK21). Association for Computing Machinery, New York, NY, USA, 43–54. doi: 10.1145/3448139.3448144

5. Martinez-Maldonado, R., Gašević, D., Echeverria, V., **Fernandez-Nieto, G.**, Swiecki, Z., and Buckingham Shum, S. (2021). What Do You Mean by Collaboration Analytics? A Conceptual Model. Journal of Learning Analytics, 8(1), 126-153. doi: 10.18608/jla.2021.7227

6. Roberto Martinez-Maldonado, PhD; Vanessa Echeverria; Katerina Mangaroska; Antonette Shibani; **Gloria Fernandez-Nieto**; Jurgen Schulte; Simon Buckingham Shum. International Journal of Artificial Intelligence in Education (IJAIED'21). Moodoo the Tracker: Spatial Classroom Analytics for Characterising Teachers' Pedagogical Approaches. doi: 10.1007/s40593-021-00276-w

7. Roberto Martinez-Maldonado, Vanessa Echeverria, **Gloria Milena Fernandez-Nieto**, and Simon Buckingham Shum. 2020.From Data to Insights: A Layered Storytelling Approach for Multimodal Learning Analytics. In CHI Conference on Human Factors in Computing Systems CHI'20. 15 pages. doi: 10.1145/ 3313831.3376148.

WORKSHOP PAPERS:

6. **G. Fernandez-Nieto**, K. Kitto, and R. Martinez Maldonado, Four Challenges in Crafting Multimodal Collaboration Analytics for non-Data experts, 2019. In CSCL 2019, Collaboration Analytics. available here

7. Miguel A. Ronda, Olga C. Santos, Roberto Martinez-Maldonado and **Gloria Fernandez-Nieto**. Exploring Emotional Reactions in Teamwork using Multimodal Physiological Data. MAIEd'21. 12 pages.

WORKSHOP:

8. Vanessa Echeverria, Lu Lawrence, Yi-Shan Tsai, Shaveen Singh, **Gloria Fernandez-Nieto**, Roberto Martinez-Maldonado. A Tutorial on Data Storytelling for Learning Analytics Dashboards. LAK21 Workshop.

9. Roberto Martinez-Maldonado1 and **Gloria Fernandez-Nieto**. Multimodal Analytics for Classroom Proxemics. ALASI 2019.

ETHICS:

The studies presented in this thesis were conducted under ethics approved by the University of Technology Sydney's Human Research Ethics Committee, and are based on projects ETH17-1411: Learning Analytics for understanding small-group collaborative processes; ETH17- 1415: Measuring Adoption and Acceptance of Learning Analytics Tools and ETH17-1502: Learning Analytics in clinical simulation. A revised version of the ethics project ETH17-1502 has the protocol number ETH18-2278. The most recent participant information sheets and consent forms can be requested by email1.

One of the studies was run at Monash University as part of the research project "Teamwork analytics in clinical simulation", project ID: 28026. Students and academics consented to participate in both, data collection and follow-up interviews.

SOURCES AND ORIGINAL WORK:

Original material of my own from the above publications has been included in this thesis. Such prior publications when used in the thesis are explicitly cited where appropriate and are not used in entirety. Publications of external authors are credited throughout the thesis with citations in text and reference at the end of the thesis. Figures from external sources were author granted permissions for usage are cited in their captions.

Table of Contents

List of Figures

LIST OF TABLES

CHAPTER 1

INTRODUCTION

We are drowning in information and starving for knowledge.

–Rutherford D. Roger

This thesis aims to support embodied team learning using automatically generated dashboards and reports that facilitate student reflections on performance. This chapter is structured as follows: Section 1.1 presents the context and motivations of the thesis, Section 1.2 explains the research questions driving this research, and Section 1.3 outlines the research goals. Then, Section 1.4 summarises the major contributions of the thesis.

Figure 1.1: Nursing simulations: Example of embodied team activity in Healthcare

1.1 Context and Motivation

Embodied learning foregrounds the importance of considering both the body *and* the cognitive processes involved in learning (Skulmowski and Rey, 2018). Models and theories grounded in **embodiment** aim at *explain interaction* in terms of practical engagement with the social (e.g., communicate with peers) and physical environment (e.g., use resources in the classroom). Embodied learning is critical for a number of professional sectors such as emergency response (Petrosoniak et al., 2019), safety training (Viktorelius and Sellberg, 2022), and healthcare (Cooper and Tisdell, 2020). Such settings (see Figure 1.1) can present unique challenges to students and professionals because they need not only to develop theoretical and procedural knowledge, but also to learn from bodily experiences such as how to use the physical space, how to position themselves in the space, how to interact with various physical and digital devices, and even how to read and act upon their own emotions or stress levels (Kelly et al., 2019). Moreover, such scenarios often occur in multidisciplinary teams, so students are also expected to simultaneously learn effective teamwork and communication skills (Roberts et al., 2022). Furthermore, professional development and training often require that students reflect on critical events or mistakes they may have made (Schön, 1987), but these can occur in rapidly evolving and busy environments where key actions are often difficult to capture.

Some educational institutions are trying to address the above challenges by enriching physical learning spaces with digital technologies and sensing capabilities (Eyal and Gil, 2022) that can be used to capture evidence and generate technology (e.g., dashboards and reports) to support teachers in guiding discussions about embodied team activity. Various sensing devices, such as infrared sensors, video and audio recorders, physiological wearables, and indoor positioning trackers, are being used for capturing real-time multimodal behaviour data from students developing critical skills in embodied team activities (Schneider et al., 2021; Sharma and Giannakos, 2020; Yan et al., 2022). These multimodal data capabilities, embedded in the physical learning environment or worn by students, can provide us with new ways to study key learning processes such as effective teamwork (Cha et al., 2021; Dafoulas et al., 2018; Rosen et al., 2014) and communication (Chng et al., 2022; Olguín et al., 2009; Watanabe et al., 2018; Winder et al., 2020), or the effect of emotions on learning (D'Mello et al., 2018; Villanueva et al., 2018). Similar sensing approaches have also been developed to capture interactions between teachers and students in classrooms using video cameras (Ahuja et al., 2019; Raca et al., 2015) or wearables (Martinez-Maldonado et al., 2021; Saquib et al., 2018).

Beyond educational contexts, various sensing tools have also been created to track people in physical environments, for example, to analyse spatial interactions of people with digital devices (Brudy et al., 2018), and with each other in office spaces (Lee et al., 2021). Together, this set of previous work points to the potential opportunity of using sensor data to help researchers and analysts identify effective embodied learning and teamwork strategies from patterns found in such data. Despite this promise, there is a gap in the literature regarding automation. *Gap 1: the dearth of studies automatically capturing and analysing data from embodied team activity, according to the learning design*. To date little work has focused on automatically supporting *end-users* (e.g., students and professional trainees) by providing data interfaces tailored to their needs, goals, and varying expertise in interacting with data.

Several challenges persist in the field of Learning Analytics (LA) that make it difficult to create effective *data interfaces* to support embodied team learning (Khosravi et al., 2022). First, if the goal is to directly support educational stakeholders (e.g., teachers or students) instead of researchers, the former generally have limited training in data analysis and interpretation (Maltese et al., 2015; Raffaghelli and Stewart, 2020) and thus require feedback and pedagogical explanations rather than interfaces that invite data analysis and exploration (Echeverria et al., 2018a). Furthermore, although there are some tools that enable non-experts to explore and interpret sensor data exist in contexts such as personal health (Raj et al., 2019), physical activity (Tang and Kay, 2017), and smart home monitoring (Moore et al., 2018), the wide diversity of the *multimodal* sensor data that can be captured from a physical learning environment makes it hard to visualise and communicate analytics outputs in ways that non-experts can understand Worsley (2018). A disconnection is also often reported between low-level logged data (e.g., x and y coordinates) and higher-order educational constructs (e.g, closeness to the patient as an association of patient-care) in several data-intensive educational innovations (Milligan and Griffin, 2016; Shute and Ventura, 2013). One of the main reasons for this problem is the intrinsic lack of meaning in the data and their representations (e.g., dashboards, recommendation systems, feedback tools) (Alhadad, 2018), which can be used to make decisions about the learning activity without considering the context in which it unfolds (Knight et al., 2020) or the teachers' pedagogical intentions (Echeverria et al., 2019). This thesis addresses this human-data interaction problem in the context of embodied team learning. These challenges in providing evidence to support reflection shows a second gap in the field. *Gap 2: The lack of studies investigating how insights extracted from learning analytics can be effectively communicated to support reflection.*

Figure 1.2: Thesis diagram: overview of the context, contributions, and validation of this thesis. Small circles in the diagram indicates requirements. For example, the generation of prototypes (RQ2 box) require previous modelling (blue circle) in order to be generated

This thesis engages research in *Multimodal Learning Analytics* (MMLA), *Human-centred Learning Analytics* (HCLA), and *Data Storytelling* (DS). MMLA is a subfield of Learning Analytics (LA) that aims to support teaching and learning by extracting the meaning of physical activity (e.g., embodied team activity) from multimodal data. HCLA approaches support the design of tools that are suitable for the purpose and are pleasant to use (Buckingham Shum et al., 2019b). HCLA approaches can be used to better contextualise multimodal data, as they reinforce the need for human inference and deep awareness of the context, ensuring that teachers and students' voices are taken into consideration (Worsley et al., 2021b). To support multimodal data interfaces for educational stakeholders, this thesis builds on the literature of *Information and Visualisation* (InfoVis) and the **Guidance** visualisation paradigm that aims to support casual users, or users with low analysis expertise, to narrow the gap of data visualisation interpretations (Schulz et al., 2013). **Data Storytelling** (DS) is one way to provide guidance, as a compression technique, to help an audience effectively understand what is important in an visualisation, communicating key messages combining *data*, *visualisations*, and *narratives* (Ryan, 2016). 'Telling stories' with data in these ways enables the elicitation of deeper reflections in an effective manner.

Therefore, this thesis progresses in tackling the different challenges outlined above, by investigating the applied problem that can be formulated as follows: **How can salient aspects of embodied team activity be automatically identified and derived insights be communicated to support timely, productive reflection?** The thesis map in Figure 1.2 represents a summary of this thesis in relation to its context, research questions, contributions, and validations, which are briefly summarised below. The following sections describe the four research questions that motivated this thesis (Section 1.2), the research goals (section 1.3), and the contributions of this thesis (section 1.4).

1.2 Research Questions

Motivated by the two gaps identified in the literature: i) *a dearth of studies automatically capturing and analysing data from embodied team learning*, and ii) *lack of studies investigating how insights extracted from learning analytics can be effectively communicated to support reflection.*, four research questions were derived.

RQ1: What modelling techniques can enable identification of salient aspects of multimodal embodied team learning according to the learning design (teachers' pedagogical intentions)? This question tackles the challenges of contextualising low-level multimodal data within the semantics of embodied activity in a given context.

RQ2: How can insights be extracted from multimodal sensors and communicated to students and teachers to support teaching and reflection on embodied team activity? This second question seeks to explore alternative ways to communicate outcomes from analytics, which may be complex to interpret for non-expert data users.

RQ3: To what extent can students and teachers reflect on embodied teamwork activity using MMLA interfaces? This question addresses the absence of information in regard to the use of Multimodal data stories, specifically, it will explore if the proposed prototypes can support students and teachers in reflecting on their embodied activities.

RQ4: To what extent can MMLA interfaces for students and teachers be automatically generated? The last question addresses the need to validate the feasibility of implementing a solution to scale up the automatic modelling and visualisation process from the beginning to the end. Also, it refers to alternatives to bring teachers into the loop with tools to co-configure the automated feedback they want to provide to student teams, based on the data they have and the learning activity context.

1.3 Research Goals

Against the research questions above, the main goals of the thesis are formulated below.

1. *Formalising stakeholders' requirements into specifications that can be used by the system to contextualise and automate generation of meaningful Multimodal data stories to support students' and teachers' reflections.* To scale-up solutions that can support different requirements of stakeholders. For instance, ideating ways to capture teachers' pedagogical intentions to drive analysis and the automatic generation of tools to support students' reflections.

2. *Implementing modelling techniques to map low-level multimodal data into teamwork aspects.* Eliciting the teachers' requirements may lead to defining specific strategies to interrogate data in search of insights into student behaviour or patterns during a embodied team activity. Theory and the learning design may also lead to the identification of relevant and meaningful teamwork aspects to support productive reflection.

3. *Designing and evaluating alternative ways to communicate multimodal data stories.* Different ways of generating student data stories may reveal how stakeholders interpret them and can suggest the use of specific types of graphs or visuals to communicate particular insights. Exploration of visual alternatives can also provide a more robust corpus of data visualisation types and techniques to communicate data insights.

4. *Providing automatic tools for teachers to define their pedagogical intentions and to generate Multimodal Data Storytelling.* The main goal of this thesis is to investigate how to automatically generate meaningful MMLA interfaces that teachers can take control of to provide relevant feedback to students about their embodied team activities.

1.4 Contributions

This thesis main contributions are briefly described as follows.

1. *A methodology to elicit teachers' pedagogical intentions.* This contribution assists in the formalisation of stakeholders' requirements. It includes a step-by-step methodology and a tool for teachers to capture their pedagogical intentions in a structured way which can be interpreted by the computer to run further analysis and plot enhancements on the visual outcomes (Chapters 4 and 7).

2. *A modelling toolkit to map from low-level indoor positioning data to higher order constructs.* This contribution brings a set of modelling techniques to interrogate multimodal data and find salient aspects that may be relevant for the embodied team activity (Chapter 4).

3. *A set of alternative multimodal visual interface designs to communicate student salient aspects of embodied team activity.* This contribution provides a set of eight MMLA interfaces designed for teachers and students to use for reflection. Four of the alternatives explored included visual elements to enhance the MMLA interfaces and support stakeholders in making sense of them. Four of them were low-fidelity prototypes and four were fully automated and functional prototypes (Chapters 5 and 6).

4. *Qualitative evaluations of MMLA interfaces.* The exploration and perceptions of teachers and students were captured as part of this thesis to investigate to what

extent prototypes are effective in communicating insights about the embodied team activity to support reflection. Five in-the-wild qualitative studies support this contribution (Chapters 5 and 6).

5. *A modular Functional Architecture*. This contribution uses a conceptual Framework (section 7.1), used to map all the modelling techniques and the generation of prototypes, as a reference to define an architecture, which can be a reference for further implementations.

6. *Reference Implementation to automatically generate multimodal data stories.* This contribution demonstrates how an automated solution can be implemented to scale up the automated generation of data storytelling to communicate insights from multimodal data (Chapter 7)

The thesis is structured in eight chapters. *Chapter 2* introduces the literature review and previous work, that supports the identification of the thesis gaps and research questions. *Chapter 3* presents the conceptual framework (section 3.1), that guides the modelling and generations of data stories; the learning contexts (section 3.2), that supports the evidence to answer the different research questions, and the research methodology (section 3.3). *Chapter 4*, presents contribution 1, the methodology to map teachers' pedagogical intentions, and contribution 2, the modelling techniques explored. Contributions 3 and 4 are explained in *Chapters 5* and *6*, where the prototypes designed and the results of their evaluation with students and teachers are described, respectively. Contributions 5 and 6 are introduced in *Chapter 7*, which describes the functional architecture and reference implementation to automate the creation of data stories to support embodied team activity. Finally, *Chapter 8* covers the discussions, limitations of this research, and future work.

CHAPTER 2

Literature Review

Four bodies of research literature provide the context for this thesis: i) embodied team learning, ii) learning through reflection, iii) data storytelling, and iv) multimodal learning analytics interfaces.

2.1 Introduction

There are clearly many forms of professional activity that require the mastery and integration of embodied cognition. While this starts with one's own body, in teamwork, it also implies coordination, practical interactions (social and physical ones), and effective communication with others. Embodied learning is critical for professional sectors such as emergency response, safety training, healthcare, or co-teaching. Training and professional development of students often requires students to reflect upon their behaviours, and potentially upon mistakes that they have made, but in common embodied settings key actions are often hard to capture such that they can provide evidence to support these reflections. Sensor and technology capabilities, embedded in the physical learning environment or worn by trainees, are improving and becoming more readily available. That motivates researchers to exploit the data they provide to study embodied team learning.

This thesis considers different areas of research that attempt to gain insight from multimodal data collected in scenarios involving embodied teamwork. Then it attempts to design and communicate effective evidence of embodied team activity to support students reflection on their physical practices. Figure 2.1, shows the key research areas relevant to this research. In line with the research questions, Section 2.2 introduces foundational research on embodied team learning (Section 2.2.1), the importance of reflection for learning (Section 2.2.2), and Learning Design (LD) (Section 2.2.3). Section 2.3 reports on the foundations of Multimodal Learning Analytics (MMLA) and previous work exploring ways to extract meaning from multimodal data. Moreover, this section presents a review of current Multimodal Learning Interfaces (e.g., visual outcomes or dashboards) to support learning.

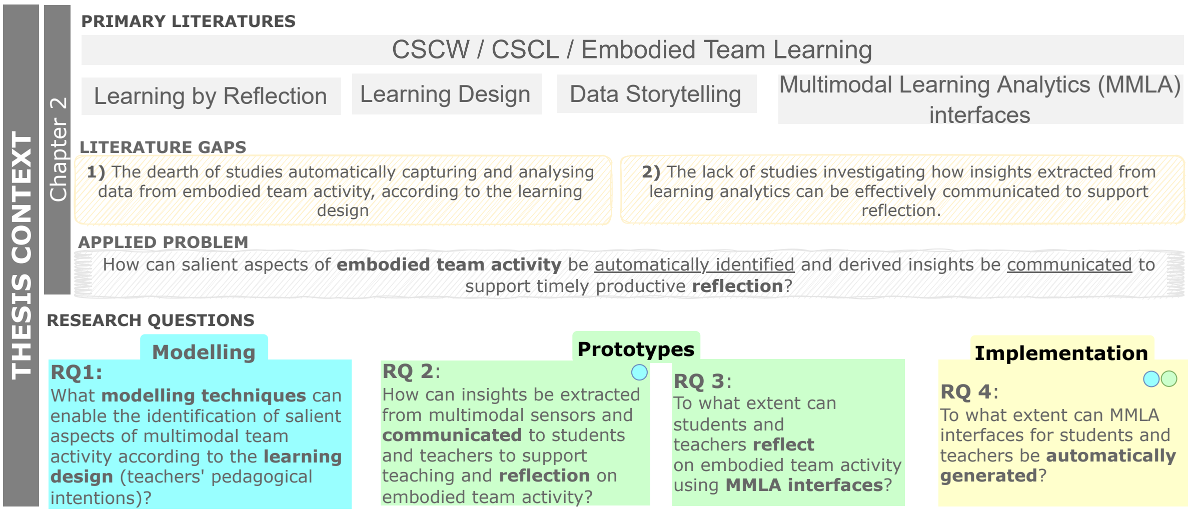


Figure 2.1: Partial thesis diagram that represents the research context and key research questions this thesis addresses.

Next, this chapter presents a review of data storytelling (DS) as a strategy to better communicate complex data to educational stakeholders, using DS to guide their interpretations. Section 2.4 explores the foundations and previous work on DS to guide educational stakeholders in making sense of visual interpretations. Finally, Section 2.5 explains the applied problem that I solved during my thesis and lists the research gaps that I have identified in this previous literature, which in turn have guided the definition of my research questions.

2.2 Background and Foundations

This thesis focused on providing evidence-based tools that are effective in supporting reflection processes on embodied team activity. Thus, this section presents the background and foundations of embodied team learning, the importance of learning by reflection, and the definition of Learning Design (LD).

2.2.1 What is Embodied Team Learning?

Embodiment involves understanding interactions in terms of practical engagement with the social and physical environment. Rogers (2012, Chapter 6) explains that embodiment brings different theories or aspects to account for different behaviours (e.g., in describing what actions occur in a physical shared space and in encouraging students to learn from physical manipulation of objects). Although important, embodied activity is not the only mental but also physical (Stolz, 2015). Being in a collocated scenario highlights the need to consider physical, emotional, and social aspects to understand the learning environment as a whole. Embodied learning is closely connected with the idea of learning by doing and engaging with the environment, which are significant aspects of experiential learning. Examples of embodied learning are activities that focus on the active role of students where emotions play an important role (for example, simulations and role-playing) (OECD, 2018). Because of the possibility for students to practice in real scenarios, embodied learning is particularly suited to teaching creative skills such as curiosity, sensitivity, multi-perspective taking, risk taking, and metacognitive and executive skills that foster learner achievement (Treffinger et al., 2002). A variety of skills and attributes are supported during an embodied team activity, for instance, physical competencies (e.g., coordination and fluency, kinaesthetic awareness), personal qualities (e.g., responsibility and leadership, communication), collaboration (e.g., teamwork), and cognitive skills (e.g., problem solving, decision making).

Teamwork in embodied environments remains an important competency in many disciplines and professions. Although in the literature the terms group and team are sometimes used interchangeably (Cohen and Bailey, 1997), it is important to realise that there are differences between the two. For instance, Forsyth (2018) (p. 351) stated: “Not all groups are teams. Teams require more from members in the way of collaboration and coordination”. This point is emphasised by Salas et al. (2000) (p. 341), who defined a team as “two or more individuals with specific roles interacting adaptively, interdependently, and dynamically towards a common and value goal, within an specific time span” . Importantly, Zoltan and Vancea (2015) extended this position, claiming that a team can be considered a specialised type of group and defined as “a structured group working well-defined goals that require coordinated interactions; and mutual trust and accountability to complete certain tasks” (Forsyth, 2018). This thesis uses the term team to describe a group where specific roles have been defined (e.g., main teacher or team leader), each member has been assigned a role, members exhibit high interdependency on their actions, and a shared goal has been defined toward the completion of the embodied activity.

During an embodied teamwork activity, the role of interaction and collaboration with others is critical (Okita, 2012). Collaborating in physical, face-to-face settings (f2f) provides unique benefits that are not easy to achieve in digitally mediated forms of teamwork (Johnson et al., 2000). The literature suggests that rich multimodal communication channels in the f2f mode can promote social bonding (Nardi and Whittaker, 2001), trust building (Wilson et al., 2006), increase creativity (Gloor et al., 2012), productivity (Olson, 2002), or being an effective member of a team (de Lima and de Souza, 2017). Additionally, teamwork skills include behaviours such as adaptability, performance monitoring, leadership, communication patterns, and interpersonal coordination, as well as the participants' feelings about the team (e.g., team cohesion, mutual trust) (Dickinson and McIntyre, 1997).

The dimension of teamwork can be very difficult to measure and evaluate, as it involves different and complex qualities such as effective communication, team orientation, team leadership, monitoring, and coordination. For example, in their work Cannon-Bowers et al. (1995, Chapter 10) conceptualised and made explicit various team competencies with the aim of specifying training requirements for effective teams. The authors link the different competencies with task and situational factors, making the complexity of teamwork training clear. Considering the various factors teamwork involves in f2f settings, learning to engage in embodied teamwork and teaching it are, therefore, difficult tasks. This is because the activity students engage in during embodied practices is ephemeral. When ephemeral it is difficult for students and teachers to remember what they did in detail, and hence challenging to self-assess. At the same time, teachers must be able to orchestrate their classes and provide effective and accurate feedback to individuals and (often multiple) teams practising at once. But this can be difficult to provide due to the complexity of the situation in which the students are often acting (which can include, for example, multiple teams performing complex activities).

A number of frameworks have been developed to understand both online and embodied team activity. For embodied learning environments, they have been increasingly used to support different dimensions of human teamwork (e.g., cognition, skills, and attitudes). For example, the Blended Interaction Framework (Jetter et al., 2014) structures the design space of Computer Supportive Cooperative Work (CSCW) into four spaces: individual and social interactions, the task, and the physical space. Other frameworks such as Activity-Centred Analysis and Design (ACAD), proposed by Goodyear and Carvalho (2016), provided a three-dimensional view of group and team activity: 1) set, which includes physical and digital spaces and objects (e.g., input devices, screens, software, material tools, or furniture); 2) epistemic, which includes both implicit and explicit knowledge-orientated elements that shape the participants' task and working methods; and 3) social, which includes the variety of ways in which people might be grouped together (e.g., dyads) and the scripted or emerging roles, and division of labour. Another critical foundational theoretical dimension to understand the CSCW work, is the affective aspect, although it can often remain invisible (Roth et al., 2016). From the point of view of Echeverria et al. (2019) group and team activity could include at least four dimensions; physical, epistemic, social, and affective. As will be elaborated in more detail in the following sections and in Chapter 4, this thesis uses the dimensions of the ACAD framework (epistemic, physical, and social) to frame the embodied team activity. In addition, the affective dimension was included, as technology is making it possible to explore inner processes that can impact embodied team activity. To contextualise the dimensions used for this research, the following three subsections introduce the foundational and background work on these four dimensions.

2.2.1.1 The epistemic dimension of embodied team learning

The ACAD dimensional view explained that a learning activity (e.g., embodied team activity) is, at minimum, physically, socially, and epistemically situated (Goodyear and Carvalho, 2016). Goodyear et al. (2021) (p. 448) explain that the student's activity is epistemically situated, which from their view includes:

> "both (a) the role of knowledge-laden task specifications in giving students suggestions about directions to travel and about good things to do on the way, and also (b) a recognition that students, as people, are always already doing various things, in which various forms of knowledge and ways of knowing play a role"

. I agree with the ACAD vision, which argues that student activity is epistemic because what teams do is significantly influenced by the task they are set. Typically, team development patterns are largely the result of environmental factors such as time constraints and task characteristics (Fransen et al., 2013). Given the task at hand and the deadlines to deliver results, solving these task-related problems will have a greater influence on team development than the dynamics of interpersonal relations (Gersick, 1988).

The role of the task in teamwork is mentioned in the Team Evolution and Maturation (TEAM) model (Salas et al., 2009), which combines existing theories and ideas into a general team-development model. In fact, Fransen et al. (2013) explain the stages of teamwork using the TEAM model and the convergence of task-related skills and team-related skills during team maturation. In their view, a team does not have to go through all stages and may start at different stages, according to the past experiences of the team and its members.

In short, taking into account the literature above and previous exploration on the epistemic dimension as a potential modality to gain insight into the embodied team (Echeverria et al., 2019), I acknowledge the important role of the task during the embodied team activity. The expected (e.g., teachers expectations of what students will do in response to a proposed task) or emergent tasks (e.g. individual and team spontaneous tasks performed as a result of the f2f interaction) affected the team results. For that reason, I investigated the epistemic dimension as part of this thesis. This will be developed in Chapters 4, 5, and 6.

2.2.1.2 The social dimension of embodied team learning

Embodied team activity is socially situated, which means that what the team (and the individuals that it is composed of) does is influenced by what people around them are doing (Goodyear et al., 2021). The embodied team activity is intrinsically social because different social aspects take place while an activity is unfolding. For example, teamwork can have social aspects such as cultural norms, interactions with peers and others in cooperative work, definition of roles, division of work, definition of learning networks with peers, formation of social arrangements, or definition of individual and team agreements on how to achieve team tasks (Goodyear and Carvalho, 2016).

In an attempt to use technology to understand the social aspects of teamwork, Erickson and Kellogg (2000) developed the notion of Social Translucence. It refers to the use of computer-mediated systems to provide social cues (visibility, awareness, and accountability) that can be lost as a result of moving away from interaction in physical spaces into the digital sphere. Social Translucence is relevant because it aims to make people's activities visible to others. In physical spaces, invisible aspects of the social dimension, which includes the variety of ways in which people might be grouped together

(Echeverria et al., 2019), might include: people activities, team behaviour, scripted or emerging roles, communication among peers, divisions of labour, or social presence.

This thesis explored different social aspects of embodied team activity, including physical social presence (developed in Chapter 4.2.2.2), social arrangements (developed in Chapter 4.2.2.3), and scripted and emerging roles (developed in Chapter 4.2.1). Although, for this thesis, audio was collected with the intention of understanding communication patterns during embodied team activity, none of the modelling techniques discussed in Chapter 4 focused on audio.

2.2.1.3 The affective dimension of embodied team learning

Although the ACAD framework did not explicitly consider the affective dimension, it considers the mental and emotional aspects in its definition of *activity*, referring to it as "what the student is actually doing during a period of time in which they are meant to be learning something" Goodyear et al. (2021) (p. 446). Affective aspects have been identified in the foundational theoretical work (Vygotsky, 1997) as critical to understanding the work of CSCW, although they can often remain invisible (Roth et al., 2016). For example, previous studies have found that physiological features (e.g., electrodermal activity, hypoxic ventilatory response) may serve as indicators of affective states at individual (Haataja et al., 2018; Neubauer et al., 2016) and group levels (Dich et al., 2018; Haataja et al., 2018). Other physiological signals (Galvanic Skin Response -GSR-, Heart Rate -HR-, Respiration Rate -RR- and Skin Temperature -ST-) are also used to identify emotions (Basu et al., 2015). The affective dimension is of interest to researchers as it is now possible to shed light on affective aspects using video recordings or physiological sensors.

Physiological data can provide a deep understanding of the affective dimension of teamwork. This is because changes in physiological data can influence student decision making, perception, human intelligence, or team interaction. Realistic simulated scenarios can lead teams to: experience different emotions (e.g., anger or happiness) in safe scenarios, and achieve their learning goals (Berragan, 2011). These stressful experiences are meant to reflect the types of pressure students will encounter in authentic workplaces, thus helping to induce authentic training experiences. There is strong evidence that stress has an important effect on student engagement and motivation and, consequently, influences learning outcomes (Pearsall et al., 2009; Rousseau and Aubé, 2010). Teamwork relies heavily on handling various aspects such as sharing work load (Carvalho et al., 2007). Human workload is related to human performance (Desmond and Hoyes, 1996), for example, an appropriate mental work load could reduce human error and enhance team performance (Skitka et al., 2000).

However, the *stress* generated in these close-to-real situations can also be confronting and hinder learning. For example, according to Schwabe et al. (2012), stress markedly impairs memory retrieval, causing, for instance, the risk of underachieving tasks. High levels of stress can cause decreases in motivation leading to low academic performance (Pascoe et al., 2020). As stress and cognitive load can affect embodied teamwork in positive or negative ways, a deeper understanding of aspects such as stress levels during teamwork activity can drive future research on the implications of stress in learning.

Therefore, considering the role of aspects such as stress and cognitive load in learning during embodied team activities, the affective dimension was considered as part of the exploration of this thesis. This work will be elaborated on in more detail in Section 4.2 and Sections 5.1 and 6.3.

2.2.1.4 The physical dimension of of embodied team learning

The physical dimension, also presented in the ACAD framework (Goodyear, 1991), was revised and is part of the context of this research to explore and better understand the physical aspects of the embodied teamwork activity. The embodied team activity is physically situated and involves tools and other resources that come to hand in specific spaces (e.g., simulation wards or laboratory), the space, and interactions happening in the space affect the activity which is being performed. The modelling of the physical dimension for this thesis was mainly grounded on the Theory of Proxemics, which will now be discussed in detail.

A team's performance depends not only on how members perform their individual task, but also on how they coordinate and cooperate with each other, and how effectively they use the resources and space available. Extensive research has illuminated more precisely how this works. For instance, in the review conducted by Hall et al. (2018), the authors illustrated how teams that have worked together to successfully build trust and develop an understanding of who knows what, who does what, and how things are done may facilitate high productivity. This is partly why proxemics has been used as a lens to study complex and dynamic interactions between team members who embody a teamwork activity, particularly in contexts where the use of space is critical to complete certain tasks.

Proxemics can be broadly defined as the study of the ways people use physical spaces and interpersonal distances to mediate interactions according to their cultural context (Hall, 1966b). The theory of proxemics has been applied to a variety of fields (such as robotics (Fiore et al., 2013), architecture (Hall et al., 1968), education (Martinez-Maldonado et al., 2020b), and urban planning (Harvey, 2005)) with the purpose of analysing spatial behaviours and mobility dynamics in both indoor and outdoor spaces (Dael et al., 2015). For instance, patient and doctor' interactions can be measured based on body distances and rotation to assess how patient satisfaction is affected by non-verbal cues (Montanari et al., 2018). When proxemic constructs are integrated into the work of an ethnographer, it is possible to generate a better understanding of the physical dispositions of people that naturally emerge during interactions in relation group tasks and the physical structures of the collaborative space (Marshall et al., 2011). These observations can lead to the identification of spatial behaviours and tools that maximise connectivity and interaction possibilities (Marquardt and Greenberg, 2015).

In fact, in the proxemics lexicon, Ciolek (1983) defines constructs which encapsulate different aspects of the study of space in social contexts. For example, Cristani et al. (2011) used the notions of interpersonal spaces and distances to identify the establishment of social ties among group members. Setti et al. (2015) focused on the proxemic construct of formations or arrangements, based on people's proximity and body dispositions, to examine how people establish conversation groups in informal settings.

In the teaching and learning space, Martinez-Maldonado et al. (2020b) proposed the notion of classroom proxemics, to explore the meanings that certain spaces in the classroom take up according to the proximity of teachers' teams to students and classroom resources (e.g., desks, whiteboards, and personal computers). This is something that can be explored using various analytics services.

In short, proxemics have been used as a lens to analyse socio-spatial interactions in embodied teamwork settings. I will refer to this particular application of proxemics as **teamwork proxemics**. This thesis will build on the proxemics lexicon to identify critical *proxemic constructs* that can guide the modelling from indoor positioning data to meaningful representations of spatial team behaviours (see modelling in Sections 4.2.2.2, 4.8, 4.2.2.4 and prototypes in Section 5.2).

2.2.2 What is learning through reflection?

According to Dewey (1997), **reflection** involves retrospective observation of the experience to discern explanations for what happened. Reflection is defined as "a personal process that can deepen one's understanding of self and can lead to significant discoveries or insights" (Desjarlais and Smith, 2011)(p. 1). Reflection is a process that involves playing back a period of time related to previous valued experiences in search of significant discoveries or insights about oneself, one's behaviours, one's values, or knowledge gained. An important goal in reflection is to bring attention to an indeterminate situation (Dewey, 1986) by gaining clarity and fully experiencing what has happened. It is important to gain closure during reflection and not to ruminate repeatedly about the experience.

Raelin (2002) and Amulya (2004) both stress the importance of reflective practice but propose different reasons why it is important. Raelin (2002)(p. 1) describes reflective practice as "the exercise of periodically stepping back to ponder the meaning of what has recently transpired to ourselves and to others in our immediate environment". He presents it as a public and open process by which an individual's interpretations, evaluations, and assumptions are subjected to the review of others in order to avoid bias and errors in perceptions of reality. Amulya (2004), on the other hand, focuses more on the process in general and less on whether it is an individual or collective experience. She states that the purpose of reflection is to learn from experiences. She describes certain experiences that can provide learning opportunities through reflection: struggles, dilemmas, uncertainties, or breakthroughs. In fact, recent work on reflection emphasises the need to study reflection "as an individual and social process" (Noffke and Brennan, 2005)(p. 74). In line with the literature, I have investigated reflection sessions with both individuals (one student / teacher at a time) and teams (all team members).

Additional work on reflection explained different moments when this process can take place. Schön (1987)'s influential work on "reflective practitioners" contrasts reflection-in-action and reflection-on-action as critical components of the development of expertise. Reflection-in-action takes place during an action, and reflection-on-action takes place after an event has occurred. Key competencies can be developed while the embodied team activity unfolds (Schön's notion of reflection-in-action) as has been demonstrated in teaching, healthcare, and team training (An et al., 2019; Sellberg et al., 2021; Zhang and Sarcevic, 2015). However, Schön emphasises the importance of reflection-on-action to recognise how our knowing-in-action may have contributed to unexpected outcomes or how it can contribute to the further development of competencies to be demonstrated in future events. For example, teachers often reflect on evidence after a class to understand how their actions may have impacted student learning (Russell, 2003). This thesis focusses on understanding how students reflect after their embodied team activities. I have chosen to do this because the embodied team activities that I have investigated may benefit from this type of reflection.

Becoming a reflective practitioner implies the need to actively strive to continually improve your practice (Helyer, 2015). However, significant challenges remain in generating effective ways to support active reflection on embodied activity. Researchers have explored the kind of reflection processes that commonly occur in physical spaces. Many of these studies have mentioned that reflection in the classroom is low-level (Ingvarson and Rowe, 2008), and generic (Brooks et al., 2019). In fact, in the work presented by Hattie and Yates (2014), about ways to promote learning (e.g., reflection scenarios), the authors explain how students perceive them as nil while teachers allege promoting them relatively high-level and routinely. A possible explanation for these contradicting points of view, concern the assessment of team processes. In physical spaces, it is difficult to monitor the multiple interactions that occur among team members. Researchers have proposed different approaches to understand team dynamics (e.g, salient aspects of embodied team activity) with the aim of generating solutions to help teams learn and work effectively (Forsyth and Diederich, 2014; Rousseau et al., 2006). While direct observation is one approach, consisting of manually documenting interactions (Goodyear and Carvalho, 2014), this has serious limitations: is time consuming for teachers and probably does not always represent behaviour in its totality, but only an extract of what happened during the task. Another approaches focused on video-based products to support this reflection, they are often impractical for classroom use, resulting in students rarely using this evidence to inform reflection (Mariani and Doolen, 2016).

Given the limitations of direct observation and video-based, a complementary strategy is to collect evidence from the embodied team activity using advanced technology and software developments such as sensors and pervasive devices. Technology provides opportunities to deliver scalable tools for reflection on embodied team activity. Although reflection has been identified as a potential process to support learning there is a need for automated evidence-based reflection tools in physical spaces to support embodied team activity. Further sections in this chapter will illustrate the role of technology in supporting teaching and learning.

2.2.3 Learning Design

Learning design (LD) can be used to refer to the artefact or product resulting from the design process, which may be a plan or a formal description of the sequence of learning activities that can be carried out on an e-learning platform (Oliver et al., 2007). It can also be seen as a resource to help teachers devise effective learning experiences aimed at achieving defined educational objectives in a given context (Mor et al., 2015). In fact, according to Conole (2012)(p. 118),

> "learning design has developed as a means of helping teachers make informed choices in terms of creating pedagogically effective learning interventions that make effective use of new technologies. Learning design representations enable teachers to document, model, and share teaching practice. Learning design also encompasses both the process of designing learning experiences as well as the product, that is, the outcome or artefact of the design process".

This resource could also be used to assist teacher enquiry into student learning as an integral part of teacher learning (Hansen and Wasson, 2016). For this thesis, the LD is considered as a strong artefact, which contains teachers' plans for a set of learners to achieve specific learning outcome goals (including the necessary environment and resources). The LD is relevant for this thesis because it provides the educational objectives that contextualises the embodied team activities.

A learning design can represent different levels of granularity, from a whole course down to an individual learning activity (e.g. simulation). In addition, it can be a formal representation which is computer-runnable or simply a semiformal way of describing the learning intervention. Yang and Goodyear (2014) use the related term educational design, which they define as "the set of practices involved in constructing representations of how to support learning in particular cases or the set of practices involved in constructing representations of how people should be helped to learn specific circumstances" (p. 167). This also means that part of the plan incorporates the different means used to promote learning in class, for example, by leading a debriefing session using specific tools. In that regard, I consider the LD important for teachers to plan how to include technological resources to support their practice (e.g. using a data visualisation resource to guide discussion in class).

In this thesis, the learning designs of each learning context used to validate and evaluate the research questions will be described in detail in Section 3.2. These learning designs were used to inform both the modelling techniques (Chapter 4.2) and the design (Chapters 5 and 6).

The next section turns to the use of Learning Analytics in physical spaces and visual interfaces to deliver technological solutions to the problems that have been discussed here. Both opportunities and challenges will be discussed later in the revision of previous work.

2.3 Data-driven aids to reflection on embodied teamwork

Having introduced embodied teamwork, reflection, and learning design, I now turn to the way in which data can be used to support an effective reflection process on embodied teamwork. Specifically, this section focuses on previous research that uses technology for educational purposes (teaching and learning) in physical spaces to support teamwork. The following sections provide a guide to allow the reader to recognise the gaps in the existing body of work, which in turn motivates the four research questions that drove my project (see Section 1.2).

According to the Society for Learning Analytics Research (SOLAR)[1], Learning Analytics (LA) refers to the measurement, collection, analysis, and reporting of data on learners and their contexts, for the purposes of understanding and optimising learning and the environments in which it occurs. SOLAR claims that:

> "Learning Analytics sits at the convergence of Learning (e.g., educational research, learning and assessment sciences, educational technology), Analytics (e.g., statistics, visualisation, computer/data sciences, artificial intelligence), and Human-Centred Design (e.g., usability, participatory design, sociotechnical system thinking)."

LA solutions began, and are still most widely associated with, analysing student activity within online learning technologies accessed via desktop, laptop, and mobile devices (e.g., Learning Management Systems, online tutors, online courses, etc.). However, during the past 10 years interest has grown in analysing embodied, collocated learning in physical spaces (e.g., classrooms, laboratories, class practices etc.), reflected in the emergence of Multimodal Learning Analytics (MMLA). The following sections describe previous work on MMLA highlighting the ways in which it has yet to fully support embodied team activity.

2.3.1 Multimodal Learning Analytics

MMLA supports learning beyond the clickstreams and keystrokes of conventional personal and collaborative computing. Although multimodal analysis has always been part of the LA agenda, it was formally defined in 2013 by Blikstein (2013) (p. 105), as:

[1] https://www.solaresearch.org/about/what-is-learning-analytics/

> "a set of techniques that can be used to collect multiple sources of data in high frequency (video, logs, audio, gestures, biosensors), synchronise and code the data, and examine learning in realistic, ecologically valid, social, mixed-media learning environments."

In MMLA approaches, sensor data is collected from spatial environments and bodies, coupled with analytics informed by educational theories to understand and support learning and teaching processes (Di Mitri et al., 2018). Data collected from different types of devices, measurement techniques, experimental setups, and other sources are called *multimodal data*.

A multimodal approach, for Learning Analytics purposes and for this thesis, recognises the plurality of approaches that learners can use to demonstrate or experience learning (Worsley et al., 2021b). Thus, I argue that by expanding the set of modalities, I might better understand complex phenomena such as embodied teamwork and support the automated generation of evidence-based reflection tools that people can interact with (e.g., through haptic, audio, and visual mechanisms). In this way, sensors can make certain aspects of teamwork visible and persistent that otherwise remain invisible and ephemeral. Furthermore, each modality can add value in improving the interpretability, robustness, and uniqueness of the whole phenomenon or system that has been observed (e.g., learning) (Worsley et al., 2021b). A range of benefits of multimodal data for education are described in Di Mitri et al. (2019), who propose that, in principle, multimodal data can (i) allow institutions to complement the digital representation of students and (ii) add additional information on cognitive states or metacognitive factors that influence learning (see Chapter 4 for details on multimodal data modelling techniques). Therefore, MMLA uses and triangulates between traditional and non-traditional forms of data to characterise or model student learning (learning behaviour detection and learning construct estimation) in complex learning environments (Worsley, 2018). Evidence about different modalities of student interaction enables the creation of new ways of studying learning in physical contexts, such as understanding how embodied team behaviours relate to learning outcomes (Blikstein and Worsley, 2018) or finding patterns that can be used to personalise instruction (Sottilare et al., 2017). For this thesis, the MMLA approach aims to provide a deeper understanding of both learning and teaching processes during embodied team activity, as is now becoming possible.

Applications of MMLA techniques in diverse learning settings have demonstrated the potential to use them to understand collocated phenomena. For instance, Ochoa et al. (2013) use video, audio, and pen strokes to extract features that discriminate between experts and non-experts in groups of students solving mathematical problems. In their results, the expertise of students was estimated using: i) the percentage of time that the students use the calculator, ii) the speed at which the student writes or draws, and iii) the percentage of time that the student mentions numbers or mathematical terms. Similarly, Prieto et al. (2016) use multimodal features and models to extract teaching activities (e.g., explanation vs. questioning). For this research, the teacher wore several sensors (using a single-electrode portable electroencephalogram -EEG-, mobile eye tracking Googles, plus a smartphone) to capture relevant teaching practice data. These previous work exemplify the work in MMLA and illuminate various challenges of multimodal approaches. I have identified two main challenges (Fernandez-Nieto et al., 2019) i) the challenge of focusing too much on sensors and multimodal data rather than the purpose of research, to interpret multimodal data in the lens of educational needs and ii) the challenge of creating effective multimodal interfaces for non-data experts, referring to the way multimodal data should be fused, processed, modelled, and effectively communicated to support user interpretations.

To create MMLA interfaces for non experts we first need to give meaning to multimodal data, and then work on interfaces that this user group can easily use to support their practice. The next sections present previous research completed within the MMLA community to partially address these challenges. Then, I will introduce the gaps that I have identified in the literature on MMLA to provide tools to support reflection on embodied team activity.

2.3.2 Giving meaning and modelling multimodal data

This section introduces the work completed to tackle one of the challenges in the materialisation of MMLA solutions, namely understanding multimodal data for educational purposes.

To date, most of the current work in MMLA has adopted a bottom-up approach. This means that the focus has been on what sensors can tell us about collaboration or individual learning rather than on identifying higher-order constructs that are important in educational terms (Prieto et al., 2017). For the case of embodied team activity, a bottom-up approach means that researchers may want to collect sensor data first (e.g., posture or movement using a Kinect sensor or voice or conversation patterns using a microphone) to then use analytics (e.g., machine learning algorithms or statistical analysis approaches) to interpret (e.g. classify) behavioural patterns and associated them to team dynamics. This approach gives prominence to technology and leaves aside stakeholders and educational requirements. In contrast, a top-down approach would focus on understanding the purpose of research and prioritising higher-order educational constructs over technology (Prieto et al., 2017). In a top-down approach, researchers would invite stakeholders to participate in the definition of potential MMLA solutions. In this way, the technical aspects, such as sensors, data, and software to be used to analyse such data, would be shaped based on the learning design and the actual needs of teachers and students.

There is a small amount of work in MMLA focused on making sense of multimodal data to provide meaningful educational insights that can be relevant to teachers and students in the context of their embodied activity (Yan et al., 2022). Specifically, three approaches have worked to provide meaning to multimodal data collected from physical spaces. Such approaches consider human inputs in the modelling or the use of constructs to contextualise raw data.

The first approach focuses on the human annotation of multimodal data to find meaningful patterns (Di Mitri et al., 2019), through the Visual Inspection Tool (VIT). This allows users to include annotations in time intervals to express their interpretations of portions of multimodal data. In this way, the data captured are contextualised with human observations by providing a meaning according to the embodied activity. A second approach proposed by Worsley (2014) makes use of the notion of "Epistemic Frames" (EF) as a way to characterise combinations of student features (e.g. by fusing data such as posture and gaze) into frames. Epistemological Framing "is a construct developed in anthropology and linguistics to describe how and individual or teams forms a sense of 'What is it that is going on here?' " (Scherr and Hammer, 2009)(p. 149). Using EF Worsley (2014) present an algorithm to (a) identify multimodal behaviours (e.g., using human hand/wrist movements into gestures instead of using raw data), (b) compare this behaviour over time, and (c) seek to understand the behaviour exhibited by different individuals (under experimental conditions). Although this research focused on identifying correlations between multimodal data and learning outcomes, it also incorporates constructs to gain meaning from raw data. Buckingham Shum et al. (2019a) propose the need for a careful mapping from multimodal data to High-Order Constructs (HOCs) to give meaning to multimodal data during the design process rather than in use. In their research, the authors introduce the notion of a **Multimodal Matrix (MM)**. The MM is an *M x N* data structure in which each data modality M is coded into N columns of the matrix, which are called multimodal observations. This research also mentioned that complex contexts, such as education (teaching and learning) and social sciences in general, require mix-method methodologies, because, this kind of methodology brings additional perspectives for data interpretation. In the context of MMLA, multimodal data can also be understood using qualitative analysis, which can be associated with constructs that humans can easily interpret (Johnson et al., 2007; Onwuegbuzie and Leech, 2005). For example, the MM is designed so that both machines and humans could write new columns of interpretations.

The first two approaches were used for research purposes in controlled settings, but not in a real-life scenario. Moreover, only the third example explicitly refers to the embodied team activity. This research thesis builds upon previous approaches that aimed to connect or interpret multimodal data to provide meaningful educational information, specifically in the work by Buckingham Shum et al. (2019a). However, finding new approaches to support non-data experts in making sense and understanding multimodal data about multiple collaborators is a required step towards the creation of new ways to facilitate the provision of feedback and assessment by using MMLA interfaces. Therefore, the first gap from the literature motivating this thesis is described as follows:

Gap 1: the dearth of studies automatically capturing and analysing data from embodied team activity, according to the learning design.

For this thesis, the LD is a tool that provides context and meaning to multimodal data. Commonly, the LD also defines the learning objectives or goals of the embodied team activity. Thus, including the LD in the process of gaining meaning of multimodal data may potentially generate relevant and meaningful aspects of embodied teamwork to support learning and teaching goals. Therefore, considering the gap above and the opportunities explored, leads to the first research question explored in this thesis:

RQ1: What modelling techniques can enable the identification of salient aspects of multimodal team activity according to the learning design?

In chapter 4, a set of techniques to model multimodal data is introduced as a contribution of this thesis.

2.3.3 Multimodal Learning Analytics Interfaces

Few MMLA interfaces have been designed for teachers or students, user groups who are not necessarily skilled in the interpretation of data. The data novice status of these users makes it especially challenging to automatically provide interfaces based on multimodal data that provide feedback about performance in embodied teamwork settings. Even less emphasis has been given to providing support to students that prompts reflection so assisting them with learning how to become part of a high performing team.

The analysis of multiple data streams introduces challenges in data integration, interpretation, and visualisation (Blikstein and Worsley, 2018). MMLA tools can very easily generate complex interfaces, which explains, in part, the dearth of suitable MMLA user interfaces for teachers and students. Another explanation for this scarcity is that the infrastructure is not yet widespread. Nonetheless, as MMLA continues to grow as a field, it is likely that more interfaces will emerge to support students and teachers. For example, Echeverria et al. (2019) designed four visualisations, each related to one modality: speech (Figure 2.2); arousal (Figure 2.5); positioning (Figure 2.3) and logged actions (Figure 2.4), but did not combine them into a single interface to facilitate reflection.

Ochoa et al. (2018) attempted to create interfaces by visualising logs of students' activity around a tabletop. The multimodal data used included logged actions, verbal participation, gaze direction, and emotional traits. The initial teacher feedback was positive, but the prototype was not evaluated in authentic contexts with the students (see Figure 2.6). The preliminary work of Vujović and Hernández Leo (2019) investigated how to compress arousal and noise information during meetings (see Figure 2.7), but this interface was designed for interpretation by educational researchers, not teachers or

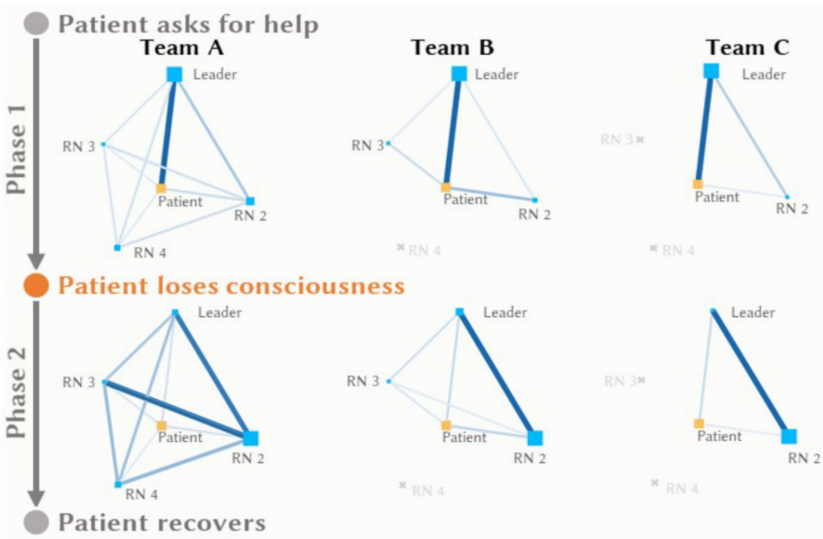


Figure 2.2: High-fidelity prototype of Multimodal Interfaces. Social proxies in the physical space. The orange node represents the patient and blue nodes the nurses. Edges represent communication among nurses and with the patient (Echeverria et al., 2019)

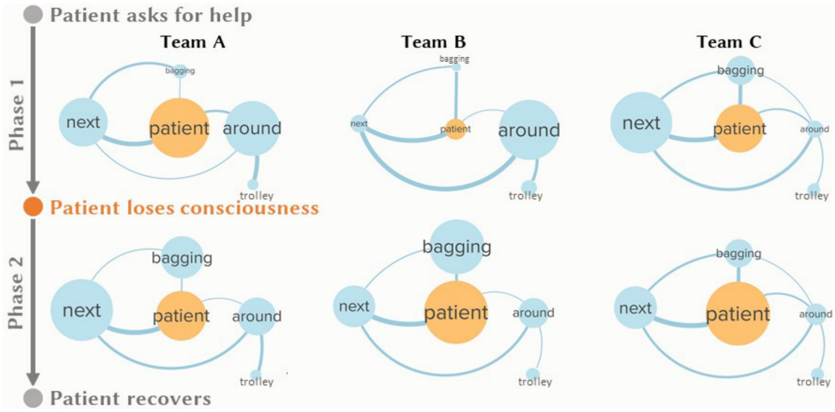


Figure 2.3: High-fidelity prototype of Multimodal Interfaces. Physical localisation and proximity proxies. Circles represents zones of interest around the patient's bed, while links represent transitions among zones (Echeverria et al., 2019)

students.

In contrast, a tool designed to provide support for teachers called EduSense (Ahuja et al., 2019) provided classroom data to teachers about their own practices. Using networked 4K cameras to capture audio and video, body and face keypoints are visualised using icons to represent students' hands raised, upper body poses, smiling, mouths open, and sitting/standing (see Figure 2.8). EduSense unified a solution to provide measures such as total hands rise, total silence, total student speech, and total instructor speech. Although it was implemented in controlled and in-the-wild scenarios, there are no automated interfaces to support teachers and students in their practice yet. Recently, the work of Alzoubi et al. (2021) has explored the design of dashboards using the data generated by EduSense. The interfaces used in this research are evolving mock-ups that

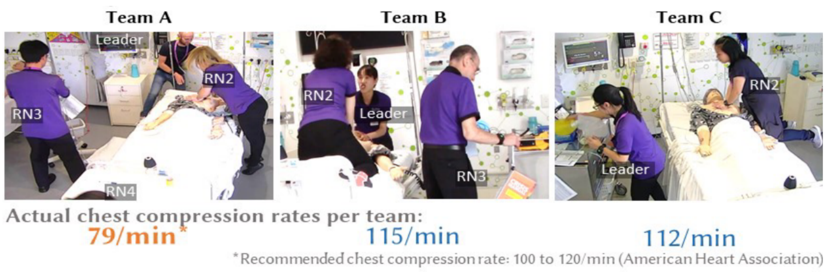


Figure 2.4: High-fidelity prototype of Multimodal Interfaces. Log data, three different ways in which nurses performed chest compressions (Echeverria et al., 2019)

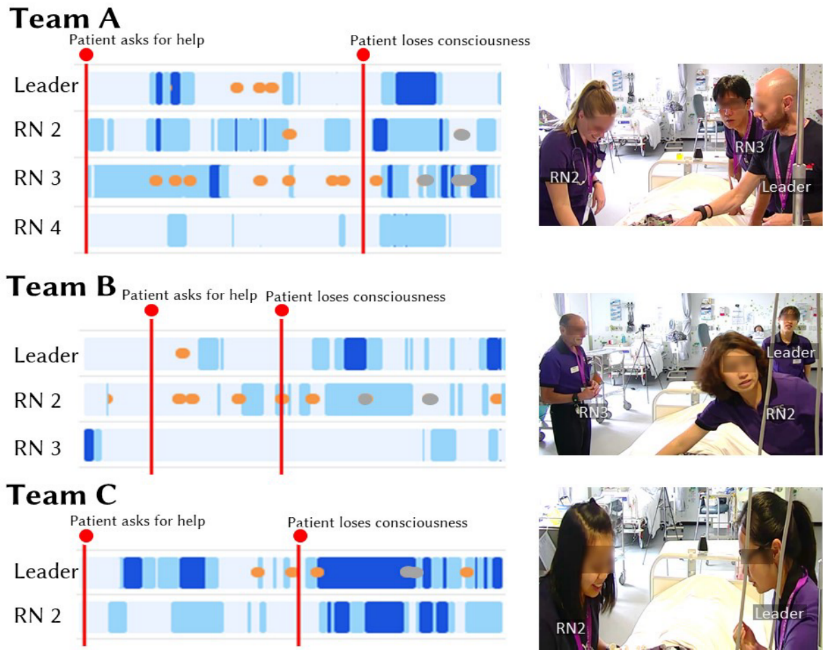


Figure 2.5: High-fidelity prototype of Multimodal Interfaces. Timeline proxies of EDA indicating: EDA peaks (orange dots); b) physical intensity (represented by different shades of blue); and c) EDA peaks that may be affected by intense physical activity (grey dots). (Echeverria et al., 2019)

have been evaluated with teachers but have not been automated or used in the classroom to support reflection.

The work presented by Di Mitri et al. (2021) implements a solution to provide feedback to students on cardiopulmonary resuscitation (CPR) performance, using 3D kinematic data of body joints and electromyographic data (EMG sensors) to recognise the movement of arms and hand gestures. Feedback consists of vibrations and audio indicating aspects about compression such as depth or the use of body weight. From the qualitative evaluation of the feedback tool with students, some aspects, such as the clarity or opportunity of feedback, need to be improved.

Continuing the focus on individual student support, but for slide presentations, Domínguez et al. (2021) uses audio, video, and a slide deck presentation to provide automatic individual visual feedback on aspects of their presentations, regarding posture, gaze direction, voice volume, and slide quality. For posture and gaze, a classifier is used to indicate 'good' or 'bad' performance (e.g., open postures are a 'good' feature), as well as pie charts indicating the percentage of time that students were looking at certain locations

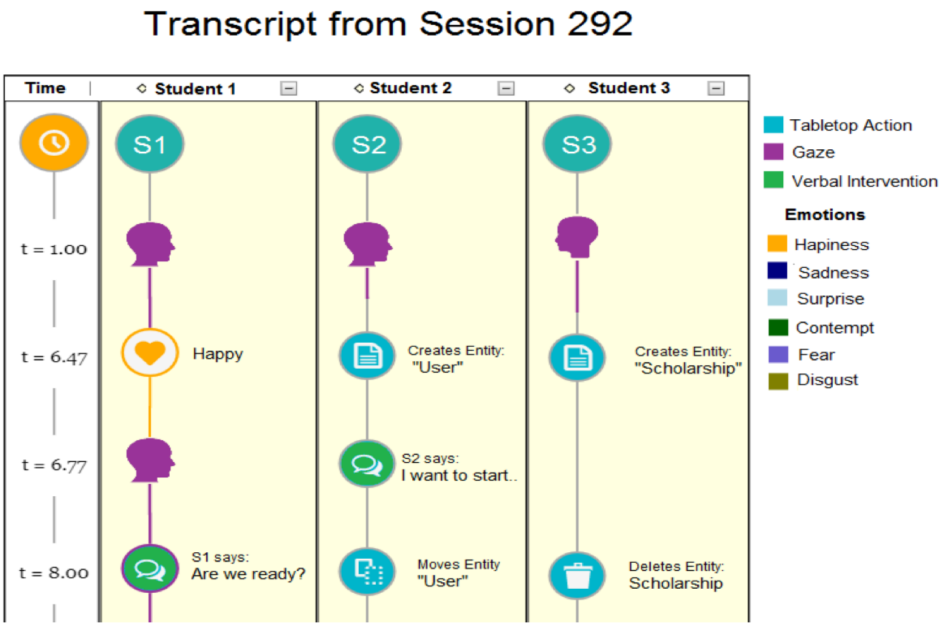


Figure 2.6: An excerpt of the multimodal transcript from a group interaction (Ochoa et al., 2018)

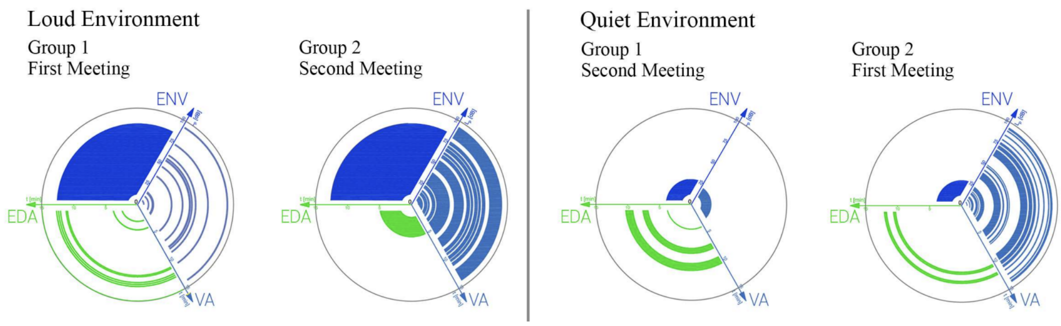


Figure 2.7: Visualization three measured parameters in a unique graphic presentation (Vujović and Hernández Leo, 2019)

(e.g., the audience, as seen in Figure 2.9), while a classifier categorizes volume of a voice as high or low. In the same way, the presentation training system (Sensai) (Kurihara et al., 2007) using a microphone and a camera provides feedback to the presenter about speed, pitch, and eye contact, combining speech and image processing techniques (Figure 2.11). These last three examples illustrate the potential of MMLA but the reports are restricted to individual student behaviour, rather than providing feedback on teamwork.

However, there is growing interest in using multimodal sensor-data to disentangle teamwork. In their exploration of prototypes for research purposes, Worsley et al. (2021a) used a visual tool (BLINC) to instantiate their view of different collaboration concerns elicited from 131 university students. Using audio, the authors capture conversations during team activity and display approximated visual indicators about when the group was discussing, asking questions, or being silent, as well as direction of the conversation (e.g., front or back) and keyword detection. Similarly, Praharaj et al. (2021b) prototyped a dashboard to explain what students talk about in collocated collaboration settings. In their approach, the authors analyse the student's conversations by visualising, in an interactive network graph, the strength of the linkage between words and phrases (see Figure 2.10). The authors indicated that their dashboard can be easily scaled and is fully dynamic and interactive for research purposes. However, the authors also acknowledge the need for an additional evaluation of how teachers or team members make sense of the dashboard. However, the complexity of multimodal data means that the interfaces generated were mainly explored for research purposes, and less work has worked to find methods for building interfaces that support productive students' reflections on complex phenomena such as embodied team activity.

Although previous research on MMLA has focused on gaining meaning from low-level data mainly for research purposes, less attention has been paid to the design aspects of the interfaces used to communicate insights, and the interpretation difficulties of educational stakeholders that have been documented in the literature (Bodily and Verbert, 2017b; Landherr et al., 2010; Verbert et al., 2013). Corrin and De Barba (2014) found that students often find LA dashboards very *unfamiliar*, making it *difficult* for them to understand. As reported in the review conducted by Bodily and Verbert (2017a),

Figure 2.8: Top row: Example classroom scenes processed by EduSense (image data is not archived; shown here for reference and with permission). Bottom row: Featurized data, including body and face keypoints, with icons for hand raise, upper body pose, smile, mouth open, and sit/stand classification (Ahuja et al., 2019).

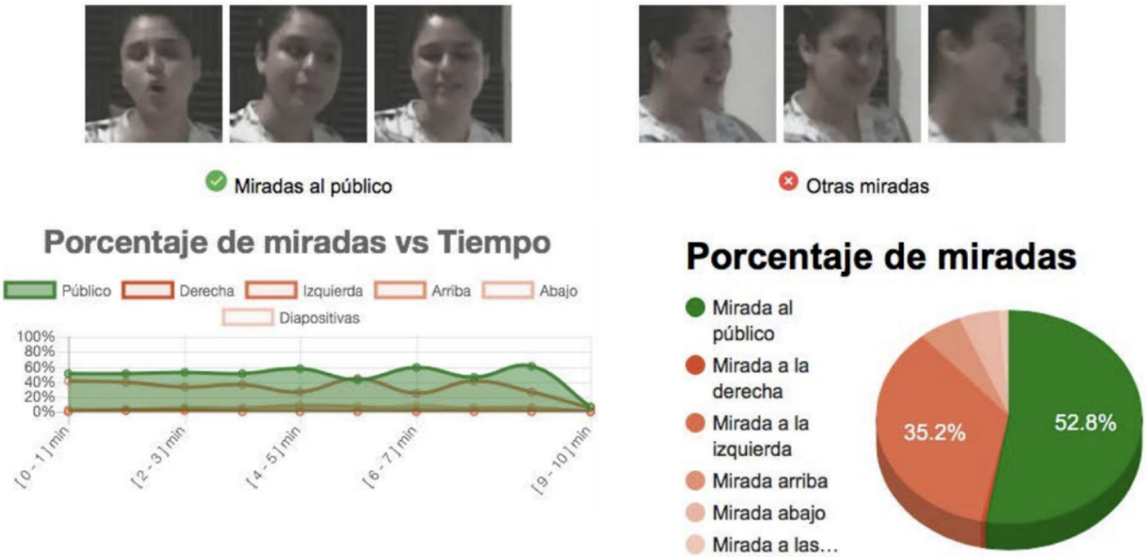


Figure 2.9: Gaze direction section of the original feedback report (Domínguez et al., 2021).

only 6 articles have generated tools for students to provide relevant learning material, which is critical to achieving the students' goals. Likewise, Pokorny and Pickford (2010), identified very interesting elements when students talked about their perceptions of visual interfaces in the classroom. For example, es explained by Pokorny and Pickford (2010) students felt that they did not have the possibility to compare their performance with standards, they found very difficult to use general feedback (e.g. number of visits to the material) to evaluate their learning goals.

This lack in the existing literature opens up the second gap that motivates this research:

Gap 2: the lack of studies investigating how insights extracted from learning analytics can be effectively communicated to support reflection.

In short, the current interpretation challenges found in many student/teacher —facing LA interfaces, in general, are even more profound for the case of MMLA innovations, which commonly deal with complex, heterogeneous data streams. There is a small but growing interest in creating MMLA interfaces for non-data experts (Echeverria et al., 2019). Therefore, my second research question is motivated by the lack of studies investi-

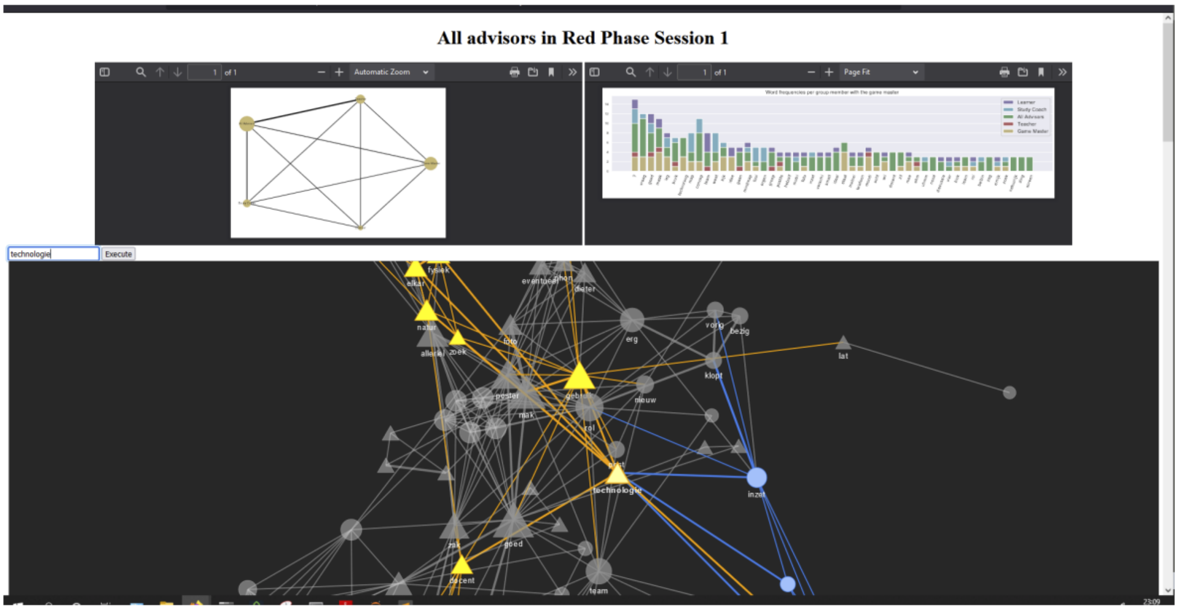


Figure 2.10: Dashboard with social and epistemic component of students' conversations (Praharaj et al., 2021b)

Figure 2.11: Online feedback. (Left) Real time monitor. (Trafic signals) Visual Alerts (Kurihara et al., 2007).

gating how insights extracted from learning analytics can be effectively communicated to support reflection.

RQ2: How can insights be extracted from multimodal sensors and communicated to students and teachers to support teaching and reflection on embodied team activity?

Chapters 5 and 6 describe a set of MMLA interfaces crafted for students and teachers to reflect on embodied team activity.

The work presented in the next sections addresses the lack of guidance for students to interpret and reflect upon their own multimodal data in the context of their embodied team activities (e.g., nursing simulation).

2.4 Foundations of Data Visualisation and Storytelling

Constructs from cognitive psychology and visualisation science, such as visual attention, perception, judgement, and decision-making, are critical to designing effective LA end-user interface designs (Alhadad, 2018). Combining storytelling with visualisation to generate visuals that provide context with relevant information has been of interest to Information and Visualisation (InfoVis) researchers. This section summarises some key foundations of data visualisation and storytelling, in relation to these constructs, in the form of principles.

First of all, there is a distinction to draw about the types and purposes of data visualisations. For this thesis, **exploratory analysis** is "what you do to understand the data and figure out what might be noteworthy or interesting to highlight to others" (Knaflic, 2017, Chapter 1). As Knaflic (2017) indicates, exploratory analysis is "like hunting for pearls in oysters". On the other hand, when a clear message or insight needs to be communicated, it is better to use an approach grounded in **explanatory** visualisation. In that case, explanatory visualisation can be used to communicate a specific story you want to tell (one of the pearls).

The ultimate objective of data visualisation in LA is to ease the cognitive load of educational stakeholders and facilitate the interpretation of the visual elements used to encode educational data (Alhadad, 2018). One way to support the teacher and student interpretation process is by using salient visual features to influence attentional capture **(Principle 1)**. Attention capture is the phenomenon in which attention is directed involuntarily toward a target stimulus based on the characteristics of that stimulus (Knaflic, 2017; Matzen et al., 2017). This means that pre-attentive attributes, which are properties of visualisations that people notice without using conscious effort (such as variations in colour, shape, enclosure, intensity, hue, or size), can actually be configured/designed for helping ‘end-users’ interpret the key information that is relevant to them (Barrera-Leon et al., 2020). In order to emphasise some visual elements, others need to be de-emphasised. Therefore, it is critical to avoid visual clutter as it can severely degrade attention and impair comprehension **(Principle 2)**. Reducing visual clutter, or the excess of information and lack of order, enables the ‘end-user’ to allocate more cognitive capacity for essential processing (Alhadad, 2018; Munzner, 2014).

In addition to highlighting and de-emphasising elements, the use of text to communicate insights can directly support inference-making **(Principle 3)**. Visual narrative refers to the use of visual and textual elements to convey a story (Segel and Heer, 2010). Generally, the viewer’s attention is orientated toward text (Matzen et al., 2017). This can be used for explaining the meaning of certain data points (if known) (Echeverria et al., 2018b) or to emphasise sections in a visualisation that require the viewer’s attention (e.g., using text to also achieve Principle 1 (Knaflic, 2017)). Finally, some types of visual representations inherently help or hinder the understanding of the data and subsequent inferences made **(Principle 4)**. This means that choosing the right type of visualisation representation for particular goals and educational contexts is key (Alhadad, 2018). I use these four principles to derive the designs of the MMLA interfaces described in Section 6.3.

As mentioned above, students and teachers can generally be considered non-data experts (users who have little or no data analysis experience (Schield, 2004)). Empirical research has shown that even students of Science, Technology, Engineering, Arts, and Mathematics (STEM) often face difficulties in interpreting and understanding charts (Maltese et al., 2015). Moreover, students can also be considered casual users of LA systems (Cuff, 1980) since, first, LA tools are commonly a novel form of tools for most students, and second, LA tools will typically be used sporadically to reflect on work (Tan and Koh, 2017).

Research in information visualisation (InfoVis) argues that **guidance** should be provided to support casual users, or users with low analysis experience, to interpret data visualisations (Schulz et al., 2013). Guidance can be defined as “a computer-assisted processes that aim to narrow the data interpretation and exploration gap encountered by end users" (Ceneda et al., 2017)(p. 2). Schulz et al. (2013) described different ways in which this concept can be materialised, such as by enhancing the charts using visual cues, allowing users to select from various visualisation techniques, and guiding users through prescriptive data exploration workflows or via **Data storytelling** (DS).

DS is a suite of information design and "compression" techniques to help an audience effectively understand what is important in a visualisation (Ryan, 2016), communicating key messages clearly and effectively through the combination of data, visuals, and narratives (Dykes, 2015). Ryan (2016) and Knaflic (2017) distilled the following DS principles (which are in line with the principles described above to reduce the cognitive load of educational stakeholders):

- **DS1. DS is goal-orientated**. A data story should have a very specific goal, which enables the identification of the data that should be visually emphasised. This can help designers and researchers establish clear boundaries about what needs to be communicated and what does not. Although it may sound simple, many learning data visualisations can be designed to invite students or teachers to explore the data without a clear goal in mind.

- **DS2. A data story should be based on a suitable chart type**. Some charts work better for certain purposes. For example, line charts can effectively show changes over time Ryan (2016). In contrast, Knaflic (2017) devotes a whole chapter to justifying why pie charts should not be used.

- **DS3. A data story should first be stripped down**. Decluttering is a critical step to reduce the complexity of visual representation. This involves removing headers, borders, grids, and data points that are not central to the story. Other design aspects are important; for example, Knaflic (2017) proposes the addition of titles to drive the intent of the visual, and *captions* to explain important features of the data. The general principles related to alignment, use of *colour*, *shape*, and texture are information design decisions that can have a strong impact on sensemaking.

- **DS4. A data story should guide attention**. Only data points that are critical for communicating a story should be emphasised using visual or narrative elements. This can be achieved by 1) *adding enhancements* such as arrows, lines, symbols, or enclosures; 2) changing *color*, *contrast* or *thickness*; and 3) *annotating salient data* features and *adding headline titles* that summarise the message of the story. These DS principles will be illustrated through the prototypes in Chapters 5 and 6, below.

An additional data storytelling principle important for this thesis is the use of text to communicate insights directly to support inference making **(DS5)**. I will use this DS principles later to describe the designs of the MMLA interfaces described in Chapter 5, Section 5.1 and in Chapter 7 to define the architecture.

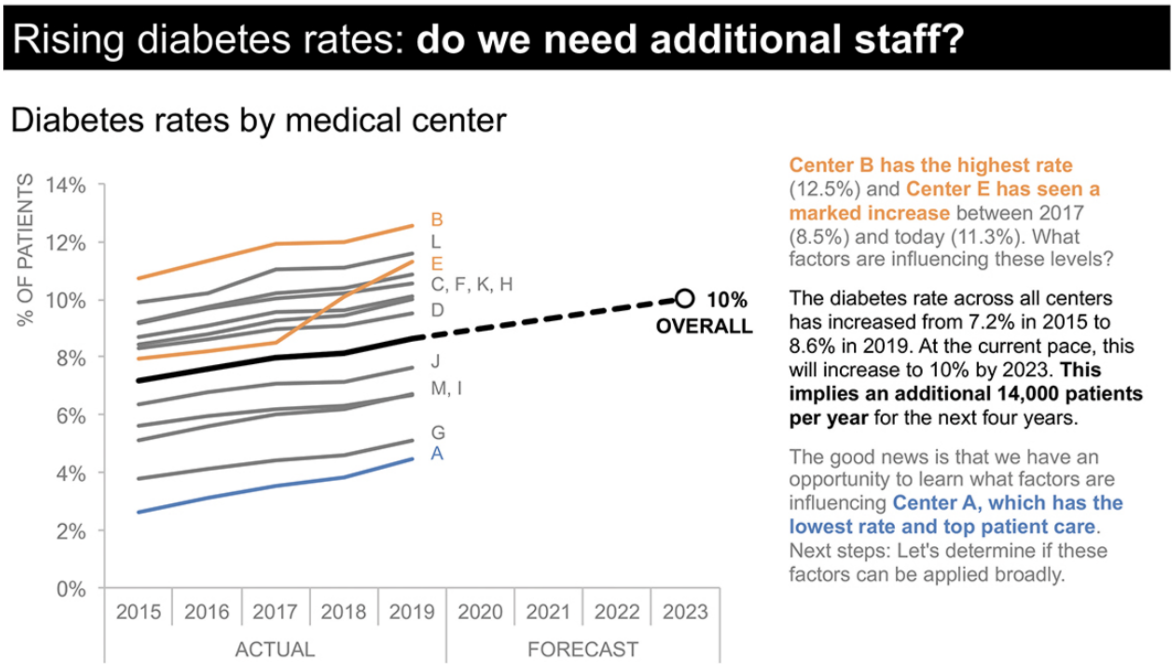


Figure 2.12: Storytelling with data example extracted from (Knaflic, 2017, Chapter 7)

Figure 2.12 illustrates the principles presented above as follows: the visualisation uses a title to explain its purpose, which is to communicate the need for additional staff (DS1); the type of visual used for this visualisation is a line chart, as it represents a time series and changes over time are important in the story (DS2); the visualisation removes unnecessary elements such as grids or not relevant data labels (DS3); the use of colour is intentional to guide attention to messages in orange and blue (DS4); and the visualisation uses text to communicate messages in a straightforward manner (DS5).

In their exploratory research, Figueiras (2014) presented different case studies on how to combine visuals, annotations, and narrative to convey stories, although their prototypes were not fully automated, they are a good reference for the strategies and benefits of adding storytelling for data visualisation purposes. Similarly, the exploration of data stories awarded by Ojo and Heravi (2018) provided a reference of techniques and the exploration of the benefits of DS to communicate insights. The next section presents examples of LA research in which some authors have started to apply the principles and techniques of infoVis to their LA tool designs.

2.4.1 Using Data Storytelling in MMLA to support student learning

Data Storytelling has only started to be recognised in LA. For example, Chen et al. (2019) proposed an approach to highlight and annotate video elements and slideshows to present visual data to teachers to understand student progress. Similarly, Echeverria et al. (2018b) showed how enhanced visualisations with DS elements can drive the focus of teachers' attention and lead to deeper reflections on the student's data. The same authors proposed that the instructional design of the teacher should drive the visualisation design (Echeverria et al., 2018a), but to date they have only reported laboratory-based trials with teachers in experimental settings.

Although this and the previous work discussed above examined the participation of teachers in the design, it is also critical to understand how students can directly benefit from multimodal data (as illustrated in Echeverria et al. (2018a,b); Martinez-Maldonado et al. (2020a)). Although there have been great advances in the field, this literature review has highlighted the need for alternative ways to evaluate how useful MMLA interfaces can be in supporting education. There is a gap in the literature regarding the lack of *empirical evaluation with educational stakeholders validating the usefulness of MMLA interfaces to support learning and teaching.* This gap motivates the third research question explored in this thesis.

RQ3: To what extent can students and teachers reflect on embodied team activity using MMLA interfaces?

This thesis will report empirical findings from five authentic qualitative studies with nursing students using seven (8) high-fidelity prototypes, presenting multimodal learning stories for them to reflect on their errors, physiological arousal, and Teamwork Proxemics during medical ward simulations.

In line with the previous gap,

This lack of research on the design and evaluation of multimodal interfaces to support educational stakeholders, rather than researchers, can be caused by **infrastructure** issues (e.g., not scaled-up solutions).

In the work of Chen et al. (2019) (slideshows, explained above), the authors defined an architecture where data processing and narrative systems are the main components. A visual interface is also provided with different types of visual enhanced with text narratives, which are generated based on a predefined question-list format about video, assignments, forum, general actions, and learner tasks. Although this approach was implemented for Massive Open Online Courses (MOOC), not for physical spaces, it is a good example of how to use storytelling tools to assist teachers in the collection and organisation of student data to produce data stories to facilitate finding learning patterns. For physical spaces, Martinez-Maldonado et al. (2020a) introduces a conceptual multimodal layered storytelling approach, whose objective is twofold: i) to categorise the underlying multimodal data (sensor data) into meaningful layers of information and ii) to apply data storytelling to drive visual attention to key events of the learning activity. Although there are some conceptualisations and initial attempts to incorporate DS improvements to support educational stakeholders, there have been no implementations of fully automated DS interfaces to support reflection on embodied team activity.

The promise that DS could help to guide students reflections motivates this research to prove if the prototypes that are becoming possible can be implemented in a scale-up way. The available infrastructure is not yet widespread, but once it is possible, more interfaces can emerge to support students and teachers. Although Martinez-Maldonado et al. (2020a) generates prototypes that were not fully automated, the approach adopted there motivates my final research question:

RQ4: To what extent can MMLA interfaces for students and teachers be automatically generated?

The conceptual approach used to address the four research questions in this thesis is fully described in Chapter 3. The *Functional Architecture* and the *Reference Implementation* to address RQ4 are described in Chapter 7.

2.5 Applied problem and gaps from literature

To this point, this chapter explained: the primary literature that provides context to this thesis; the gaps identified from the literature review; the applied problem; and the four research questions that motivate this thesis. Figure 2.13 summarises the context of this investigation.

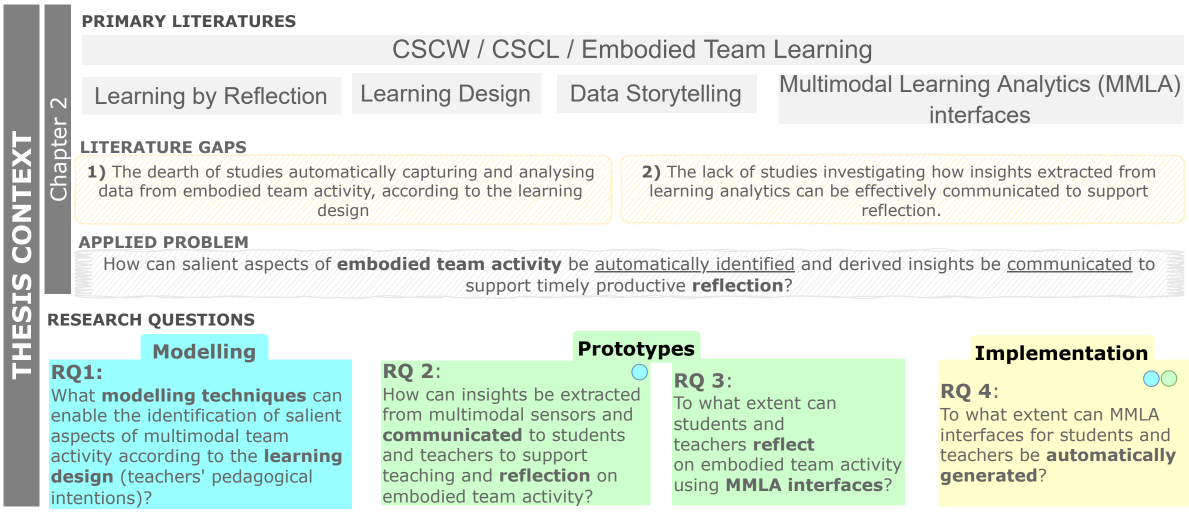


Figure 2.13: Thesis context: literature, gaps, applied problem and research questions -RQs.

The two gaps identified as a result of the literature review and evaluation are:

- A dearth of studies automatically capturing and analysing data from embodied team activity (according to teachers' pedagogical intentions)
- The lack of studies investigating how insights extracted from learning analytics can be effectively communicated to support reflection

Considering the previous gaps the applied problem that motivates this thesis research states as follows: *How can salient aspects of embodied team activity be automatically identified and derived insights be communicated to support timely productive reflection?*.

The work presented in this thesis goes beyond previous work in: i) exploring other modelling modalities and constructs to gain a deeper understanding of embodied teamwork (Fernandez-Nieto et al., 2021a,c) (chapter 4); ii) designing and evaluating new explanatory DS interfaces (Fernandez-Nieto et al., 2021b; Fernandez Nieto et al., 2022) and exploratory MMLA interfaces (Fernandez-Nieto et al., 2022, 2021c) (chapters 5 and 6); and iii) defining and explaining architectural considerations and illustrating its technological implementation in an in-the-wild setting for automatically generating MMLA interfaces for teachers and students (chapter 7).

CHAPTER 3

Conceptual Framework and Research Methodology

This chapter[1] describes the conceptual framework (see Figure 3.1) used to guide the modelling and the generation of MMLA interfaces (Section 3.1). The learning contexts that supported the evidence to address the different research questions of this thesis (Section 1.2) are presented in Section 3.2. Then, based on the conceptual framework and the learning situation, the research methodology used to develop this research is explained and justified in Section 3.3.

3.1 Conceptual Framework

Motivated by designing meaningful analytics, this thesis explored an approach to drive: analysis and modelling based on the learning design and the generation of automated MMLA interfaces. The conceptual framework guiding this thesis is inspired by previous research: Echeverria et al. (2018a)'s learning design driven conceptual model, that uses the learning intentions to define **rules** which can be read by a data processing system, their work focused on generating data story elements based on the rules and

[1]This chapter is based on: R. Martinez-Maldonado, V. Echeverria, **G. M. Fernandez-Nieto**, and S. Buckingham Shum. 2020. From Data to Insights: A Layered Storytelling Approach for Multimodal Learning Analytics. In CHI Conference on Human Factors in Computing Systems CHI'20. doi: 10.1145/3313831.3376148.

Martinez-Maldonado et al. (2020a), whose work introduces a conceptual multimodal layered storytelling approach for physical spaces (see Figure 3.1).

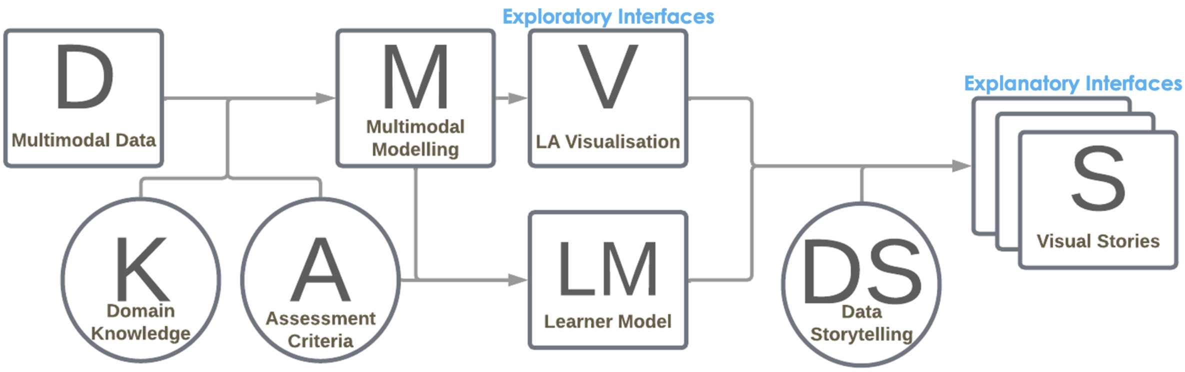


Figure 3.1: Layered Multimodal Data Storytelling Approach. Conceptual framework used to guide the modelling and design of multimodal data storytelling

The framework described in Martinez-Maldonado et al. (2020a) (see Figure 3.1) was developed in collaboration with other researchers and myself. Our approach aims are two-fold i) categorising the underlying multimodal data (sensor-data) into meaningful layers of information and ii) applying data storytelling to drive visual attention to key events of the learning activity. I used this framework to map all the modelling techniques and the generation of prototypes for this thesis. Furthermore, my research goes beyond this conceptual approach by exploring how to automatically elicit educational intentions from teachers and how to automate the entire approach. In the conceptual framework *boxes* represent artefacts such as data structures and interfaces. *Circles* represent input parameters expressed as rules or templates to generate data structures or configure visualisations. In our approach, *multimodal low-level learning activity data* **(D)** are encoded into a meaningful information structure, as part of a *multimodal modelling* process **(M)**, based on *domain knowledge* **(K)** (e.g., students' expected behaviours from the learning design). This structure can be directly rendered visible as **exploratory visualisations (V)**. However, our approach was enriched by including: *assessment criteria* **(A)**, used to generate a higher-order information structure; the *Learner Model* **(LM)**, an structured representation of student performance, misconceptions or difficulties; and *data stories* **(S)**, which are visualised in a layered **explanatory interface** using LM and operationalising DS principles (described in Section 2.4).

The following chapters will present how this approach was used to gain insights (M) from multimodal data (D) using the teachers' pedagogical intentions and the learning design (K,A) (see Chapter 4), to generate students (see Chapter 5) and teachers (see Chapter 6) MMLA interfaces (V, LM, DS, S), and an implementation reference to automate the whole approach (see Chapter 7).

3.2 The learning situation

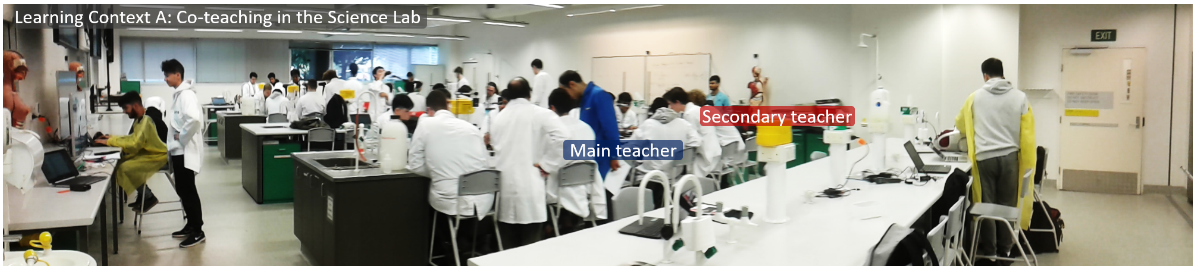


Figure 3.2: Learning contexts from which positioning data were captured. Top: a science laboratory. Bottom: a nursing simulation classroom.

This section describes the learning contexts that supported the evidence to answer the main research question of the thesis (in Section 1.2).

This thesis focused on four learning situations that evolve embodied teamwork activity, three healthcare simulations, and one co-teaching scenario in a science laboratory. These two context are of interest for this thesis because they both focused on *training professionals* during embodied team activity. Nurses train to take care of a patient in different situations while teachers work together to co-teach a a team in a science lab.

3.2.1 Healthcare Simulations

Healthcare simulation is a pedagogical approach that uses a constructivist learning model to provide students with opportunities to experience teamwork and patient situations without compromising the care of real patients (Berragan, 2011). *A simulation is an embodied activity in which nurses interact in teams in a social and physical environment* *to provide care to a patient*. During the simulations nursing students should learn bodily experiences such as how to use the physical space, how to position themselves in the space, how to interact with various physical and digital devices and how to interact with others.

Simulations often start with a description of learning goals, followed by the simulation itself, concluding with a debrief aimed at provoking students' reflection on performance and errors made. Nursing simulations supports professional training in aspects such as increasing professional self-confidence immediately after training (Fuglsang et al., 2022), gaining and learning theoretical concepts, developing critical thinking abilities (Lei et al., 2022), and improving competency and emergency response ability (Guo et al., 2022). Although errors in simulations can cause negative feelings in some students, current studies suggest that constructively addressing errors can help learning (Palominos et al., 2019). Although there are video-based products to support this reflection, they are often impractical for classroom use, resulting in students rarely using this evidence to inform reflection (Mariani and Doolen, 2016).

Various simulations are conducted as part of the curricula of any undergraduate nursing program. Simulation classrooms are equipped with 5-6 beds with a patient manikin on each. Students are commonly organised in teams of 4-5 members to take care of each patient in a hypothetical scenario.

3.2.1.1 Simulation 1: Recovery from abdominal surgery

Simulation 1 was run in 4 classes by the same teacher (the subject coordinator). A total of 19 students in their third year (all females, one team in each class) volunteered to participate in the study and for their data to be recorded. The goal of the simulation was providing care to a patient after an abdominal surgery. Students in each team played the roles of team leader, recovery nurses (RN1, RN2), scribe (RN3) and the patient not tracked (see Figure 3.2, bottom). According to the assessment criteria set by the teacher, a highly effective team should have performed the following 5 *actions*:

1. Assess vital signs every 10 minutes

2. Check fluids, suction secretions, perform head tilt/chin, or add oxygen therapy after the patient presents breathing obstruction

3. Administer Fentanyl within 10 minutes after the patient complains of abdominal pain

4. Administer a second bolus of Fentanyl after the patient complains of severe abdominal pain

5. Administer Ondansetron within 10 minutes after the patient experiences nausea.

Based on the critical actions described above, the simulation can be divided into 6 phases:

- *Phase 1:* patient assessment (from the beginning of the simulation to the moment that nurses realise that the patient presents breathing obstruction);
- *Phase 2*: IV fluid preparation and IV fluid administration (Fentanyl);
- *Phase 3:* IV fluid preparation and IV fluid administration (Fentanyl second dosis);
- *Phase 4:* patient experiences nausea (after Fentanyl administration);
- *Phase 5:* Ondansetron administration; and
- *Phase 6:* patient recovery.

3.2.1.2 Simulation 2: Allergic reaction to antibiotic

Simulation 2 was run in 5 classes taught by 3 teachers (including the same subject coordinator). A total of 25 students in their third year (21 females and 4 males) volunteered to participate. The purpose of this simulation was to help nurses learn how to react when a patient has an allergic reaction to some medication. Similar roles to those in Sim 1 were assigned to members of each team. According to the assessment criteria, a highly effective team should have performed the following 6 *actions*:

1. Perform an initial set of vital signs, after the teacher reads the initial handover
2. Administer the intravenous fluid -IV antibiotics
3. Perform another set of vital signs after the patient complains of chest tightness
4. Stop the IV antibiotic after the patient reacts to chest tightness
5. Perform an ECG after the patient complains of chest tightness
6. Call the doctor after stopping the IV antibiotic.

Based on the critical actions described above, the simulation can be divided into 5 phases.

- *Phase 1:* patient assessment (from the beginning of the simulation to the moment nurses realise that the patient needs IV antibiotic);
- *Phase 2*: IV fluid preparation;
- *Phase 3:* IV fluid administration;
- *Phase 4:* patient adverse reaction (since the patient starts complaining about the allergic reaction until the moment the nurses stop the IV antibiotic); and
- *Phase 5:* patient recovery.

According to the subject coordinator, the simulation have 3 critical phases of interest to understand students arousal levels:

- *Phase 1:* patient assessment (from the beginning of the simulation to the patient's chest tightness)
- *Phase 2*: critical patient deterioration (since the patient starts complaining about the allergic reaction until recovery)
- *Phase 3:* patient recovery (from the patient recovery to the end of the simulation)

3.2.1.3 Simulation 3: Prioritisation of beds

Simulation 3 was run in 38 classes by different instructors. A total of 254 students in their third/fourth year volunteered to participate in the study and for their data to be recorded. The goal of the simulation was to provide care to four patients and prioritise the care of each bed as a team. The students in each team played the *roles* of Graduate Nurses (GN) 1 and 2 (primary nurses) and the Ward Graduate Nurses (WN) 1 and 2 (secondary nurses) and 4 manikins played the role of patients (see Figure 3.2 bottom). According to the assessment criteria set by the subject coordinator, a highly effective team should have performed the following 6 *actions* in the main bed:

1. Administer oxygen after patient respiratory depression;
2. Assess vital signs every 5 minutes;
3. Cease PCA (patient-controlled analgesia) after patient altered conscious state;

4. Activate MET (Medical Emergency Team) calls after patient deterioration;
5. Administer Naloxone timely.
6. Students are supposed to take care of the other 3 beds; they have to prioritise care.

According to the subject coordinator, the simulation have 3 critical phases of interest:

- *Phase 1:* patient assessment (from the beginning of the simulation to the patient's respiratory depression)
- *Phase 2*: patient altered conscious state (since the patient starts feeling sleepy until ceasing the PCA)
- *Phase 3:* patient recovery (after ceasing the PCA to the end of the simulation)

3.2.2 Science Laboratory

In the educational sector, teaching guides (e.g., (Jones et al., 2007) and professional support staff or peers (Britton and Anderson, 2010) often recommend or prescribe to teachers embodied strategies such as learning how to position themselves in specific locations in the classroom (termed instructional proxemics (Chin et al., 2017)). These guides and peer feedback are important for many teachers, particularly those in higher education (HE), who *rarely receive pedagogical training* (e.g., teaching assistants, tutors, and academics) or feedback on how to position themselves or effectively approach students while delivering classes (Hagman et al., 2016). In the science laboratory, two or more teachers are jointly involved in the educational process that manages classroom behaviour. *Co-teaching is an embodied team activity that requires coordination of the instructional practice and understanding of own and others social and physical aspects in the classroom*. Teachers can benefit of training on co-teaching by learning how to distribute their roles and workload (e.g., even/uneven distribution of roles), identifying the different models of cooperation between educators (e.g., territoriality), or recognising effective patterns for co-teaching.

3.2.2.1 Laboratory 1. Co-teaching

Laboratory 1' learning context involves weekly 2 and a half-hour laboratory classes of a first-year undergraduate science unit at University of Technology Sydney, in which students run physics experiments. The *lead teacher* and a *teaching assistant* co-teach

each laboratory with typically 30-40 students working in 10-13 teams of 2-3 students each. The positioning data of both teachers was captured from 18 laboratories randomly chosen (1-18). All laboratories were conducted in the same 16.8 m x 10 m classroom (see Figure 3.2, top) equipped with workbenches, a lectern, a whiteboard, and multiple laboratory tools. A limited number of photos were taken during the classroom sessions.

The laboratory session was divided in three phases (the study was focused on phase 2):

- Phase 1 includes the main teacher explaining the goals and giving instructions to the students from the teacher's desk (average duration 11.5 +/- 8 minutes, n=12).
- Phase 2 corresponds to the period in which all students start working on the experiment(s) of the day in small teams (1.36 hours +/- 25 min).
- Phase 3 corresponds to the time when some teams complete their experiments and start leaving the classroom (32 +/- 25 minutes).

3.3 Methodology

The methodology selected to address the research objectives of this thesis is **Mixed-Methods Research**. This methodology is commonly used for research processes to integrate quantitative and qualitative methods of data collection and analysis (Plano Clark and Ivankova, 2016). Complex contexts, like education (teaching and learning) and social sciences in general, require the use of mix-method methodologies, because this methodologies brings additional perspectives for data interpretation. Mixed methods research is, generally speaking, "an approach to knowledge (theory and practice) that attempts to consider multiple viewpoints, perspectives, positions, and standpoints (always including the standpoints of qualitative and quantitative research)" (Johnson et al., 2007). The mix of methods felicitated the linkage or integration of data seeking to have a panoramic view, or having views from different perspectives through diverse research lenses. In the context of MMLA, multimodal data can also be understood using qualitative analysis, which can be associated with constructs that humans can easily interpret (Johnson et al., 2007; Onwuegbuzie and Leech, 2005). For instance, mixed methods can provide opportunities for participants to share their experiences and expertise across the research process facilitating avenues of exploration that enrich the evidence and enable questions to be answer more deeply (Shorten and Smith, 2017).

As described by Ivankova and Stick (2007) the sequence of design of a mixed method research commonly includes the following steps: (a) quantitative data collection; (b) quantitative data analysis; (c) case selection, interview development protocol; (d) qualitative data collection; (e) qualitative data analysis; and (f) integration of quantitative and qualitative results. The procedures and expected products of each phase defined for this doctoral thesis are explained below.

3.3.1 Quantitative data collection

Quantitative multimodal data **(a)**, was captured from authentic classrooms during embodied teamwork activities (health simulations and physics laboratory) using different sensors and devices. The selection of the devices depends on the dimensions (epistemic, physiological, affective, physical, and / or social) that teachers must include to provide evidence-based reflection tools. Data collection (full details are provided in the next subsections) was driven by the learning design of the healthcare simulations and the teachers' pedagogical intentions (e.g., learning objectives). This research also collected qualitative data (e.g., human observations during the activity) to complement quantitative data collection and for further analysis, as will be explained in Chapter 4. The next subsections specifies the quantitative data collected from the nursing simulations and the science laboratory.

3.3.1.1 Data collection for Healthcare simulations

The qualitative data collected during the simulations are described in Table 3.1 and illustrated in Figure 3.3. The table 3.1 presents a description of the sensor data collected, illustrative examples of the salient aspects of embodied teamwork that can be measured with them, and the pre-processing considerations used for the different studies (e.g., how the data was stored).

Students' low-level **positioning data** was captured through wearable tags[2]. Tags, carried in waist bags or armbands, were worn by students during the simulation (see indoor positioning trackers in Figure 3.3). The x and y location and the body rotation of students and the teachers were capture over the simulation.

The students' **physiological data** was captured through (Empatica e4) wristbands (see device picture in Figure 3.3). These record electrodermal activity (EDA) at 4Hz. EDA signals are composed of both skin conductance responses (SCR) and skin conductance

[2]www.pozyx.io

Figure 3.3: Quantitative data collection for healthcare simulations: positioning trackers, physiological bracelets, microphones, video, and patient manikin logs.

Sensor data	Illustrative Embodied Team Activity Metrics	Pre-processing
Indoor positioning	Students' spatial behaviours: 1) Co-presence in interactional spaces, 2) Socio-spatial formations, 3) Presence in spaces of interest (Fernandez-Nieto et al., 2021a)	Captured at 2-3Hz and then down-sampled to 1Hz, data points ware formatted as follows: *{tagId, timestamp, x, y, rotation}*
Physiological data	Stress reactions during teamwork (Noroozi et al., 2019; Ronda-Carracao et al., 2021), mental work load (Vanneste et al., 2021)	Electrodermal activity (EDA) at 4 Hz. Detect student arousal peaks (sudden changes in skin conductance level). Formatted as follows: *{timestamp, EDA, amplitude (amp), Skin Conductance Responses width}*
Log of actions	What students were doing as a team and individually (Fernandez-Nieto et al., 2021b), assessing teamwork (Schraagen et al., 2010), observe social processes (Cross and Clayburn Cross, 1995)	Observations were captured and formatted as follows: *{role, timestamp, action description}*
Audio	Effects of communication on performance (Fletcher and Major, 2017), collaboration detection (e.g., talk vs silence) (Martinez-Maldonado et al., 2013), total speaking time (Bachour et al., 2010), detect collaboration levels (Lubold and Pon-Barry, 2014) (intensity and pitch of sound), predict collaboration quality (Bassiou et al., 2016) (e.g., speech overlap duration)	Log of starting and stopping time of video recording as follows: *{start timestamp, stop timestamp}* Audio files were stored indicating the simulation session id and the role who wore the device: *{sessionId, roleId}*
Video	Detect indicators of collaboration (Davidsen and Ryberg, 2017; Scherr and Hammer, 2009) using gesture, posture, gaze, eye contact, or other behaviours	Video files were stored indicating the simulation session id.

Table 3.1: Sensor data collected and illustrative examples of its importance for embodied team activity

levels (SCL). The analysis of such components can be used to signal student-measures of stress (Noroozi et al., 2019) or cognitive load (Vanneste et al., 2021).

Some students **actions**, such as assessment of vital signs, were detected by the mid-fidelity manikin (Laerdal Nursing Anne). Other actions performed by each student (e.g., stopping intravenous-IV fluid, writing on charts, and calling the doctor) were manually logged by an observer (a researcher), but it could also be a student using a web application, or a higher-fidelity patient manikin not available in all the classrooms of the hosting universities. Finally, **video** and **audio** data were collected using in-house web applications with third-party open-source libraries (see video camera in Figure 3.3). The audio data are captured with lapel microphone (see Figure 3.3) or wireless microphones and an audio interface (e.g., TASCAM US-16x08[3]).

3.3.1.2 Data collection for the Science Laboratory

Figure 3.4 illustrates the devices used for the data collection in the science laboratory while teachers were co-teaching. For this learning context, only the *x*-*y* coordinates and body orientation data of the two teachers in each Science Laboratory (Context A) were automatically recorded. The sensors used and the way the devices were worn are similar to what was described above. Eight 'anchors' were installed on the classroom walls to triangulate the (x and y) positions of the badges.

Figure 3.4: Quantitative data collection for the science laboratory: positioning trackers

[3]https://tascam.com/us/product/us-16x08/top

3.3.2 Summary of multimodal modelling techniques (quantitative data analysis)

The objective of this phase **(b)** is to *translate raw data into educationally meaningful insights*. However, mapping the huge amounts of data that sensors can generate to *meaningful constructs* related to embodied teamwork (e.g., effective communication, coordination, and leadership) remains a key challenge. Although logs can illuminate what users *do*, they often say much less about *why* (Dumais et al., 2014) and *how*. Modelling techniques that incorporate context as part of the analysis help *provide meaning* to the data and inform the analytics. This thesis addressed this challenge of enriching quantitative data streams with the qualitative insights needed to make sense of them by adopting a **Quantitative Ethnography** (QE) approach (Shaffer, 2017). QE was used to identify the salient aspects of the specific embodied teamwork activity.

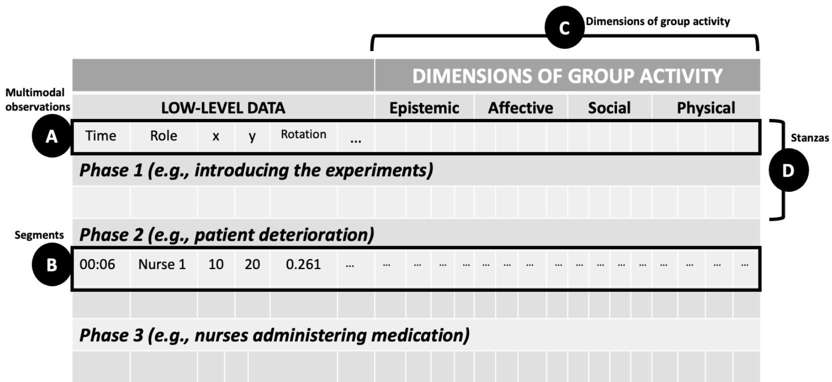


Figure 3.5: Multimodal matrix used

To apply the QE approach for the quantitative analysis the Multimodal Matrix (MM) Buckingham Shum et al. (2019a) (introduced in section 2.3.2), is generated to provide context to multimodal data. The MM is a data structure (m by n) in which each data modality m is coded into n columns of the matrix that are called *multimodal observations*. **Columns** in the matrix represents the multimodal observations (literal A in Figure 3.5), for example, observation of the critical actions or the count of nurses arousal peaks. *Segments* (m rows) are the smallest units of meaning considered for analysis and contain instances of team behaviours (literal B in Figure 3.5), in this thesis every row represent 1 second of the embodied team activity. Columns can be group into **dimension of group activity** (literal C in Figure 3.5), which can be epistemic, affective, physical, or social (as explained in Section 2.2.1). Finally, the critical actions or phases of the learning context

Multimodal data	Modelling technique	Section
Actions	Identification of students' sequence, timeliness, and frequency errors and asserts	Section 4.2.1.1
Physiological data (EDA data) and actions	Classification of students' arousal peaks to stress labels	section 4.2.1.2
Indoor positioning data and actions	Teamwork Proxemics: Co-presence in interactional spaces	section 4.2.2.2
Indoor positioning data and actions	Teamwork Proxemics: Socio-spatial formations	section 4.2.2.3
Indoor positioning data and actions	Teamwork Proxemics: presence in Spaces of Interest	section 4.2.2.4

Table 3.2: Multimodal modelling techniques explored as part of this research work

can help to define moments of interest to group segments into **Stanzas** (literal D in Figure 3.5).

Multimodal observations and segments vary depending on the multimodal data. For example, for actions (epistemic dimension) the absence or presence of certain key actions is represented as binary flags (1, 0). For the physiological data the number of arousal peaks is identified per minute, counted and recorded in the matrix for each team member. The modelling techniques used to contextualise actions, indoor positioning data, and physiological data will be described in more detail in Chapter 4. Table 3.2 summarises the multimodal data and modelling techniques used in this thesis and reference the section where the modeling technique is explained in more detail.

The analysis or modelling techniques used to gain insights from the multimodal data are aligned to the feedback that the teachers require to provide to support the embodied activity (simulation or laboratory). The analysis is normally guided by teachers' pedagogical intentions (see Section 4.2.2.1) and theory (e.g., theory of Proxemics).

3.3.3 Selection of in-the-wild qualitative studies

Five in-the-wild qualitative studies (step **c**, select cases for the mixed-method approach) were designed and conducted as part of this thesis to address the research questions presented in Section 1.2. Only one of the studies considered the co-teaching context (study 3, described in Section 6.1) the rest of them were focused on nursing simulations. Figure 3.6 presents the list of studies, the first 2 focused on explanatory prototypes and the other 3 on explanatory prototypes (see details about purposes of data visualisation in Section 2.4). In four out of the five studies either teachers or students evaluated the prototypes. The full description of the studies is presented in Chapters 5 (evaluation

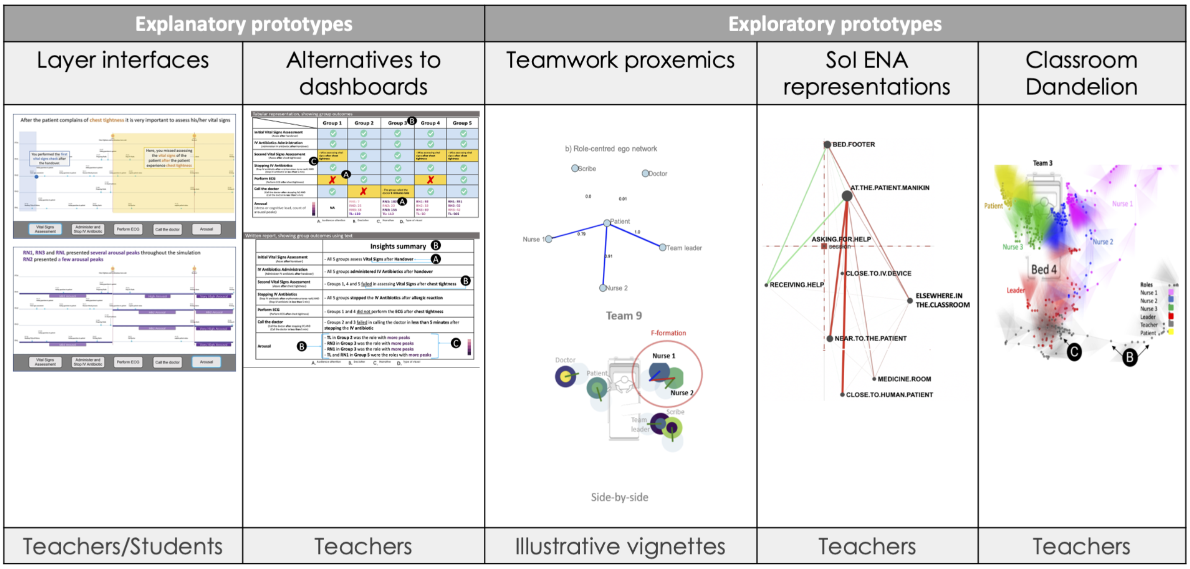


Figure 3.6: In-the-wild qualitative studies overview

with students and vignettes) and 6 (evaluation with teachers).

Please use table 3.3 as a reference to navigate thought the 5 studies in this document. An additional study corresponding to a **proof of concept** is explained in Chapter 7. For the proof of concept prototypes 1 and 2 were used to validate how feasible they are to be implemented.

3.3.4 Qualitative data Collection

All qualitative studies (step **d** of the mix-method methodology) were conducted using LATEP (Learning Analytics Translucence Elicitation Process), an elicitation protocol to understand how non-data experts envisage the use of LA systems (Martinez-Maldonado et al., ress). An illustrative example of the protocol used for the qualitative study 1 is presented in the Appendix section of this document A. Reflection sessions (think aloud and focused group) were recorded on audio tape, fully transcribed and coded using NVivo. Five empirical studies (explained above) were conducted; the protocol used for each evaluation changed according to the objectives that the empirical study was aimed to address (see Chapters 5 and 6 for additional details).

Study	Learning Context	Participants	MMLA Interfaces	Purpose
1. Layered interfaces	Simulation 1 (Section 3.2.1.1) Simulation 2 (Section 3.2.1.2)	Students	- Students facing errors and physiological data (Section 5.1)	Explanatory
2. Teamwork Proxemics	Simulation 2 (Section 3.2.1.2)	Na. Illustrative Vignettes	Teamwork Proxemics Vignettes (Section 5.2): - Interactional space between nurses and the patient (Section 5.2.1.1) - Positioning of the team leader (Section 5.2.1.2) - F-formations while stopping IV-antibiotics (Section 5.2.2.1) - How different teams used the spaces of interest (Section 5.2.3.1)	Exploratory
3. Classroom Dandelion	Simulation 2 (Section 3.2.1.2) Laboratory 1 (Section 3.2.2.1)	Teachers	Classroom Dandelions (Section 5.2): - Co-teaching (Section 6.1.4) - Teachers mentoring or intervening (Section 6.1.5.1) - Spatial team dynamics 6.1.5.2	Exploratory
4. Spaces of Interest (SoI) ENA representations	Simulation 2 (Section 3.2.1.2)	Teachers	Teachers' facing students spatial abilities (Section 6.3): - ENA representations (Section 6.2.4)	Exploratory
5. Alternatives to dashboards	Simulation 2 (Section 3.2.1.2)	Teachers	Teachers' facing MMLA visualisation narrative interfaces (Section 6.2): - Visual Data Slices (Section 6.3.4.1) - Tabular visualisation (Section 6.3.4.2) - Written report (Section 6.3.4.3)	Explanatory

Table 3.3: Summary of qualitative studies and the proof of concept

3.3.5 Qualitative data Analysis

Following best practices of qualitative research analysis (step (e) of the methodology) ((McDonald et al., 2019) p. 13), and given the direct alignment between the study protocol and the analysis themes, statements of interest were coded (Braun and Clarke, 2006) by two or more researchers according to the pre-set themes of the study protocol:

1. Teachers or students *perceptions* on different MMLA interfaces while exploring them.

2. Open questions regarding the *potential uses* of each MMLA interfaces to support their reflection on embodied teamwork activities to improve their teaching and learning practices.

The resulting coded statements were examined by the researchers (one of them the author of this thesis), who had several discussions to select instances that illustrate the opportunities and concerns raised by the teachers and students.

3.3.6 Integration of qualitative and quantitative results

Finally, to integrate the qualitative and quantitative methods (step **(f)** of the methodology) interpretations and explanations of the results are presented at the end of each study (see Chapters 5 and 6 for additional details on the results). General results of this thesis are presented as discussions, implications and future research (see chapter 8). Based on the evidence collected from the empirical studies, claims about the research are generated and described in chapter 8. The conclusions are drew base on that claims as contributions of this research. This process is described in the Aktinson and Hammersley (1998)' framework.

CHAPTER 4

EMBEDDING TEACHERS' PEDAGOGICAL INTENTIONS IN A MULTIMODAL MODEL

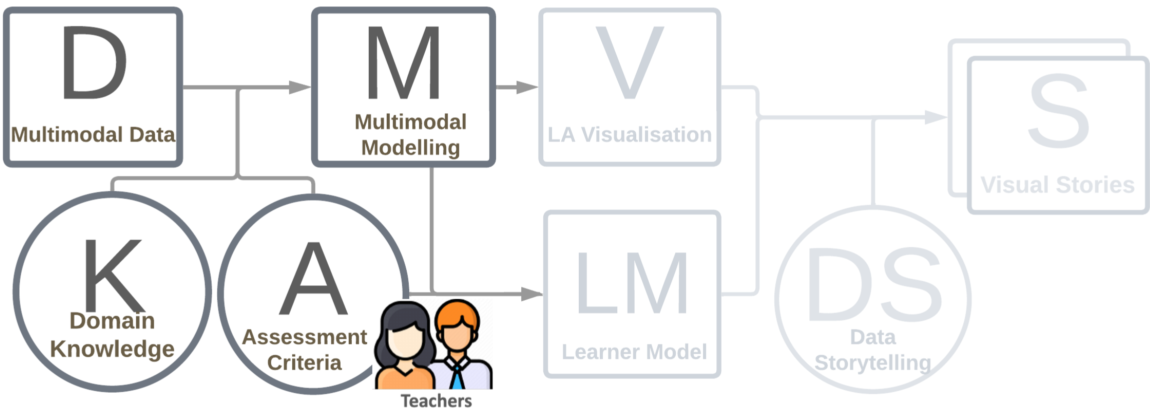


Figure 4.1: Mapping from low-level data to the assessment criteria or teachers' pedagogical intentions

The first and second contributions of this thesis are described in detail in this chapter. This chapter focused on the first component (Figure 4.1) of the conceptual model introduced in Chapter 3 (see Section 3.1). Thus, the chapter presents the **multimodal modelling techniques** *(M)* that were used to assign meaning to the **multimodal data** *(D)* collected from the physical class settings. This was achieved by applying a model to the data based on the **learning design** *(K)* of the scenario and the expectations of the teachers about student performance, represented by the **assessment criteria** *(A)* they set. Section 4.1, introduces the methodology for mapping from low-level data to assessment criteria or teachers' pedagogical intentions. Section 4.2 describes the modelling techniques used to gain meaning from indoor positioning data, physiological data, and epistemic (action) data. Section 4.2.2.1 introduces a methodology and approach to elicit nurses spatial behaviours expected by teachers.

4.1 Embedding teachers' assessment criteria and pedagogical intentions in the multimodal model

Using the Quantitative Ethnography (QE) (Shaffer, 2017) approach presented in Section 2.3.2 the multimodal data collected was modelled to map low-level data to the assessment criteria or teachers' pedagogical intentions. Thus, QE was used to gain meaning of multimodal data about the three dimensions of embodied team learning: epistemic (Section 2.2.1.1), affective (Section 2.2.1.3), and physical (Section 2.2.1.4).

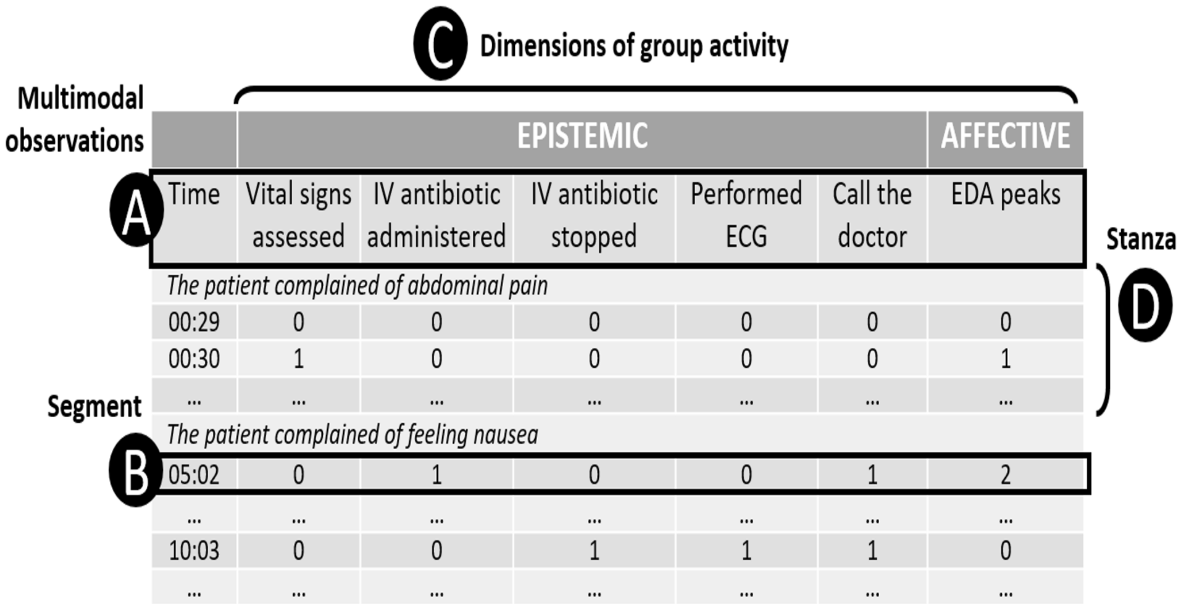


Figure 4.2: Application of the Multimodal Matrix modelling technique on the data of one team member. (A) Multimodal observations, (B) Segments, (C) Dimensions of group activity, and (D) Stanza

Different instances of the *Multimodal Matrix (MM)* proposed by Buckingham Shum et al. (2019a), will be presented in this chapter. For sensor data, each row can represent a time window (e.g., one second) of the team activity. This way, the content of each cell expresses an attribute of a given team member at that moment. Figure 4.2 is and example of a MM representing epistemic and affective dimensions of group activity. The MM shows a simplified representation of the modelling performed on the data of one team. Figure 4.2A shows the critical actions (e.g., assessment of vital signs and administration of IV antibiotic) and physiological arousal of a team member represented as columns in the matrix. They correspond to multimodal observations of the epistemic and the affective dimensions. For time series data, such as physiological data, each row (segments) can represent a time window (e.g., one second in our studies) of team activity (Figure 4.2B). Finally, for this study segments were grouped into stanzas to represent meaningful associations using phases or milestones in the learning contexts, Figure 4.2D.

This MM representation enables automated coding of low-level quantitative data into qualitative behavioural markers that can be grounded in generic features of teamwork or other constructs and the specifics of the activity. For this research, the exploration of modelling techniques was focused on the epistemic, affective, and physical dimensions of teamwork (described in Section 2.2.1). The physical dimension was one of my main interests during the investigation, that is why most of the techniques are focused on making sense of indoor positioning data.

Regarding social aspects of group activity, this thesis explored nurses' physical social presence (develop in Section 4.2.2.2), nurses' social spatial arrangements (developed in Section 4.2.2.3), and nurses' scripted and emerged roles (developed in Section 4.2.1).

The reference framework (Figure 4.1), was designed to implement QE approaches. Thus, it facilitates the consideration of qualitative aspects to enrich the multimodal data. For the modelling techniques presented in the next sections, I used the learning design *(K)* or the assessment criteria or the teachers' pedagogical intentions *(A)*.

In order to contextualise the low-level multimodal data and drive the multimodal modelling, this thesis incorporates: i) **learning design** (e.g., phases, expected actions from effective teams), ii) **teachers' expectations** or **pedagogical intentions**, and iii) modelling of **high-order constructs** (HOC) (e.g., patient care or co-teaching). In a higher-level, a HOC is defined in terms of one or many sub-constructs; a sub-construct is described as one or many behavioural markers; and one behavioural marker is determined by one or many data/analytics. The multimodal modelling techniques explored as part of this research are explained in the following sections.

The purpose of contextualising the modelling is to enrich the analysis from a qualitative perspective to generate meaningful outcomes, which are closely related to the embodied team activity. This can be possible by using the learning design (e.g., healthcare simulations 3.2.1) created by the unit coordinator or by eliciting teachers expectations of students. This research, follows a **Human-centred Design** (HCD) approach. Such approaches (e.g., context design), help researchers and designers to observe people in a natural context (often a work context), to discuss their observations in a multi-disciplinary product development team setting, and to translate these observations into specifications for an improved or new product or service (often an Information Communication Technology application) (Holtzblatt, 2007). Deeper understanding of the context can be done by observations, exploration, and design research (Tschimmel, 2012) with stakeholders. This research engages research in Human-centred Learning Analytics (HCLA), as it reinforces the need for human inference and deep awareness of the learning context, ensuring that teachers' and students' voices are taken into consideration during the modelling and the design process. That way, the analytical outcomes are relevant and meaningful for the students to reflect upon them and for teachers to support their practice.

The following sections explored the modelling techniques used to map epistemic (actions) data into errors (Section 4.2.1.1), physiological data (EDA peaks) into arousal levels respectively (Section 4.2.1.2), and indoor positioning data into teamwork proxemics (Sections 4.2.2.2, 4.2.2.3, and 4.2.2.4).

4.2 Multimodal modelling techniques

The following sections present the specific techniques used to model the epistemic, physical, and, physiological dimensions of team activity [1].

4.2.1 Epistemic and physiological dimensions

Figure 4.2 illustrates the Multimodal Matrix (see Section 2.3.2) used to model epistemic and affective data collected from Simulations 1 and 2 3.2.1.1 and 3.2.1.2. For this modelling technique, I used both data modalities (activity logs and physiological) collected from Simulations 1 3.2.1.1 and 2 3.2.1.2. The following sections provide additional details about each modelling technique.

[1] Content based on the work presented in: **Gloria Milena Fernandez-Nieto** et al., "Storytelling With Learner Data: Guiding Student Reflection on Multimodal Team Data", in IEEE Transactions on Learning Technologies, doi: 10.1109/TLT.2021.3131842

Type of Rule	Purpose	Example	Data modelled
Sequence	Gave feedback base on sequence of actions	Provide oxygen **after** the patient respiratory depression	Activity logs
Timeliness	Gave feedback base on timeliness of actions	Stop the IV device in less than **5 minutes**	Activity logs
Frequency	Gave feedback based on frequency of actions	Validate vital signs **every** 5 minutes	Activity logs
Arousal Peak Level	Categorise the arousal experienced by a student as very low, low, mild, high or very high)	Team leader presented **very high arousal peaks** after the patient deterioration	Activity logs and physiological

Table 4.1: Types of rule-based algorithms used to model multimodal data

4.2.1.1 Activity logs modelling - from actions to errors

The five expected actions described in Simulation 1 3.2.1.1 and the six described for Simulation 2 3.2.1.2 were observed as indicated in Section 3.3.1.1.

The absence or presence of certain key expected actions described in Section 3.2.1.2 is represented as binary flags (1, 0). Thus, one register in the MM log (a segment in the MM) will indicate the time (a particular second during the sim) and whether a particular critical action was performed (1) and not (0). Based on these and the assessment criteria of the learning task, I created a rule-based algorithms to automatically detect errors in the order or timeliness of student actions in the simulation task. Three types of errors were automatically identified (see first three rows in Table 4.1). The *sequence* errors are flagged if the group performed a critical action using the wrong sequence. For example, if students forget to perform a *vital signs assessment* after the patient has complained of serious chest pain. *Timeliness* errors are identified when students reacted slow and performed certain actions too late according to healthcare guidelines. For example, this happens if they take too long before calling the doctor after a patient's crisis or if they take too long to stop a medication that is causing an adverse reaction, which should be done in less than 5 minutes in this simulation. An error related to *frequency* is assessed by calculating the timestamp difference between two key logged actions that are meant to be repeated, for example: assessing patient's vital signs at least every 10 minutes.

4.2.1.2 Physiological data modelling - from electrodermal time series to arousal levels

This modelling technique considered the assessment criteria and phases specified for Simulations 1 3.2.1.1 and 2 3.2.1.2.

The recorded EDA at 4Hz, was passed through the EDA Explorer [2] algorithm to detect **non-baseline-based arousal peaks** (sudden changes in skin conductance levels) in each student's EDA signals (Taylor et al., 2015). The classifier implemented by Taylor et al. (2015) includes features such as amplitude and first and second derivative of the EDA signal. Particularly for Simulation 2, the number of arousal peaks per minute are counted and recorded in a matrix for each team member. Then, a switch algorithm was developed to assess the *arousal peak level* (Table 4.1, row 4). This categorises nurses' levels of arousal during each of the 3 phases by calculating the ratio of arousal peaks for each nurse compared to the highest ratio of arousal peaks experienced by a single student that we have detected in all of our nursing simulation studies, including previous work (5 peaks / minute) (Fernandez-Nieto et al., 2021a). This maximum ratio is divided into quintiles of equal size which are used to categorise the arousal experienced into very low, low, mild, high, or very high levels. Thus, for instance, teachers can know if the student playing the role *RN1* presented *high arousal peaks* during *phase 2*. Figure 4.2 shows how the MM was generated to identify the total arousal peaks for Simulation 2. Using the notion of *stanza* (Figure 4.2 D), the count of arousal peaks can be aggregated according to the simulation phases.

4.2.2 Physical dimension - localisation

The modelling techniques presented in this section use teachers' pedagogical intentions to gain a better understanding of indoor positioning data. The methodology used to elicit teachers' pedagogical intentions is documented and illustrated with three modelling techniques. In addition, the modelling techniques illustrated in this section considered Simulation 2 (Section 3.2.1.2). [3]

[2] https://github.com/MITMediaLabAffectiveComputing/eda-explorer

[3] This section content is based on: **Gloria Milena Fernandez-Nieto**, Roberto Martinez-Maldonado, Vanessa Echeverria, Kirsty Kitto, Pengcheng An,and Simon Buckingham Shum. 2021. What Can Analytics for Teamwork Proxemics Reveal About Positioning Dynamics in Clinical Simulations? Proc. ACM Hum. Comput. Interact.5, CSCW1, Article 185 (April 2021),24 pages. doi: 10.1145/3449284

4.2.2.1 Mapping from low level positioning data to students behaviours expected by teachers

In order to design meaningful analytics, it is necessary to understand the range of team behaviours that are expected in a learning scenario. This is commonly captured by the pedagogical intentions of teachers designing a simulation and, importantly, by the criteria they intend to assess. **Pedagogical intentions**, can be often described as teachers' purpose, aims, and objectives, and they are prominent when planning for an embodied learning activity (Nilsen, 2021).

Student behaviours expected by teachers

To illustrate how to embed the students' behaviours expected by teachers in the modelling, the *students' spatial behaviours* for Simulation 2 (see Section 3.2.1.2) were elicited from teachers. The five teachers (male = 1, female = 4, average years of teaching = 12.6), who had previously taught the simulation, were interviewed to identify the spatial behaviours they expected from the students in each phase. Each interview was recorded using an online video conference platform (e.g., Zoom) and had an approximate duration of 60 minutes. The interview followed a semi-structured format:

1. The purpose of the session was explained to the teachers
2. Teachers were presented with the phases of the simulation, using the visuals representation in Figure 4.3 and according to the learning design (see Section 3.2.1.2), and
3. Teachers were asked to respond to the following question for each phase of the simulation: What spatial behaviours would you expect students or certain roles to exhibit in phase X (X, ranging from 1 to 5), if any?

The fully recorded interviews were grouped and categorised teachers' responses to identify the expected behaviours in relation to each phase using the NVivo qualitative data analysis software tool. This resulted in a set of descriptions of the *expected spatial behaviours* of students for this specific simulation that were discussed by the rest of the research team (5 researchers). The team found consistent descriptions of expected behaviours across teachers. Then, I conducted an inductive thematic analysis (Saldaña, 2015) across phases by searching emerging categories related to the types of spatial behaviours that could be tracked using positioning technologies. The following three categories emerged:

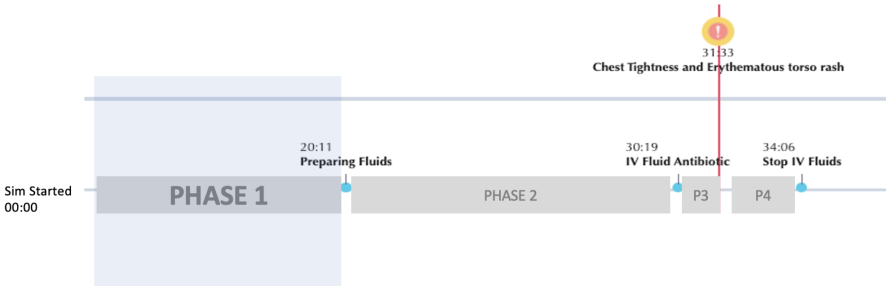


Figure 4.3: Visual representation of phases in Simulation 2 used for the interviews

i *Expected interactions* between specific roles or with the patient

ii *Expected social arrangements* during the simulation

iii *Expected spaces* where a particular nurse should position herself/himself. The resulting categories were discussed with the rest of the co-authors and were mapped to proxemics lexicon (Ciolek, 1983).

Table 4.2 summarises the spatial behaviours that were expected for each phase, and was extracted from the teachers' quotes. For example, in phase 1, all nurses are expected to *gather around* the *team leader* to plan their activity (e.g., of category *i* - expected interactions between specific roles), usually at the *bed footer* where the documentation is commonly located (*iii* - expected spaces of interest). In phases 2, 3, and 4, at least two nurses should be in close proximity to each other while they prepare (face to face or side-by-side: *ii* - expected social arrangements), provide and stop (side-by-side for one nurse to monitor what the other is doing) the IV-antibiotic, as per current national guidelines (Australian Nursing & Midwifery Federation, 2020).

Although this section illustrates the elicitation process of students' spatial behaviours, other expected behaviours can be elicited using the same method for other modalities (e.g., audio). For instance, Ronda-Carracao et al. (2021) guided their analysis of *physiological data* using the methodology described above. For that paper, we elicited nurses' *expected physiological behaviours* in terms of stress levels and cognitive load [4]. Table 4.3 presents an example of the expected behaviours of the nurses elicited from the teachers with

[4]Based on the paper: Ronda-Carracao, Miguel A., Olga C. Santos, **Gloria Fernandez Nieto**, and Roberto Martínez Maldonado. "Towards Exploring Stress Reactions in Teamwork using Multimodal Physiological Data." In MAIED@ AIED, pp. 49-60. 2021.

Phase 1	Phase 2	Phase 3	Phase 4	Phase 5
- Nurses are expected to be around the patient most of the time. (i) - Nurses should be together, around the team leader. (i) - Nurses are expected to be close to the patient performing the initial assessment. (iii)	- The antibiotic preparation should be performed by two nurses. (ii) - Fewer people should be around the patient. (i) - Some nurses should be at the medicine room retrieving the antibiotic and IV equipment. (iii)	- The patient should not be alone during this phase. (i) - Nurses are expected to be near to the patient validating that the intubation is working properly. (iii)	- One nurse must be assessing vital signs, other doing the ECG, other calling the doctor, and one nurse with the patient. (iii) - The antibiotic stopping procedure should be performed by two nurses. (ii)	- Nurses should be together, around the team leader, writing a patient report. (i)
- At least one nurse should be assessing the patient vital signs every 10 minutes during the simulation. (ii) - The scribe can be next to the patient, or at the footer of the bed through the simulation. (iii)				

Table 4.2: Teachers' expectations regarding the nurses' spatial behaviours in each phase, categorised as follows: *i*) expected *interactions* between specific roles or with the patient; *ii*) expected *social arrangements* during the simulation; and *iii*) expected *spaces* where a particular nurse should position herself/himself.

	Phase 1	**Phase 2**	**Phase 3**	**Phase 4**	**Phase 5**
Stress Potential triggers of stress/arousal	Not much stress. It can variate, it depends on the patient too (e.g.how willing is the patient to answer questions or to allow nurses to approach him/her).	Trying to figure out which medication does the patient need. Trying to figure out what antibiotic to use and the appropriate dose.	Probably if they have not' given antibiotics, or using the equipment correctly, they should be a bit nervous. Administering the IV-antibiotic is a technical skill. It could be the first/second/third time for them practicing. For this course, they should have practiced before. Some of them are confident, some not.	Critical trigger. The allergic reaction can be very surprising for nurses. Majority of stress peaks should happen in phase 4, due to changing conditions of the patient. They should be aroused all the time during this phase.	Depends on experience. Less stress because at this point the critical moment had happened.
Cognitive Load What can make team members experiment with cognitive load?	Reading through the notes. Validating compatibility of medicine. Validating dosage.	Working out how antibiotics work together. Reading through the notes. Validating compatibility of medicine and dosage.	Working out how antibiotics work together. Validating all equipment is adequate. Administering the IV antibiotic.	Put everything together to identify what was causing the situation. Figuring out what is going on with the patient. Coordinating the team.	Writing reports

Table 4.3: Nurses' Expected behaviours elicited from teachers - Physiological data

respect to physiological data. These expected behaviours are based on Simulation 2 3.2.1.2.

Spaces of Interest

As part of the interview to elicit nurses' spatial behaviours, teachers were asked to validate the *meaningful physical spaces* that were identified based on the characteristics of Simulation 2 (see 3.2.1.2) and the learning design. Then, these were mapped as a two-dimensional area using the coordinate system of the positioning tags. Figure 4.4, illustrates the meaningful physical spaces identified and validated with teachers. For example, being in close proximity to the *bed* or the nurse enacting the role of the *patient* could indicate interaction with the patient. Furthermore, being close to the *IV device* could indicate that nurses are starting or stopping the IV fluids, which is critical for the simulation. Also, as expressed by teachers, the *bed footer* is where the nurses are commonly interacting with each other to analyse the current situation, looking for more information, or preparing the medicine brought from the medicine room. In addition, these physical spaces could be linked to the actions of a specific role, such as the team leader being *next to the bed footer* monitoring the team members activity. The physical spaces which teachers recognised as meaningful for Simulation 2 will be fully described below in Section 4.2.2.4.

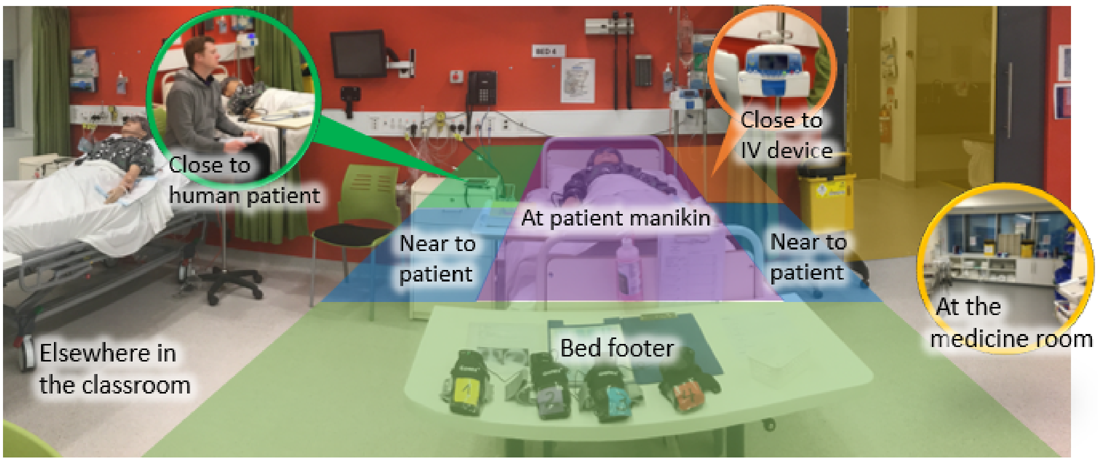


Figure 4.4: Meaningful physical spaces of interest according to teachers who conducted the simulation.

The next section describes the approach to model student spatial behaviours, by transforming low-level positioning data into meaningful proxemic constructs based on the three categories that emerged from the study presented in this section.

Modelling Approach

Figure 4.5 depicts my modelling approach for mapping from (1) low-level positioning data (described in the previous section) to (2) higher-order proxemics constructs, with the purpose of (3) addressing meaningful questions that an teacher may formulate to engage with students in dialogical feedback, or (4) questions that could be interesting for researchers to analyse teamwork activity in general. I use the notion of the *emic* perspective [32] to refer to the kinds of questions that can be addressed with positioning data from the insider's perspective, who would raise questions according to the culture of the healthcare education context. For example, these are questions that nursing teachers may have about students' activity in order to provide informed feedback, such as the categories that emerged from the study presented the previous section. The positioning behaviours that teachers expected for this simulation (presented above in Table 4.2) can be formulated as a very specific question, such as: Were at least two nurses side-by-side while stopping the IV-antibiotic? In contrast, the *etic* perspective reflects outsider questions (e.g., CSCW researchers and team scientists) studying positioning dynamics, using different, more analytical constructs.

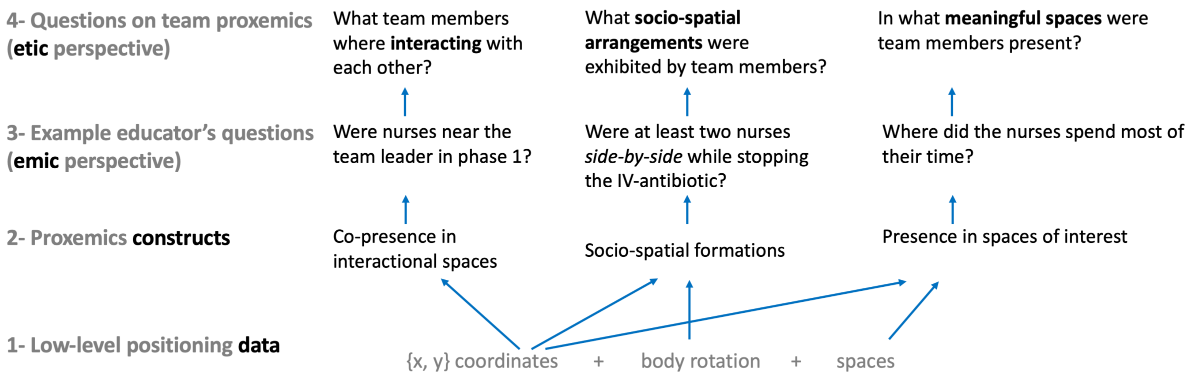


Figure 4.5: An approach to model from (1) nursing students' low-level positioning data (coordinates, body rotation and spaces of interest) to (2) higher-order proxemics constructs aiming at (3) addressing contextual questions by teachers or more (4) general questions about teamwork activity.

One of the **contributions** of this thesis is the *methodology* to elicit students' behaviours expected by teachers (in Section 4.2.2.1), and this *modelling approach*, to map indoor positioning data to higher-order proxemic constructs to address general and specific questions about teamwork activity.

Informed by the expected spatial behaviours that were derived from the interviews with teachers (in Section 4.2.2.1), the emic language used by teachers, I specified three **proxemic constructs** that can be modelled using low-level positioning data (informed by the etic perspective): i) *co-presence in interactional spaces,* ii) *socio-spatial formations,* and ii) *presence in spaces of interest*. In this section, I will ground their definitions in the literature, and detail the data used to model each of them (for Simulation 2 in Section 3.2.1.2).

4.2.2.2 Proxemic construct: co-presence in interactional spaces

The concept of *interactional space* was defined by Mondada (2013), and refers to the dynamic use of the physical space which enables verbal interactions between people. The interactional space is constituted by people mutually adjusting the arrangements of their bodies to enable close proximity and mutual attention to each other, and the objects they manipulate. This is aligned to Hall's classic work (Hall, 1966a) which outlined four types of distances, each commonly used by people for a certain type of interaction. These are: i) *intimate* (0-0.46m), where the presence of the other person is unmistakable and can be overwhelming (Ciolek, 1983); ii) *personal* (0.46-1.2m), where the majority of intensive and delicate interpersonal transactions occurs (Ciolek, 1983); iii) *social* (1.2-3.7m), where verbal transactions can occur, but it is generally considered as a distance from which strangers commonly interact Sorokowska et al. (2017); and iv) *public* (3.7+), where the other person's presence is not well-defined and it can be either acknowledged or ignored (Ciolek, 1983). Although these exact distances vary across cultures, most interpersonal interactions with acquaintances tend to occur under 1.5m (Sorokowska et al., 2017).

These types of distances are critical to identify situations in which team members may be interacting with each other. Based on classic proxemics work (Hall, 1966a; Martinec, 2001) and empirical work in healthcare (Melo et al., 2013; Moreira et al., 2017), the co-presence of two nurses within their intimate or personal spaces can be indicative of some verbal interaction or awareness of each others' actions. Similarly, if a nurse is close to the patient it can be indicative of nurse-patient interaction or care giving.

The construct of co-presence in interactional spaces was thus modelled by measuring the distance between the x, y coordinates among each team member and the patient (either the manikin or the nursing student playing the role of the patient), per second. We identified instances of close proximity (intimate or personal distances) using the parameter $d = 1.2m$ which, according to Martinec (2001), is an appropriated distance to enable direct interaction and also accounts for the (commonly 91cm) width of the patient bed (according to current standards (Wiggermann et al., 2017)) that may be between two nurses. This parameter can be adjusted depending on the team context or the simulation.

With the calculated distances, the interpersonal proxemics were indicated using the MM representation. In Figure 4.6, I indicated the interpersonal distance between the team members (RN1, RN2, and RN3), the teachers (TCH), and the patient (PTN) (roles in Figure 4.6). For this modelling technique, the phases of Simulation 2 (see Section 3.2.1.2) can serve to group segments into stanzas. After normalising the data, the aggregation of phases indicates the percentages of time that nurses spent in close proximity (intimate and personal).

LOW-LEVEL POSITIONING DATA						CO-PRESENCE IN INTERPERSONAL SPACES					
phase	Time	Tracker	Role	x	y	Proximit y to RN1	Proximit y to RN2	Proximity to RN3	Proximit y to TL	Proximity to TCH	Proximity to PTN
PH1	sec. 1	1	RN1	6132	2471	NA	Social	Intimate	Social	public	Intimate
PH1	sec. 1	2	RN2	1481	6101	Social	NA	Intimate	Public	intimate	Intimate
PH1	sec. 1	3	RN3	…		intimate	Intimate	NA	Social	public	NA
PH1	sec. 1	4	TL			Social	Public	Social	NA	intimate	Social
PH1	sec. 1	5	TCH			Public	Intimate	Public	Intimate	NA	Public
PH1	sec. 2	1	…	…	…	…	…	…	…	…	…

Figure 4.6: Multimodal Matrix representation of co-presence in interpersonal spaces.

4.2.2.3 Proxemic construct: Socio-spatial formations

The notion of socio-spatial formation or *facing-formations (f-formations)* was originally defined by Kendon (1976) referring to the ways people cluster so that they can have direct and equal access to one another (for example in side-by-side, face-to-face, in a circle or L-shapes: see Figure 4.7), and exclude a designated outer space behind them. F-formation analysis has enabled HCI and CSCW research to understand how teams coordinate and communicate to achieve tasks including collaborative information-seeking (Marshall et al., 2011), healthcare (Mentis et al., 2012), and even cooking (Paay et al., 2013). This construct has also been useful to design systems that enrich collocated groupwork with context-aware interfaces that adapt the way content is displayed according to how group members orient themselves (Marquardt and Greenberg, 2015).

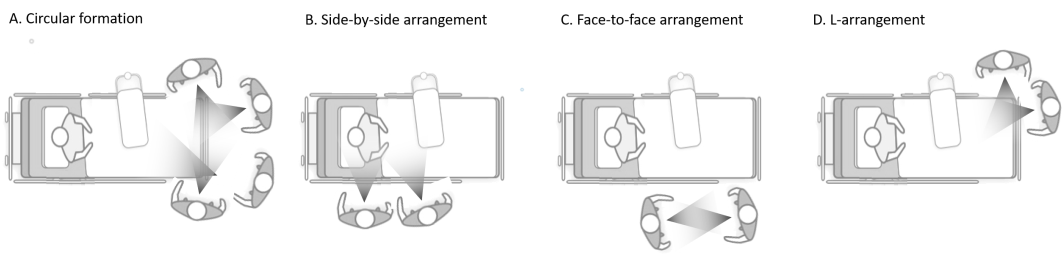


Figure 4.7: Illustrative examples of some f-formations in the context of our healthcare team simulation study.

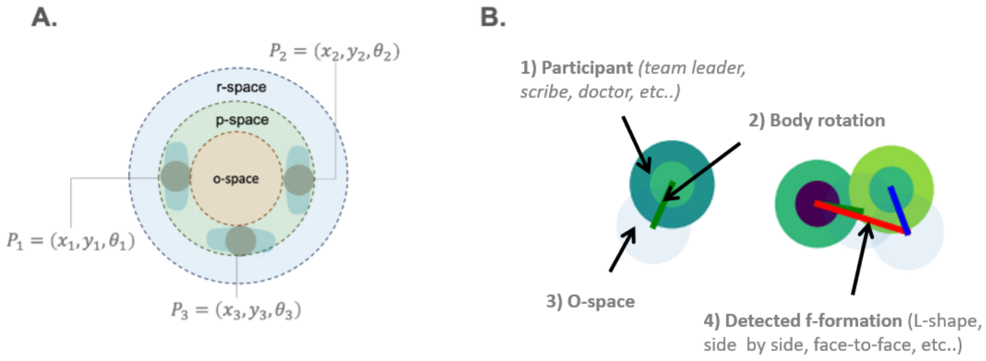


Figure 4.8: Left: Social spaces in an f-formation: o-space, p-space and r-space. Three individuals (P), with position (x, y) and a rotation angle (θ) Right: visual output of the GCFF algorithm that automatically detect f-formations.

Three concepts characterise an f-formation: the *o-space*, which is the joint interaction territory to which team members in the formation have easy access; the *p-space*, the narrow strip of space that surrounds the o-space which is occupied by the team members; and the *r-space*, which is the area that surrounds the o-space and p-space. Figure 4.8 (left) illustrates these spaces and the kind of data that would be needed to automatically model the f-formations: positioning coordinates and the angle of approach of each person. For example, in the healthcare context, nurses collaborating while preparing the IV antibiotic might require an appropriate f-formation, having the clinical instruments located at the o-space, to equally and effectively complete the task.

We used the Graph-Cuts for f-formations (GCFF) method (Setti et al., 2015) to automatically model this construct based on the positioning data. GCFF has been previously used to detect f-formations in static images. GCFF detects when two or more individuals' o-spaces intersect. This intersection is defined as the *transactional segment*, which is the area in front of the body that can be reached easily, and where hearing and sight are most effective between individuals. Thus, given the x and y coordinates of an individual or a group of individuals (see B.1 in Figure 4.8-right) and their *rotation* (B.2), GCFF calculates the probability of each individual occupying a specific *o-space* (B.3) (and thus, a f-formation) when their transactional segments overlap (e.g., see B.4). The detected formations were represented in a MM in Figure 4.9. The f-formations detected are structured in a way that it is possible to elicit if specific roles were in an f-formation during a particular moment and for how long.

LOW-LEVEL POSITIONING DATA							SOCIO-SPATIAL FORMATIONS					
phase	Time	Tracker	Role	x	y	Rotation	RN1	RN2	RN3	TL	TCH	PTN
PH1	sec. 1	1	RN1	6132	2471	0.261	None	L-shape	F2F	None	None	None
PH1	sec. 1	2	RN2	1481	6101	-0.159	L-shape	None	Side-by-Side	None	L-shape	L-shape
PH1	sec. 1	3	RN3	...	...	...	F2F	Side-by-Side	None	L-shape	None	None
PH1	sec. 1	4	TL	...	...	...	None	None	L-shape	None	None	None
PH1	sec. 1	5	TCH	...	...	...	None	L-shape	None	None	None	None
PH1	sec. 2	1	...	...	...	...	...	...	...	...	...	...

Figure 4.9: Multimodal Matrix representation of socio-spatial formations

4.2.2.4 Proxemic construct: Presence in Spaces of Interest

Certain spaces in the classroom can have multiple meanings based on the kind of activity unfolding on the site and the relative proxemics among teachers, students and objects (e.g., devices, furniture) (Pointon and Kershner, 2000). For Hall (1963), such spaces are of three types: *fixed spaces*, which have their shape and size determined by the presence of objects that cannot easily be moved (e.g., walls or screens; *semi-fixed spaces* which are established by movable features in the environment (e.g., tables, beds, curtains and clinical trolleys) that only remain unmoved and unrearranged during peoples' interactions; and *dynamic spaces*, which are formed solely by the spacing and orientation of individuals as they interact with each other.

Spaces of Interest for Simulation 2. The meanings that presence and movement in *spaces of interest* signify for a particular learning design of a simulation have been studied over the course of several years' co-research with the nursing academics, through a combination of formal interviews and prototyping (e.g., Echeverria et al. (2018a) and Martinez-Maldonado et al. (2020a)).

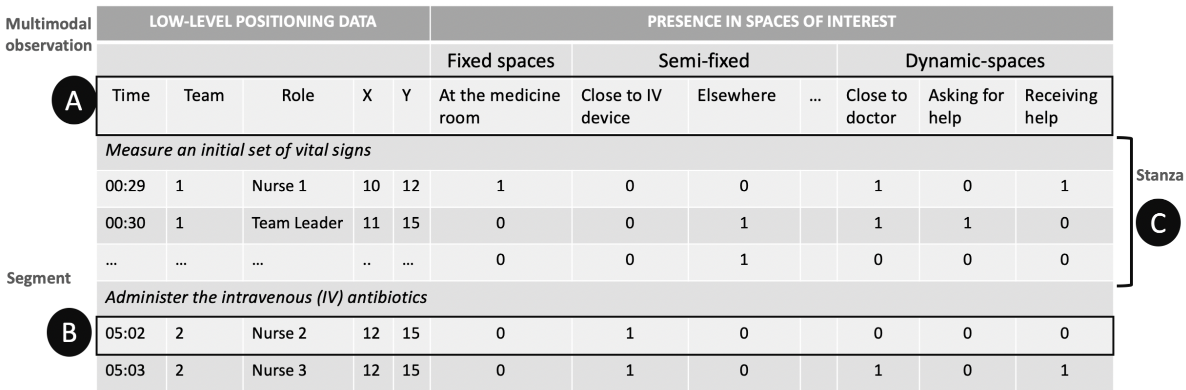
The spaces of interest for this simulation elicited from teachers are summarised in Table 4.4. For instance, the only *fixed* space for this context was *the medicine room* which is a well defined area with medical instruments and supplies (row 1). In contrast, *semi-fixed* spaces were determined by different areas depending on the position of the IV device (row 2), the student acting as the patient (row 3), and the patient's bed (rows 4-6), meaning that these spaces could change depending on the classroom or laboratory where the simulation is being enacted. Finally, *dynamic spaces* were driven by areas where the close proximity between people (e.g., nurse, doctor) could occur and are not attached to a specific location, but instead depend on participant availability. For example, the teacher acting as the doctor moves from bed to bed, and creates dynamic spaces with different nurses as she moves. Based on teachers' interviews, if nurses come close to the doctor outside of the semi-fixed spaces it is commonly for the purpose of *asking for help* (row 8) and if the teacher comes closer to nurses' work area, then she is either just supervising or nurses are *receiving help* (row 9). Rows 8 and 9 are the only ones which do not represent actual physical spaces. These spaces are dynamic because they depend both on the proximity between teachers and students and the location of their encounters. The labels used to represent these spaces indicate how such encounters are commonly interpreted by teachers. All the remaining positions in the classroom were coded as *elsewhere in the classroom*.

Row	Space of interest (codes)	Meaning	Example expected behaviour in current simulation	Type
1	In the medicine room	Here, nurses commonly get medicine and equipment they require for the patient care.	In phase 2, nurses are expected to be at the medicine room retrieving the antibiotic and IV equipment.	Fixed
2	Close to IV device	From here, nurses can check, start and stop the IV device.	After noticing the patient is having an allergic reaction nurses are expected to be close to the IV device to stop it.	Semi-fixed
3	Close to the human patient	Nurses being close to the student enacting the patient can indicate that verbal assessment of the patient is taking place.	In phase 1, nurses are expected to be close to the patient performing the initial assessment.	Semi-fixed
4	Near to patient	At these spaces nurses validate the intubation device (left) and assess vital sings (e.g., pulse, hart rate) (right)	In phase 3, nurses are expected to be near to the patient validating the intubation is working properly.	Semi-fixed
5	At the patient manikin	Being very close to or on top of the patient bed can indicate the patient is being attended. Certain clinical procedures require nurses to lean over the patient's bed.	After noticing the patient is having an allergic reaction nurses should attach the ECG device to the manikin.	Semi-fixed
6	At the bed footer	From here, the team leader monitors and delegates tasks; and nurses coordinate, read charts or write observations.	The scribe should be next to the patient, or at the head/footer of the bead.	Semi-fixed
7	Elsewhere in the classroom	Nurses can be in other spaces interacting with other nurses, finding books (e.g., the Monthly Index of Medical Specialities) validating medication, or looking for the doctor (teacher).	In phase 4, nurses have to notify the doctor that the patient had an allergic reaction.	Semi-fixed
8	Asking for help	Nurses asking for help to the doctor (teacher)	Nurses spending time elsewhere in the classroom and close to the doctor.	Dynamic
9	Receiving help	Nurses receiving help from the doctor (teacher)	Nurses being close to the teacher in any space of interest but elsewhere in the classroom.	Dynamic

Table 4.4: Codes for the meaningful spaces of interest construct.

This construct is modelled by mapping the (x and y) coordinates to fixed and semi-fixed areas identified above. The dimensions (width and height) and shapes (e.g., rectangle areas) of each fixed and semi-fixed space were mapped as two-dimensional areas to assess if a positioning data point was in any of the spaces of interest. Additionally, proximity data between nurses and the teacher was used to identify the dynamic spaces (rows 8 and 9). More specifically, if a nurse was close to the doctor and they both were elsewhere in the classroom, this was coded as asking for help. If they were both present in any of the semi-fixed spaces of interest (i.e. rows 2-6), this was coded as receiving help. This way, each datapoint of each nurse in the dataset is associated with one or two codes of spaces of interest. If more than one team is tracked in the classroom at the same time, the coordinate mappings of the semi-fixed spaces would be defined with reference to the patient bed each team is working at.

Building on the MM, this thesis's is documenting how to drive the modelling of students' positioning data from teachers' pedagogical intentions, and coding *fixed*, *semi-fixed* and *dynamic* spaces of interest within a physical learning space. Figure 4.10 shows a simplified representation of the modelling performed on the positioning data. Figure 4.10 (A) shows the raw data and the *spaces of interest*.



LOW-LEVEL POSITIONING DATA					PRESENCE IN SPACES OF INTEREST						
					Fixed spaces	Semi-fixed			Dynamic-spaces		
Time	Team	Role	X	Y	At the medicine room	Close to IV device	Elsewhere	...	Close to doctor	Asking for help	Receiving help
Measure an initial set of vital signs											
00:29	1	Nurse 1	10	12	1	0	0		1	0	1
00:30	1	Team Leader	11	15	0	0	1		1	1	0
...	...	...	..	...	0	0	1		0	0	0
Administer the intravenous (IV) antibiotics											
05:02	2	Nurse 2	12	15	0	1	0		0	0	0
05:03	2	Nurse 3	12	15	0	1	0		1	0	1

Figure 4.10: Schema showing the application of the Multimodal Matrix modelling technique to indoor positioning data to two phases in the simulation activity.

In this study, I represented the presence or absence of a student in one or more spaces of interest at a given moment (i.e., each second). For this modelling technique, the phases of Simulation 2 (see Section 3.2.1.2) served to group segments into stanzas (Figure 4.10, C).

4.3 Multimodal data modelling - toolkit (software)

The multimodal techniques used to gain insight into the physical dimension of embodied team activity (described in Section 2.2.1.4), are available for other researchers to use them:

- Nurses' levels of arousal. The EDA data captured using the E4 was processed using the EDA Explorer [5] open source code algorithm to detect the EDA peaks from the data (see Section 4.2.1.2).

[5] https://github.com/MITMediaLabAffectiveComputing/eda-explorer

- The co-presence in interactional spaces construct was implemented in python and the code is available in [6]

- Socio-spatial formations. The GCFF algorithm has been specifically developed to automatically identify f-formations from static images. For this reason, GCFF was applied to automatically identify overlaps in transactional segments of two or more people (presented in Subsection 4.2.2.3). The algorithm automatically identifies if nurses were in an f-formation and graphically represents a detected formation. (Setti et al., 2015) original implementation in Matlab is available in [7]

- Presence in Spaces of Interest (SoI). The output of the modelling described above (see Section 4.2.2.4) was processed using the online ENA tool[8] for the duration of the simulation. Unfortunatelly, for confidentiality reasons, the modelling outcomes cannot be shared.

4.4 Summary

This chapter discussed the multimodal modelling approaches and techniques used to extract educationally relevant meaning about team activity from low-level multimodal data. The techniques are driven by learning design, teachers' pedagogical intentions, theory, or a combination. The use of these techniques indicates that eliciting teachers' pedagogical intentions leads to the identification of meaningful educational insights, which are relevant for the specific embodied team activity. Thus, supporting the promise of QE approaches of enriching quantitative data with qualitative insights Shaffer (2017). This section tackles the first research question of this thesis:

RQ1: *What modelling techniques can enable identification of salient aspects of multimodal team activity according to the learning design (teachers' pedagogical intentions)?*

This thesis has demonstrated that many salient aspects of team activity can be extracted from multimodal data using these techniques. For instance, using observed actions (extracted from the learning design) it is possible to identify team errors, such as an incorrect frequency in performing critical actions (e.g., vital signs validation). Using actions and physiological data (e.g., student arousal peaks), it is possible to classify team members' stress levels during critical moments (e.g., patient deterioration). This thesis

[6] https://github.com/Teamwork-Analytics/obs-rules
[7] https://github.com/franzsetti/GCFF
[8] http://www.epistemicnetwork.org/

explored in more detail the physical dimension of team activity, demonstrating how low-level indoor positioning data in combination with teachers' pedagogical intentions and the Theory of Proxemics, opens up opportunities to explore team interaction aspects. Firstly, team co-presence in interactional spaces, was useful to identify team interpersonal interactions, which is an appropriated distance to enable direct interaction (according to the literature Martinec (2001)), where verbal transactions and the majority of intensive and delicate interpersonal interactions occurs Ciolek (1983). In addition, this construct has been used to identify establishment of social ties among team members Cristani et al. (2011). Secondly, modelling *team socio-spatial formations* was useful to identify f-formations (facing-formations), which for the healthcare scenarios is an indication of how the nursing team coordinates and communicates to achieve their tasks during the simulation (e.g., attaching an ECG device). Finally, modelling team presence in spaces of interest (SoI), allowed the elicitation of meaningful SoI for the nursing simulations and the identification of aspects such team effective patient-care and team autonomy.

CHAPTER 5

Evaluation of MMLA interfaces with students

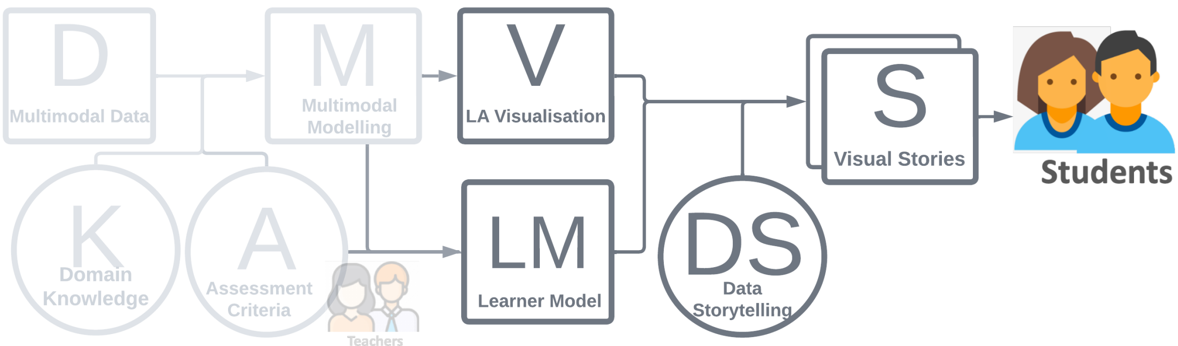


Figure 5.1: Crafting and evaluating MMLA interfaces with students

The third and fourth contributions of this thesis are described in detail in this chapter. This chapter focused on the second component (Figure 5.1) of the conceptual model introduced in Chapter 3. Thus, the chapter presents a set of **MMLA interfaces** *(V-S)* designed based on the modelling approach described in Chapter 4 and its **evaluated with students**. The framework enables the generation of two types of interfaces; interfaces that invite users to **explore** *(V)*, and interfaces that **explain** or communicate insights *(S)*. The framework introduces the learner model *(LM)* and data storytelling principles *(DS)* to enable overlay additional information or enhancement to the visualisation to guide interpretation. Section 5.1 presents the students exploration of two explanatory MMLA interfaces. Section 5.2 provide four illustrative vignettes (exploratory MMLA interfaces) and exemplar insights about team dynamics.

Each of the studies presented in the following sections were designed with various protocols and research objectives. This, with the intention to gain additional insights and perspectives from the evaluation. Figure 5.2, explains how studies 1 (Section 5.1) and 2 (Section 5.2) are addressing the main research questions of this thesis (Section 1.2). The identification of research objectives (Ro) derived from the different qualitative studies are defined based on the study number (St1-St2) and the research objective number (Ro).

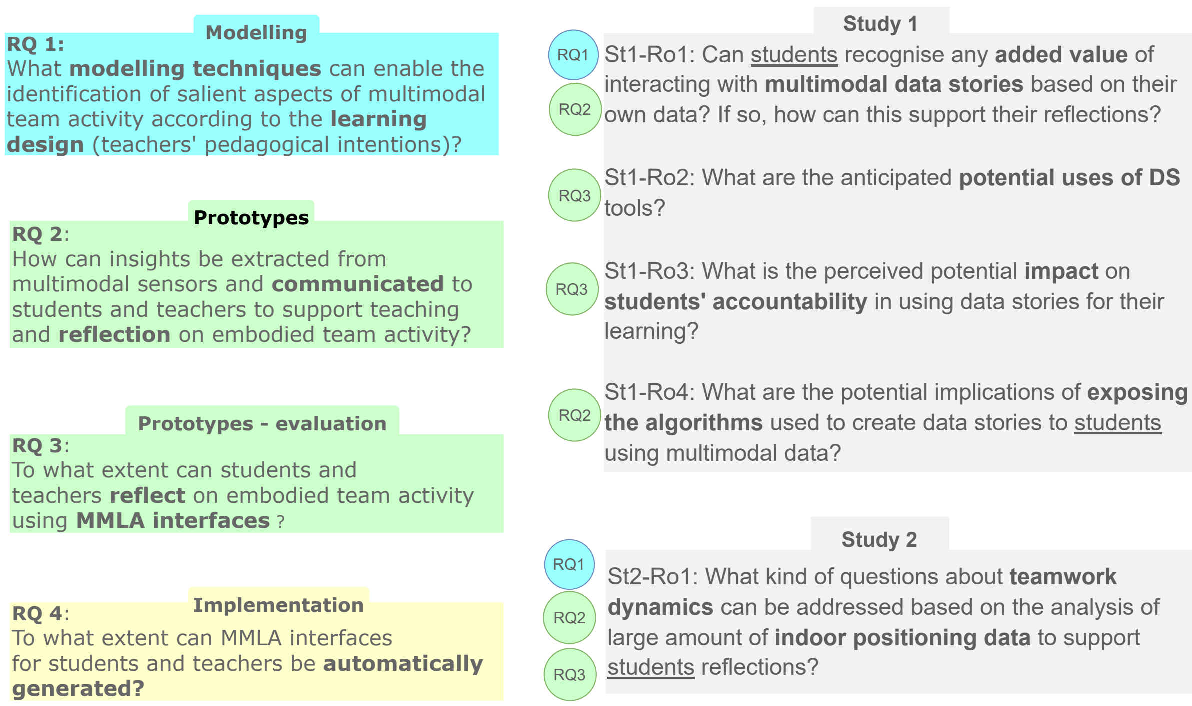


Figure 5.2: Mapping of thesis RQs and the objectives of studies 1 and 2 presented in this chapter

5.1 Study 1: Students reflecting on errors and arousal levels

The section includes two MMLA interfaces from two learning contexts (Simulation 1 3.2.1.1 and Simulation 2 3.2.1.2) and the qualitative results from students' perceptions while using the MMLA interfaces to support reflection on their embodied team activity [1]

The study was conducted using LATEP (Learning Analytics Translucence Elicitation Process), an elicitation protocol for understanding how non-data experts envisage the use of LA systems Martinez-Maldonado et al. (ress). Based on this, the study had the following four research objectives (Ro's) with a specific focus upon the student perspective.

- **St1-Ro1:** Can students recognise any *added value* of interacting with multimodal data stories based on their own data? If so, how can this support their reflections?
- **St1-Ro2:** What are the *anticipated potential uses* of DS tools?
- **St1-Ro3:** What is the perceived potential impact on students' *accountability* in using data stories for their learning?

Inspired by the growing interest in explainable Artificial Intelligence (AI) Wang et al. (2019) to support transparency and improve trust in systems, I also sought to understand: (4) What are the potential implications of *exposing the algorithms* used to create data stories to students using multimodal data?

5.1.1 Participants

Simulations 1 and 2 were conducted in regular classes of the *Integrated Nursing Practise* course in 2019 (semesters 1 and 2, respectively), and all the students were in their third year. We invited the 44 students, in total, who were recorded while enacting Simulation 1 (Section 3.2.1.1) and Simulation 2 (Section 3.2.1.2) to participate in optional team reflection sessions a week after each simulation. These were co-organised with their teachers as an optional extra activity at the end of the class time. A total of 16 out of the 19 female students in study 1 volunteered to participate in the post hoc reflections (aged 19-53 years, mean=27, std=10), and 23 out of the 25 students in study 2 (21 females and 2 males, aged 20-45 years, mean=23.5, std=5.4). Four teams of Simulation 1 *(T1-T4)* and five teams of Simulation 2 *(T5-T9)* had from 2 to 5 students attending this reflection. All sessions were conducted in a meeting room next to the students' regular classroom.

[1] Based on: **Gloria Milena Fernandez-Nieto** et al., "Storytelling With Learner Data: Guiding Student Reflection on Multimodal Team Data", in IEEE Transactions on Learning Technologies, doi: 10.1109/TLT.2021.3131842

Learning Intention	Rule pseudocode
RULE 1 – used to generate a story about an **error made** shown in Figure 5.5, example 1. Administer Fentanyl within 10 minutes adter the patient complained of abdominal pain	IF time_between(Abdominal_pain, IV_medication)>10: add_annotation("The patient needed pain relief ...") add_arrows(annotation, data_point) highlight_datapoints(data_point, "orange") add_rectangle(Abdomical_pain, IV_medication, "orange") add_title("The team administer Fentanyl"+ X + "min late") ELSE add_annotation("The team did it right! Well done!") highlight_datapoints(data_points_list, "blue")
RULE 2 – used to generate a story on **arousal** shown in Figure 5.8, example 4. How aroused was a nurse after the patient complained or experience chest pain.	SPLIT (Max_arousal_ratio/4 quartiles) SWITCH (Arousal_ration_in_a_stanza): > Quartile 4: add_annotation("Very High") in Quartile 3: add_annotation("High") in Quartile 2: add_annotation("Mild") in Quartile 1: "Low" do nothing < Quartile 1: "Very low" do nothing add_title(role + "presented several arousal peaks ...")

Figure 5.3: Rules used to highlight visual elements in example 1 and 4 in Figure 5.5 and 5.8.

5.1.2 Protocol

The reflections were conducted as a 30-minute *focus group*, which consisted of reflection sessions with students (S1-5) from each team (T1-9). This study design reflects the standard approach used in this class-based scenario, in which it is normal to conduct a clinical debrief after team simulations. The sessions were structured as follows:

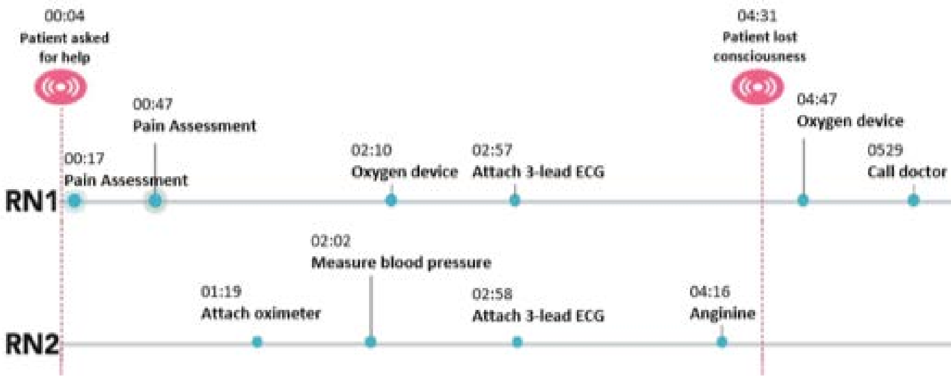


Figure 5.4: Example timeline of actions for a team of two nurses, which served as the background of both MMLA interfaces.

1) *Pretest on procedural knowledge*. Before the group reflection, students were asked

to individually list all the nursing actions that should have been performed at specific times during their simulation. A blank version of the timeline (Figure 5.4), showing only the actions of the patient, was provided on paper for them to annotate *(aim 1)*.

2) *Think-aloud team reflection*. i) In teams students were asked to think aloud while jointly inspecting the timeline (without data stories) of their own team (e.g., Figure 5.4). ii) Next, teams were asked to explore the *enhanced timelines* mocked up as interactive screens (e.g., Figure 5.6), in any order they wished. Then, they were asked about their views on these versions *(aim 1)* as follows: a) Did the data stories add value to the timeline (without stories)? and b) how do you think your performance was based on the data stories?

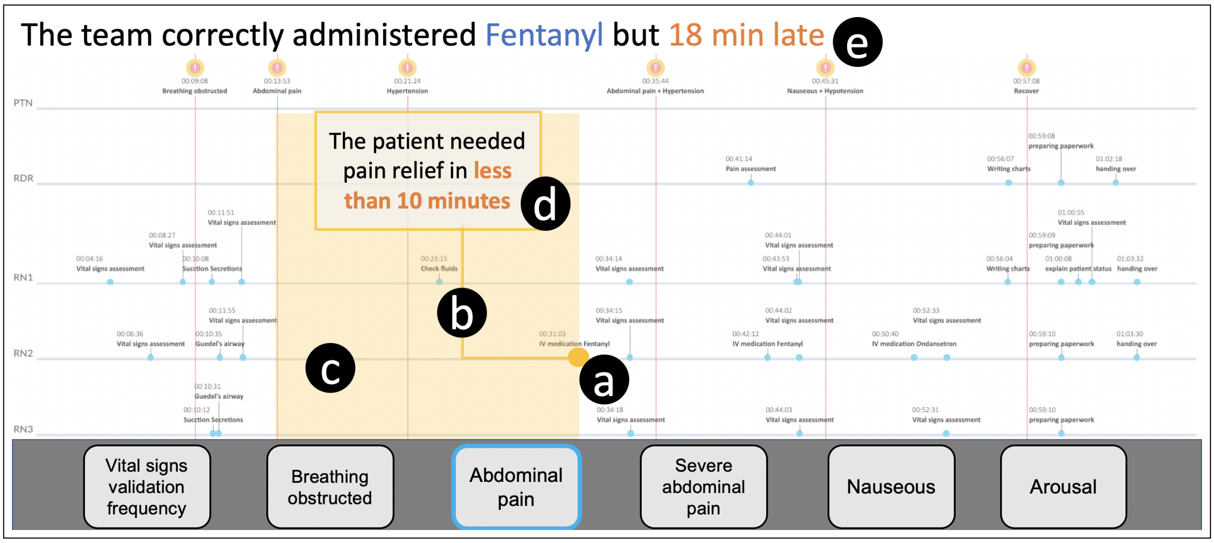


Figure 5.5: MMLA interface example 1. Showing a story on an error made in terms of time responsiveness.

3) *Usage and accountability*. Students were asked about usage and accountability opportunities and concerns *(aims 2, 3)*, as follows: i) how can the system be used in or outside the classroom? and ii) who should be able to see the interface, for which purpose and in what form?

4) *Explainability*. Students were asked to review the interface that exposes the algorithm (Figure 5.9). Students were asked if they understood the rules used to craft the stories and whether they would like to see these rules in the interface *(aim 4)*.

5) *Posttest on procedural knowledge and student perceptions survey*. Finally, students were asked to go back to the paper version of the timeline annotated in step 1, and correct/add/remove any actions they wished *(aim 1)*, using red ink on the paper. A survey

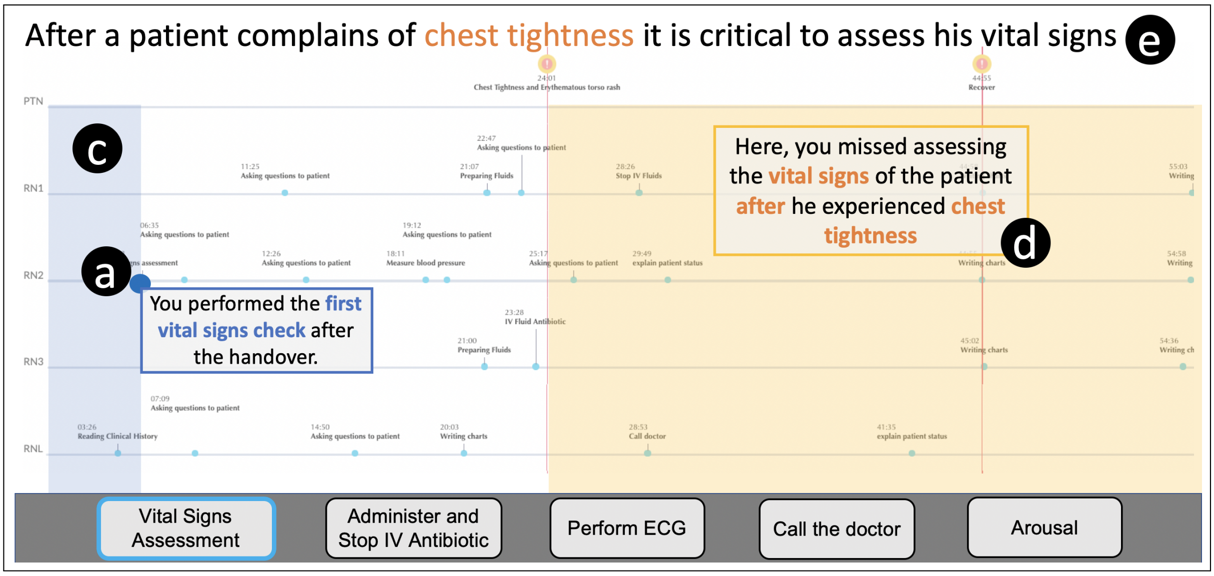


Figure 5.6: MMLA interface example 2. Showing a story on an error made in terms of actions omission.

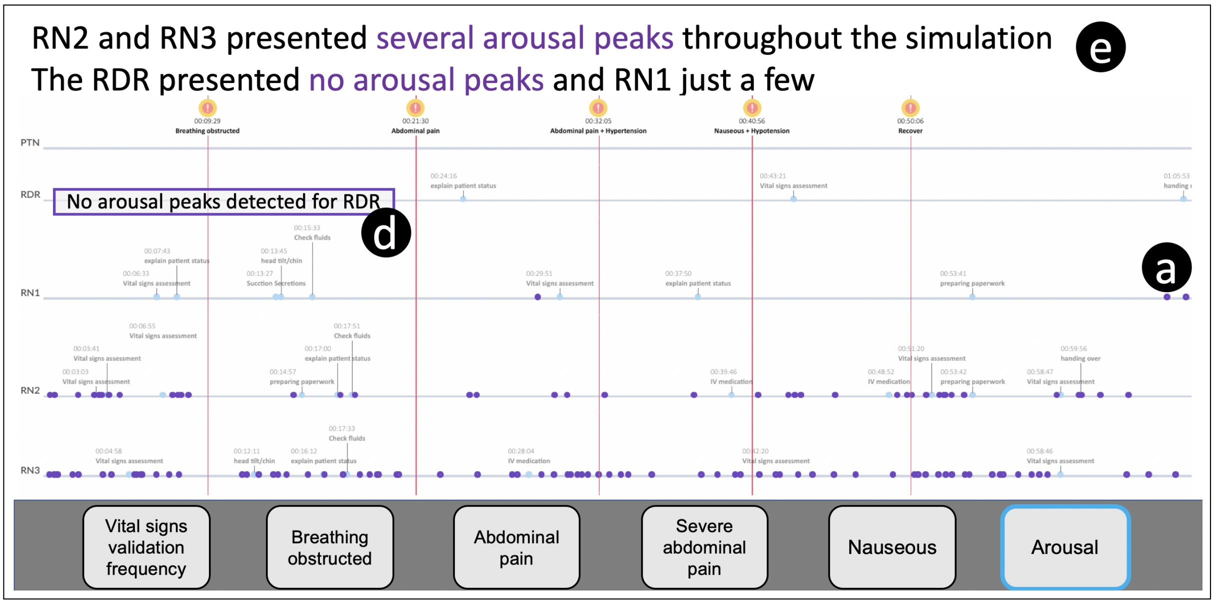


Figure 5.7: MMLA interface example 3. Showing a story on arousal with all relevant detected peaks

questionnaire was provided to the students to elicit their individual perceptions about the MMLA interfaces. The MMLA interfaces were presented to students using a 21-inch Liquid Crystal Display (LCD) display, connected to a laptop running macOS 10.14 with a wireless mouse for students to explore the stories.

5.1.3 MMLA interface: What did the team miss?

Figure 5.3 presents two exemplar rules used to create the data stories shown in Figure 5.5 and Figure 5.6 respectively. Using such rules, two storytelling MMLA interfaces were created. Following principle **DS2** (selection of an appropriate chart type, Section 2.4), both MMLA interfaces were based on a minimalistic visualization technique (**DS3**: strip out excess detail, Section 2.4), namely, a *timeline of actions* by each nurse within a team during the simulation (captured by the manikin and the observer, presented as annotated blue circles on each nurse's line of actions-see Figure 5.4). Actions performed by the patient (such as asking for help or complaining of severe pain) are represented by red vertical lines, which also divide the dataset into stanzas. Since line charts effectively show changes over time (Ryan, 2016), this visualization without any visual enhancement was used as the background in both MMLA interfaces. Student navigated both MMLA interfaces through a user interface. The MMLA interfaces were accessible via a set of buttons located at the bottom of the interface of the *timeline of actions* (see Figure 5.5 and Fig 5.6).

Endorsing principle **DS4** (Section 2.4), selected elements of the visualisation were emphasized by *adding enhancements* such as: a) icons (see Figure 5.5-5.8, b) arrows , and c) enclosing areas; *changing color*, contrast, or thickness (also see elements marked as A, B, and C); and d) *annotating* salient data points, or e) *adding titles* that summarise the take-away message.

For example, in Figure 5.5 an *if then else* rule was applied to assess the presence and timing of a critical action in a certain stanza of the logged activity. The algorithm adds visual elements to the visualization, including: (a) orange icons to highlight the relevant data points; (c) a colored enclosure area to emphasize the stanza where the error was detected; (d) an annotation where the error occurred: *"The patient needed pain relief in less than 10 minutes"*; and (e) a prescriptive title : *"The team correctly administered fentanyl but 18 min late"*.

5.1.4 MMLA interface: What were students' arousal levels?

In Figure 5.8 a more sophisticated *switch* algorithm (rule 2 in Figure 5.3) assesses the ratio of arousal peaks for each nurse in each stanza and compares it to the highest ratio of arousal peaks experienced by a single student that we have detected in all of our nursing simulation studies (5 peaks/minute). This maximum ratio is divided into quintiles of equal size which are used to categorize the arousal experienced by a nurse (as

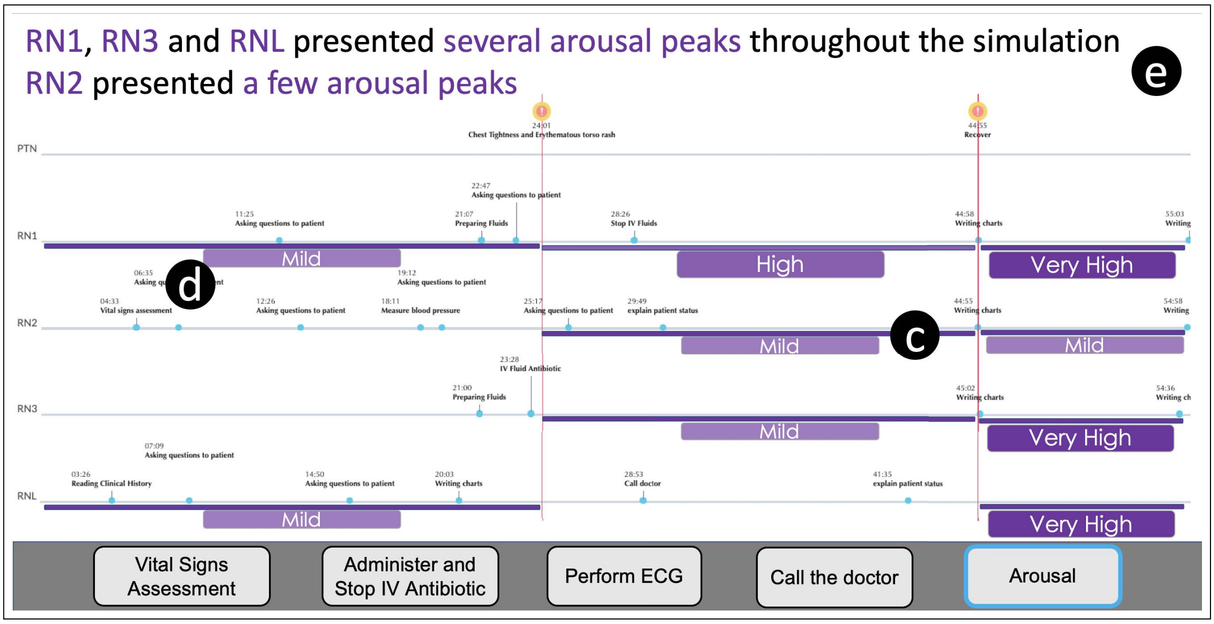


Figure 5.8: MMLA interface example 4. Showing a story on arousal showing annotations only.

very low, low, mild, high, or very high) in any given stanza. Only the last three categories are shown in the interface to provoke discussion on high levels of arousal.

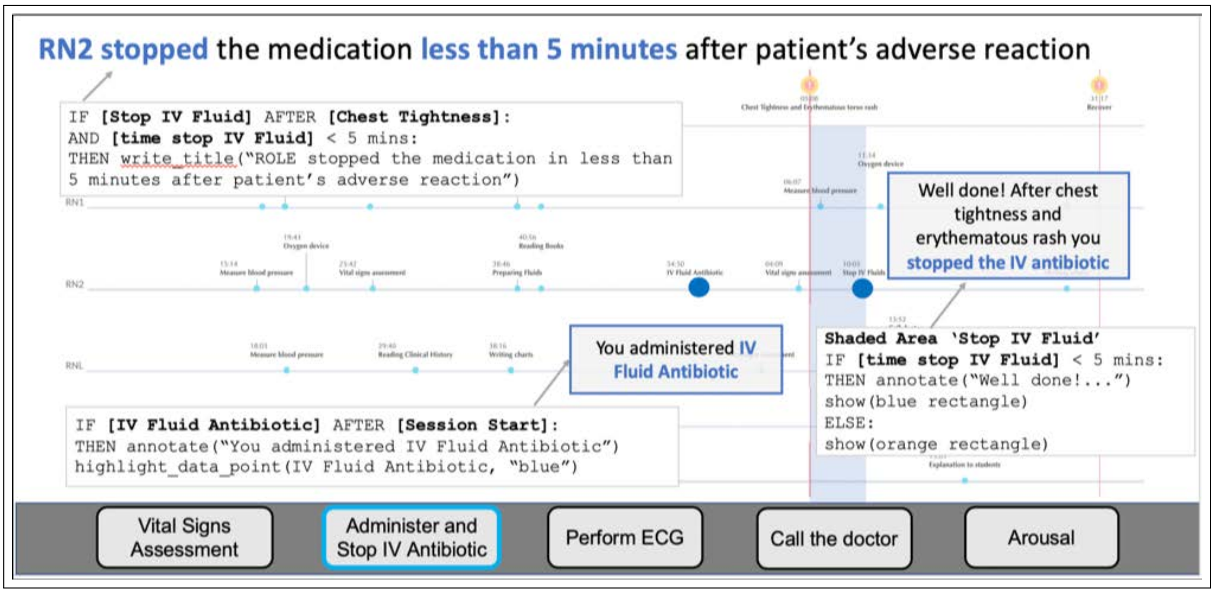


Figure 5.9: MMLA interface opening the algorithm used to enhance the timeline of events.

5.1.5 Results

This section presents the results of the analysis organised around the four research objectives presented above (see 5.1). The results of this study were generated following the qualitative analysis explained in Section 3.3.5. The emerging topics are quantified based on the number of times a 'topic' was mentioned within a team (not per individual) in the focus group.

5.1.5.1 Added Value of the Data Stories

Compared to the pre-test of procedural knowledge applied before the exploration of the data stories, all teams showed a decrease in the number of errors in the posttest (defined as omitted actions and slow responses according to the learning design). Students made an average of 3.6 errors (Study 1, out of 5 actions) and 2.3 errors (Study 2, out of 6 actions), before inspecting the data stories, reduced to 2.6 and 1.8 errors respectively, after the reflection activity. This suggests that the reflection activity helped them to identify the errors they had not correctly identified. It is possible, however, that any activity that engaged them in reconstructing a week–old simulation might have helped refresh their understanding, so deeper insights are needed to claim that the data stories were adding specific value.

During the think-aloud protocol, students explained in detail their perspectives on how the data stories offered guidance for them to interpret the timeline.

Students in all nine teams who participated in the reflection appreciated the guidance offered by each data story to be able to **focus** on the expected learning goals. For example, a student explained: *"[the data stories] divide the screen into sections [critical incidents of patient]. These sections looked like the same thing in [the timeline of actions without enhancements]. It does not highlight the things that needed to be done and the timeframes"* (T2, S2).

Other students more explicitly explained the value of reflecting through the data stories, because they were able to clearly see the mistakes made during the simulation for example: *"[data stories] highlight what you did not do [right]"* (T8, S4); and *"[the data stories] show both what we did and what we were meant to do"* (T8, S4).

Some DS elements and their usefulness were emphasized by students during their reflections. For example, two teams mentioned that **enclosing, colored regions** helped them to understand the expected timeframe in which they were meant to perform certain actions, and five teams agreed that the **annotations** informed them about how they could act upon the information: *"It is not just that we were late but we were late by that much time, so we can precisely see what we have to improve, and what we have done well. [The data stories] make a big difference"* (T1, S2).

Students also commented on the designs of the MMLA interfaces (V1, shown in Figure 5.7 and V2, Figure 5.8) and commented on how the **design decisions** impacted the interpretability. Three out of the five teams in Study 2 found that V2 was less complex than V1. Students explained their preference in terms of the simplicity of V2 to more quickly digest the arousal data, compared to looking at every arousal peak in V1. One student explained: *"We just want a quick snapshot to see how [our arousal] was. That is really interesting to see in [V2]"* (T6, S5). Another student compared both MMLA interfaces and explained: *"I tried for a long time to understand [V1]. I do not know how to understand this, but this other one [V2] is pretty straightforward"* (T9, S5).

Students in the other 2 teams suggested that both views could be combined. A student explained that *"because, the dots [in V1] can be more specific"* (T8, S2) students could first view V2 and then see the V1 details on demand. Another student suggested a third option, as follows: *"You can combine both [V1 and V2], using short and long lines to show arousal peaks"* (T9, S4).

Overall, students appreciated the benefits of augmenting the timeline of actions with annotations and visual elements that explained the errors they made, or the extent to which they experienced arousal during the reflection. Additionally, a survey was performed at the end of the interview to elicit students' perceptions of the added value of each layer, the results of which are described in Figure 5.10. The survey results support the students' interview comments. In the next section, we report more in detail how students used the data stories to reflect on the simulation, and anticipated usage strategies.

5.1.5.2 Anticipated usage strategies

This subsection is divided into three parts, presenting anticipated usage opportunities of the data stories regarding i) errors made, ii) arousal, and iii) the storytelling tool as a whole.

1) Use of the tool for reflection on errors. Students in 7 of the 9 teams considered that the data stories on errors provided them with the **opportunity to improve** their clinical practice. For example, one student stated that the data stories *"showed [them] what [they] did well and what [they] should have done to improve"* (T3, S4). Some students mentioned that each data story helped the team to focus on improvements to be

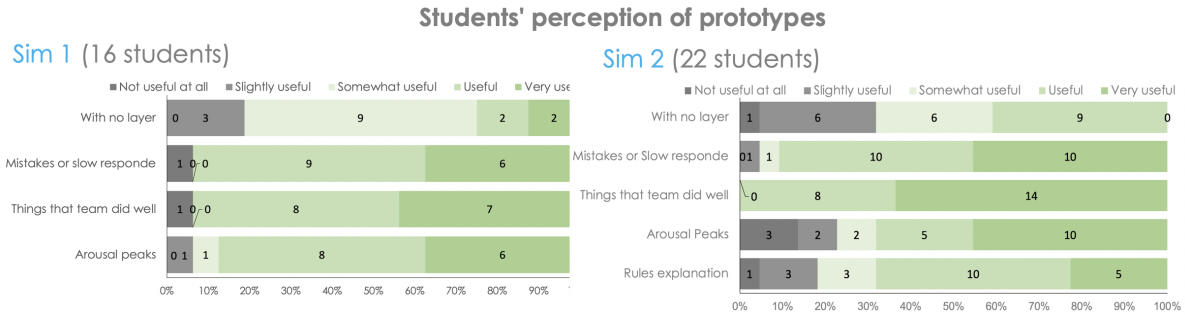


Figure 5.10: Survey results of students' individual perceptions of the added value of the layered visualization: with no layer, timeline without enhancements (Figure 5.4); mistakes or slow responses and things that team did well (Figure 5.5-5.6); arousal peaks (Figure 5.7- 5.8); and rule explanations presented to students of Simulation 2 (Figure 5.9).

made. This was stated by one student as follows: *"[the data stories] help us to reflect on practical ways to improve and work as a team"* (T3, S5). They expressed that seeing their actual performance, with errors highlighted, made them think about **specific skills and knowledge** they need to master. One student said: *"[the data stories] highlighted areas where I need to strengthen my clinical thinking"* (T2, S4). Another student more specifically listed the kind of skills she had to strengthen as a result of her reflections: *"[the data stories] showed me that I need more knowledge about PAC [post anaesthetic care] nursing, I need more knowledge about gaining trust in postoperative, hypertension etc."* (T2, S2).

Students suggested some **strategies** for making use of the storytelling tool. For example, some of the students' reflections were quite individualized, suggesting potential uses of the tool to support individual reflection or reflection on particular roles (3 teams). For example, one student reflected as follows: *"I can use the timeline to see when and what action I did during the whole simulation, and it is very useful and helpful for me to reflect and analyze my work"* (T7, S2). In contrast, students in 7 teams suggested that the tool should be used to provoke group reflection, for instance: *"The timeline helps the group to recall what we did because sometimes we may forget about the details or the sequences of our own actions"* (T7, S3).

Reflections about **collaborative** skills and **teamwork** were shared by students while looking at both their own errors and those made by others in their team. For example, students mentioned that they could see the *"importance of teamwork in clinical scenarios"* (T3, S1), *"identify collaborative work as a team"* (T2, S4), that they needed to find *"ways to improve and work as a team"* (T3, S5) and work on improving their *"critical thinking"* (T2, S4).

2) Use of the tool for reflection on arousal. Students in 5 teams who exhibited very high arousal reflected on what they were doing at those points in the simulation **actions**. For example, one student explained: *"I was writing notes. Everyone else was talking and I was writing notes"* (T4, S2), or *"[my very high arousal] was more like during the airway"* (T4, S4). Another explained that not knowing which actions to perform during the simulation, made them feel under stress: *"I was nervous because I didn't know what to do in the situation. I tried my best to do it, but when I gave the handover [to the Doctor] and he did not accept it, I got more nervous. I had to come back and do it again"* (T3, S1).

Seven students associated their arousal peaks with their **roles**. For instance, a **team leaders** who displayed very high arousal peaks explained that this was due to their engagement with the team and the responsibility with the patient. For example, the team leader in T6 explained this as follows: *"I wanted to make sure I was doing the right thing for the patient. I had to constantly remind myself and the team members about the actions that had to be performed"*. Similarly, the team leader in T8 stated: *"I am very calm under pressure. I think I started mildly aroused because I thought: I am the team leader! I was chilled then because I knew what was going on and I was very confident with what was happening"* (S1). Another team leader reported: *"At first, I felt a bit of stress. Then, I did not feel panic or anything like that, but I was just more concentrated on the task than before"* (T9, S4).

Students who performed the **scribe role** had mixed perceptions about their high and low arousal. One scribe who exhibited a higher arousal mentioned that she was very engaged with the simulation: *"I think I was trying to work out how we should deal with the situation. I got excited about that"* (T6, S5). Another scribe who exhibited a lower arousal explained she was disengaged due to the irrelevant duties for her role: *"I am only the observer. I am not the nurse" (T9, S4).*

Interestingly, students in 3 teams pointed out that **external factors** such as tiredness, fatigue, or sickness might have affected their arousal during simulations. One student explained this as follows: *"It was just later in the day, and you are tired, and want to go home"* (T3, S4). Another pair of students reflection on their arousal peaks reported that one felt very nervous as she had *"injured herself"* (T9, S3), while the other *"was sick"* (T3, S1).

In addition, students from 3 teams suggested that **wearing external devices and being recorded** could cause bias in arousal peaks. One student who exhibited higher arousal levels expressed how she felt during the recordings and data collection: *"I felt a bit stressed because [the group] was getting filmed"* (T7, S1). Similarly, one student in T8 mentioned that being recorded to some extent could cause spikes in arousal readings: *"As soon as anyone puts anything on me, like I am going to check your heart rate, I am like, no! So, my heart rate goes up in a way"* (S1). Nevertheless, the other 6 teams mentioned specifically that they felt comfortable using the devices, and felt that they behaved normally during the simulation.

3) Other uses of the tool in the classroom. In 6 teams, students described the value of the visualizations during the usual teacher–led **debriefs**. For example, a student envisaged how the tutor could help them to go through the data stories to reflect on further medical conditions and procedures as follows, which is not possible at present for logistical reasons: *"The tutor can actually explain things, if you have something to go off. Usually we do our simulation and then we forget about it and never look at it again, or they [tutors] would not know what we have done anyway, because they are not watching each person"* (T7, S3).

Additionally, students in 5 teams mentioned the possibility of using the tool for the **assessment** of their performance and the actions they are intended to perform when enacting certain roles. However, most students (in 6 teams) preferred the tool to be used to enhance the provision of feedback (formative assessment) as opposed to perform automated summative assessments. Surprisingly, none of the students argued against the incompleteness of those summative assessments. However, they were mostly worried about the additional pressure this would put on them, as stated by one student as follows: *"I feel like an assessment would be so stressful"* (T6, S2).

To summarise, this section has illustrated students' broadly positive responses to the MMLA interfaces, and the different ways in which they envisaged their use. Students agreed that the tool would be very useful to enhance the debriefs with their teachers, which are currently performed without using any evidence, relying only on what teachers can see while dividing their attention between 5–6 teams.

5.1.5.3 Accountability and Privacy

Discussions about reflecting on the data stories and the mediating role of teachers was a matter of concern for some students. Students in six teams mentioned that debriefs guided by the data stories could provide a better team discussion, even without the teacher, and that they were willing to compare their performance with other teams (6 out of 9 teams) and *"see what each group member did"* (T3, S3). However, students in these teams agreed that this social comparison should be led by the tutor: *"Do not just give it to the student, get the tutor to sit down with the group and talk it through"* (T3, S3). Students were asked to explain their views about sharing their data with other students, tutors, or teams. Students in all teams agreed that it would be fine to **share** their timeline with other students, especially *"If it helps [other groups]"* (T3, S2) to reflect on their mistakes. One student (S4) added: *"If it is a great mistake, I want to share it with others"* (T9).

However, some concerns about **anonymity** were also raised in 2 teams who thought that the timeline should only expose activities performed by specific roles, without disclosing the names of the students. One of the students emphasized: *"you [researchers] do not have to put our [team members] names on it"* (T5, S1). Only one student suggested that their data should not be shown to other students. However, her rationale was that other students would find it boring to explore other team' timelines *"because it does not relate to what they have done. They cannot learn from it. I have to learn from what I have actually done"* (T1, S1).

Surprisingly, none of the students raised concerns about sharing their arousal traces (even when asked specifically about their perceptions about data sharing), which we attribute to them feeling comfortable in the classroom as an academic space. As mentioned by (S4): *"the lab is a safe environment for humans"* (T9) and they knew their data would be used only for learning purposes.

5.1.5.4 Explainability of the Rules

When students (only the 5 teams of Sim 2) were introduced to the rules used to craft the data stories, most students mentioned feeling confused, and hesitated to respond (e.g., *"You mean all those boxes? I am a bit confused"*, T9, S5). Students in three teams requested further explanations, for example: *"If you [interviewer] do not explain it, how long would it take to understand this stuff?"* (T9, S4). Other students judged that the rules needed some basic coding skills to be fully understood (e.g., *"If you do not know coding, it is hard to understand"*, T9, S4). Some students however, demonstrated that they could understand them, "translating" one rule: *"If time is less than five minutes, the box would be blue. Otherwise, it's going to be orange because it is going to be like a mistake"* (T6, S1).

When students were asked to comment on the **added value** of seeing the rules they had contrasting views. Two students from teams 6 and 8 asked to include the rules into the timeline. One of them suggested alternatives to present them, such as adding extra explanations in the existing stories on–demand: *"Can't you [researchers] just incorporate the rules into each [story] instead of doing it in an extra interface"* (T8, S1). However, another student (team 7) argued that by inspecting the rules, students could potentially solve emerging questions about their mistakes while exploring the timeline. The student argued that if she makes any mistake she will: *"ask a question about why? Why is that wrong? Like it is [data stories] feedback or whatever it is called, it should be constructive"* (T8, S1). The rest of the students (in teams 5 and 9) mentioned they would not find the rules useful and that they would not try to change the rules (*"I would like to keep [the parameters] like that"* T9, S1). Two students stated that teachers would be in a better position to *"understand the parameters"* (T5, S2) and that they would also need *"basic [concepts of] coding"* (T8, S1) to change them. Interestingly, in the survey (see Figure 5.10) students rated the explanations in the rules layer as *mostly useful* (10 students) and *very useful* (5 students.) It is possible that they either misinterpreted what this question was asking about, or were more honest in the interviews than in the survey.

5.2 Study 2: Students' Teamwork Proxemics dynamics

The second study evaluates the modelling approach explained in Section 4.2.2 [2]. The study depicts four MMLA interfaces in the form of vignettes and illustrative insights on how students can use them to support their reflections.

This section evaluates the MMLA interfaces resulted of the modelling approach presented in sections 4.2.2, about modelling techniques of indoor positioning data. The MMLA interfaces were selected because they respond to authentic teachers' questions and serve to show how positioning data can help in telling data-informed stories about nurses' proxemic behaviour.

This study seeks to address the research objective: **St2-Ro1:** *What kind of questions about teamwork dynamics can be addressed based on the analysis of large amounts of indoor positioning data?* by using indoor positioning data to interrogate what analytics for teamwork proxemics reveal about positioning dynamics in clinical simulations. Figure 5.2 shows how the objectives of this study are addressing the main RQ of this thesis.

[2]Based on: **Gloria Milena Fernandez-Nieto**, Roberto Martinez-Maldonado, Vanessa Echeverria, Kirsty Kitto, Pengcheng An,and Simon Buckingham Shum. 2021. What Can Analytics for Teamwork Proxemics Reveal About Positioning Dynamics In Clinical Simulations?. Proc. ACM Hum. Comput. Interact.5, CSCW1, Article 185 (April 2021),24 pages. doi: 10.1145/3449284

Each group of MMLA interfaces includes: a) the teacher question being addressed (the emic perspective) about the nurses' spatial behaviours as expected in each phase (see Table 4.2); b) an illustrative analysis of positioning data from one or more of the 11 teams who participated in our study; and c) a set of insights that can be gained from the modelling of a specific proxemic construct.

5.2.1 Illustrative Vignettes: Co-presence in Interactional Spaces

Following the modelling technique explained in Section 4.2.2.2, this section illustrates the technique with two vignettes as follows:

5.2.1.1 Data vignette 1: the interactional space between nurses and the patient.

a. Context and the teachers' question. The data vignettes presented in this subsection address the teachers' question: *Were nurses around the patient during the simulation?* Throughout the simulation, nurses should assess the status of the patient regularly because strong medication is being administered. Moreover, close proximity to the patient is associated with effective communication and reassures the patient that they are being cared for, which is critical to patient-centred hospital practice (Webster et al., 2019).

b. Analysis. For this illustrative example, social network analysis (SNA) was performed on the output from the modelling of co-presence in interactional spaces described in Section 4.2.2.2. This is the aggregated time nurses spent in close proximity with both one other and the patient (manikin plus human role-player). SNA is a tool that is commonly used to investigate social structures represented using nodes (team roles and the patient in our case) and links (representing social ties of some sort) (Haythornthwaite, 1996). For this purpose, SNA can be an effective analysis technique to model presence of nurses in the patient's interactional space (i.e. connections between student roles and the patient based on physical proximity). For this particular vignette, the *links* represent the average time that students spent in interpersonal proximity to the patient, thus serving as a proxy for close patient care. Thick (dark blue) links indicate longer periods of time in close proximity (>50% of the phase duration). Red (thin) links indicate shorter periods of time in close proximity (<=50%).

Two kinds of interpersonal social graphs can be generated: *full proximity networks*, portraying physical proximity among all team members (Figure 5.11, left); and *role-centred ego networks*, revealing personal proximity between roles in relation to a central role (i.e. the patient) and the focus of these data vignettes (right). All networks were normalised based on the weighted time average of co-presence to enable comparison among phases and teams.

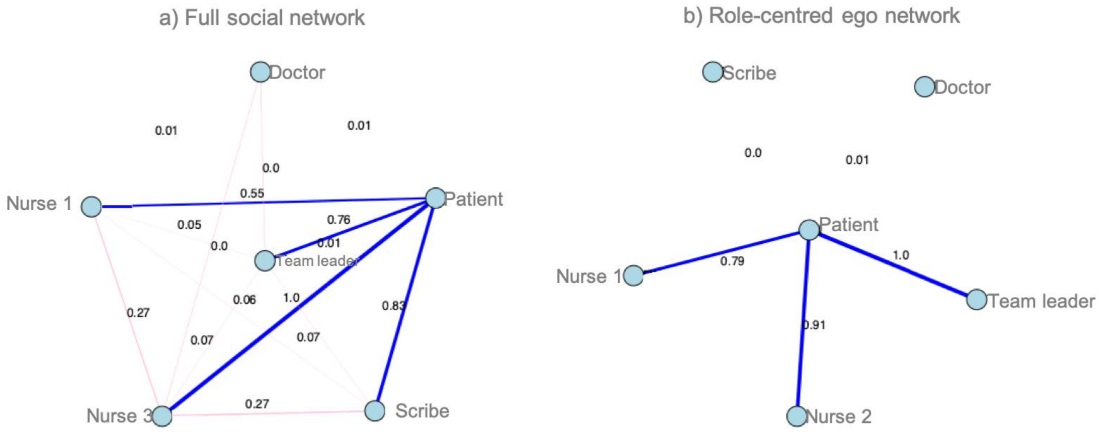


Figure 5.11: Example of social networks representing mutual presence in interactional spaces during phase 1: a) a full proximity network of team 4 (left) and b) role-centred ego networks focused on the patient in team 4 (right). The labels near each edge indicate the proportional amount of time two people were in close proximity to each other.

c. Exemplar insights. Since the patient is central according to teachers, patient-centred ego networks can be used to identify the presence of team members in the patient's interactional space. Figure 5.12 depicts social ego networks for teams 1 and 2 for phases 1-4 of the simulation. Nurses in team 1 (Figure 5.12, top) were strongly connected to the patient. During all phases, at least one of the nurses enacting active roles (i.e. the team leader and nurses 1 and 2) were at the interactional space of the patient. Although the scribe nurse, on average, spent most of her time in close proximity to the patient, she was acting as an observer and it was not intended to be performing an active role. As a result, the fact that two active nurses were always close to the patient while having an adverse reaction (with the exception of phase 2) suggests this team was effective when providing patient-centred care.

In contrast, Figure 5.12 (bottom) demonstrates that the patient-centred ego network for team two had a weak presence of active nurses in the patient's interactional space. The only strong presence of active nurses near the patient occurred during phase 1 (nurse

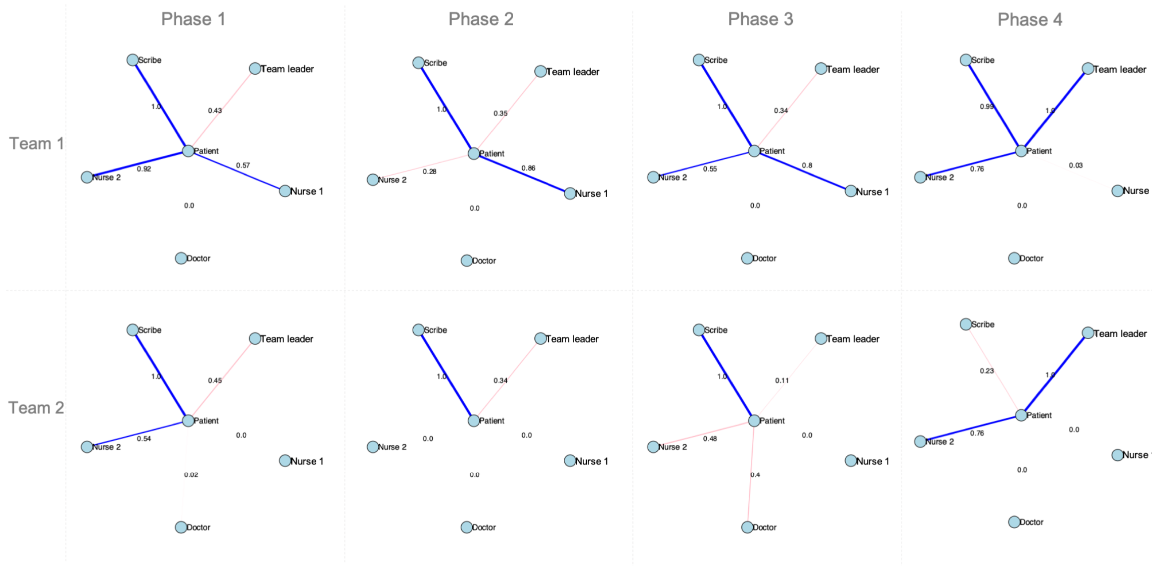


Figure 5.12: Patient-centred ego network for teams 1 and 2 from phase 1 to 4.

2) and close to the end of the simulation (nurse 2 and the team leader). In fact, phases 2 and 3 show only the scribe close to the patient (i.e. none of the active nurses). Compared to team 1, these networks suggest a weaker patient-centred attention. Moreover, the lack of connections between nurse 1 and the patient might be an indicator of disengagement of nurse 1 with the patient.

Overall, responding to the teachers' question (were nurses around the patient during the simulation?), five out of the eleven teams failed in being close to the patient in at least one of the phases, which means during that phase none of the team members were in interpersonal proximity with their patients. This is a potential area to be improved by these pre-service nurses, which was automatically highlighted by the analytics.

5.2.1.2 Data vignette 2: positioning of the team leader

a. Context and the teachers' question. Teachers expect the leader to play a central role in phase 1 of the current simulation, because this is when the team is assuming the responsibility of taking care of a new patient. Because of that, teachers may raise the question: *were nurses together (around or close) to the team leader during phase 1?*

b. Analysis. Similar to the previous vignette, SNA can also be used to model proximity ties among nurses. To analyse whether team leaders in various teams played a central role in phase 1, a full social network representing co-presence in interactional spaces can be used for comparison. These networks were also normalised to enable team comparison. From SNA, the metric *degree of centrality* was used to identify the most connected role in each team in phase 1 (Landherr et al., 2010).

c. Exemplar insights . Through the following data vignettes we compare two teams, *team 2* and *team 5*, which exhibited contrasting spatial behaviours in relation to the team leader. The leader from team 2 was not surrounded by other team members (low weighted centrality= 1, considering ties >0.5%). By contrast, the leader from team 5 played a central role during this phase (higher centrality= 3). These two teams are representative of 3 and 8 other teams in which the team leader also showed a low or high centrality, respectively.

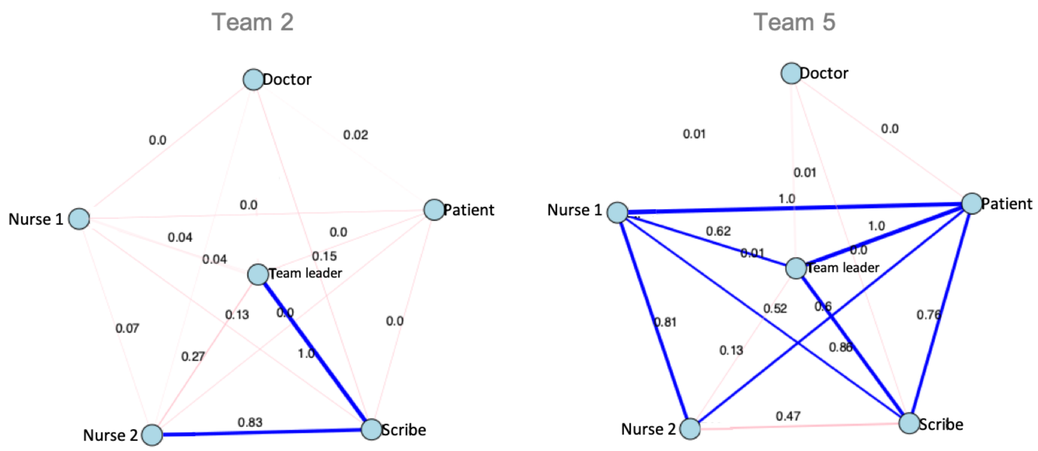


Figure 5.13: Full social networks for teams 2 and 5 during phase 1.

Figure 5.13 (left) shows these contrasting behaviours in more detail during phase 1. Although the team leader was in close proximity to the scribe during the whole phase, he/she only was in close proximity to the main nurses 1 and 2 during 0.04% and 0.27% of the time. This team leader was not close to the patient to a great extent either (0.15% of her time) making it reasonable to expect this nurse to have called the other nurses to come closer during phase 1 to coordinate their work for the rest of the simulation.

By contrast, Figure 5.13 (right) shows multiple similarly thick links between all the nurses, which are also connected to the patient. In fact, the team leader was in close proximity to the patient during the whole phase and his/her time of co-presence in other nurses' interactional spaces was 0.62%, 0.52% and 0.86% for nurses 1, 2 and the scribe, respectively.

Although this would be an expected behaviour from an effective team, this was not ideal either, since it is not recommended that the patient listens to nurses' talking about their case as they are planning an intervention. In short, a leader that is too close could result in patient discomfort.

5.2.2 Illustrative Vignettes: Socio-spatial Formations

Following the modelling technique explained in Section 4.2.2.3, this section illustrated the technique with one vignettes as follows:

5.2.2.1 Data vignette 3. F-formations while stopping IV-antibiotic

a. Context and teachers' question. From an teachers' perspective, students are expected to follow official guidelines while preparing, administering or stopping medications. One of such guideline emphasises the need to perform these tasks, at least, in pairs, with one nurse monitoring what the other nurse is doing. For this, the data vignette in this subsection illustrates how positioning data could help teachers to confirm the following question: *How was the team physically arranged while stopping the IV-antibiotic? Were at least two nurses engaged in the stopping task?*

b. Analysis. The GCFF algorithm has been specifically developed to automatically identify f-formations from static images. For this reason, GCFF was applied to automatically identify overlaps in transactional segments of two or more people (presented in Subsection 4.2.2.3). The algorithm automatically identifies if nurses were in a f-formation and graphically represents a detected formation.

c. Illustrative insights. Figure 5.14 depicts the visual outputs from the GCFF algorithm of four teams (8 to 11) while stopping the IV fluid during phase 4. Team 8 did not exhibit any specific formation as nurses were at different sides of the bed while nurse 2 stopped the IV antibiotic. Although this does not necessarily suggest an ineffective teamwork behaviour, it can signal that one nurse was probably stopping the IV by herself, which can lead to errors and unexpected patient outcomes, particularly during procedures with the patient medication. With this information, teachers could provide informed feedback during debrief sessions and help nurses reflect on why they did not comply with specific guidelines.

In contrast, the visual outputs of teams 9, 10 and 11 show particular f-formations during the IV-fluid stopping task. Teams 9 and 11 exhibited the expected behaviour:

one nurse is providing the medication and the second is monitoring. Based on their body rotations, nurses were side-by-side (team 9) or assuming an L-shape (team 11) formation, which both enable manipulation of the IV device. For the case of team 10, the algorithm detected a side-by-side formation between the team leader and nurse 2, but other team members were in close proximity to them, suggesting they were also aware of their actions before and after stopping the IV-antibiotic. These insights can be useful for teachers not only to assess whether at least two nurses engaged in stopping the IV medication but also to visualise how nursing students approached the task and to discuss with them any challenges they may have faced.

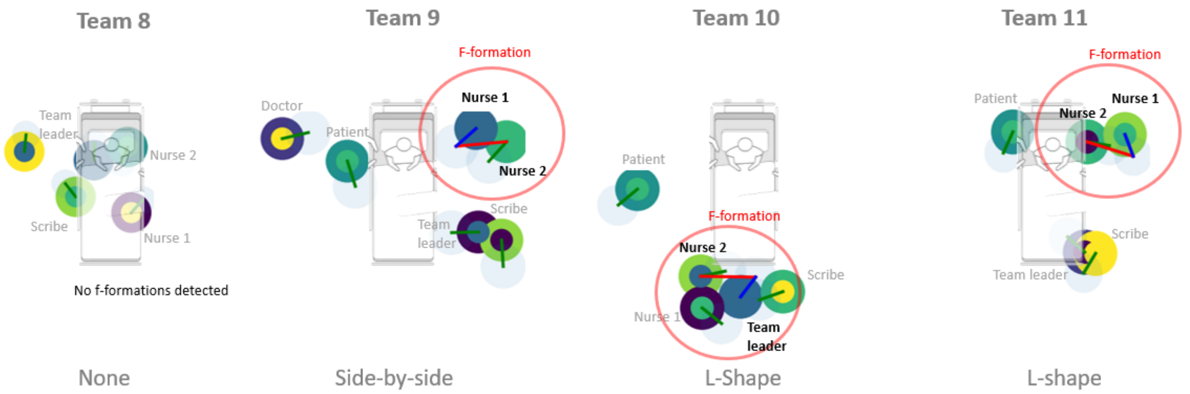


Figure 5.14: Detected formations in teams 9, 10 and 11 while stopping the IV-antibiotic.

5.2.3 Illustrative Vignettes: Fixed, Semi-fixed and Dynamic Spaces of Interest

Following the modelling technique explained in Section 4.2.2.4, this section illustrated the technique with one vignettes as follows:

5.2.3.1 Data vignette 4. How different teams used the spaces of interest.

a. Context and teachers' question. Here, we are interested in how nurses from different teams used the physical spaces in the classroom for the same team task. This problem is related to the following questions from an emic perspective: *Where did nurses spend most of their time during the simulation? Did nurses spend too much time at the medicine room? Were at least some team members near the patient during the event?*

b. Analysis. In this case, the interest is in giving meaning to individual team members' coordinates at a higher level of abstraction. SNA can be used to model interpersonal ties based on proximity data but it does not incorporate the particular places where nurses actually were. For this reason, we used Epistemic Network Analysis (ENA) to analyse nurses' presence in fixed, semi-fixed and dynamic spaces of interest (as described in Section 4.2.2.4). ENA is a novel method used for identifying connections among elements in coded data and for representing such connections through dynamic network models (Shaffer et al., 2016). ENA was originally conceived as a tool to quantify and model qualitatively coded discourse data. Yet, the method has recently been used by several data scientists to model other forms of coded data, such as social connections (e.g., (Wooldridge et al., 2018)) and digital tools usage (e.g., (Shaffer et al., 2009)), in various group settings. We believe that our work is the first to use ENA to model physical spaces.

The output of the modelling described in Section 4.2.2.4 was processed using an online ENA tool[3]. In the resulting epistemic networks, each node represents fixed, semi-fixed and dynamic spaces of interest (see Table 4.4), and each link represents the co-presence or transitions between two spaces of interest. In addition, networks are weighted: thicker and more saturated lines suggest stronger connections, whereas thinner, less saturated lines suggest weaker connections (Shaffer and Ruis, 2017). The positioning of nodes does not correspond to the actual positions of the spaces of interest on the floorplan. Instead, ENA automatically places the nodes in fixed positions to allow for meaningful comparison of patterns of connection between two or more team networks.

c. Illustrative insights. Figure 5.15 shows ENA diagrams for teams 7 and 9, mapping transitions between spaces of interest for the whole simulation.

Figure 5.15 (left) shows the epistemic network of a team transitioning between the bed footer and the patient manikin most of the time (see thick edge between nodes: *at the bed footer* and *at the patient manikin*). Moreover, this team received some help from the teacher mostly in these two semi-fixed spaces (see edges going to node *receiving help*). In contrast, although members of team 9 also remained very close to the patient manikin and at the bed footer (Figure 5.15, right), they occupied other meaningful spaces during the simulation. For instance, edges going to the nodes *near to the patient* and *close to human patient* suggest that team members also occupied the space further apart of each other but around the patient manikin, and close to the student role-playing the patient. This can be indicative of patient-centred care. Moreover, this team displayed a more

[3]http://www.epistemicnetwork.org/

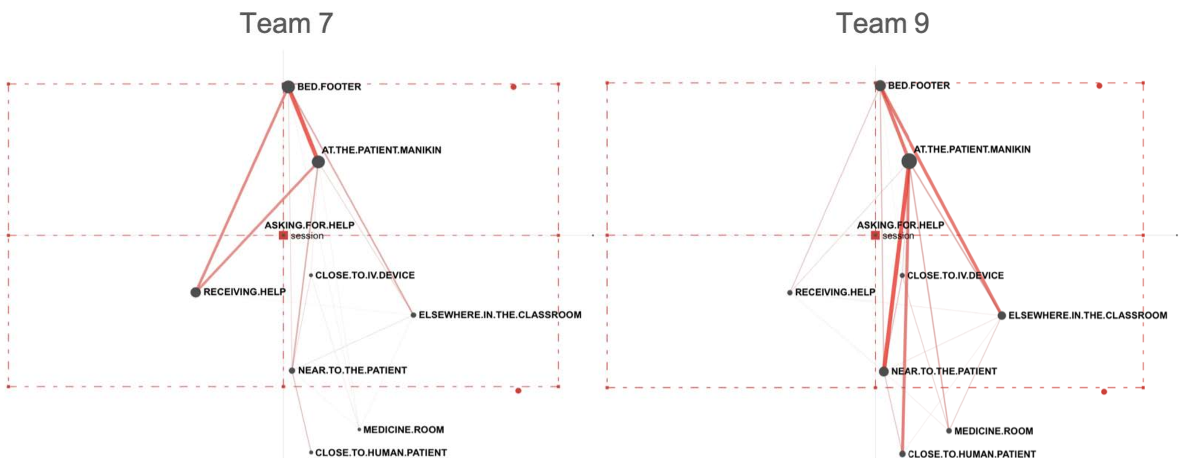


Figure 5.15: Epistemic Networks showing the spaces of interest. Team 7 (left) received help from the teacher, while Team 9 (right) was more independent.

independent and proactive behaviour, as team members commonly were elsewhere in the classroom and only received little help from the teacher (see thick edges connecting the node *elsewhere in the classroom* and thinner lines connecting the node *receiving help*).

5.3 Summary

5.3.1 Student-facing MMLA interfaces and Illustrative vignettes

Student-facing MMLA interfaces - Study 1. Students from simulations 1 (Sec. 3.2.1.1) and 2 (3.2.1.2) face and evaluate their multimodal data, using explanatory MMLA interfaces (described in Section 5.1). Three storytelling MMLA interfaces were created (one per simulation) to communicate team outcomes with regard to: i) team errors, which took account of sequence, frequency, and timeliness of actions, ii) nurses' arousal levels, which reported on students' arousal peaks, and iii) the rule-based algorithm used to generate the stories. The stories were designed using: a) the multimodal modelling outcomes (see Section 4.2); b) rules, extracted from the learning design; and DS principles (see Section 5.1 for additional details).

Students used the data stories to reflect upon: i) opportunities to improve, ii) specific skills and knowledge they need to master, iii) revise their strategies to tackle different situation in the future, iv) their accuracy of team/individual judgement about performance, and iv) their collaboration and teamwork strategies (Section 5.1.5). These results are aligned to Schon's notion of *reflection-on-action* (Schön, 1987), which explains that during the reflection practice individuals normally reflect on three aspects: i) recall and reflect upon the teamwork activity that has occurred, ii) think about what they should have done differently, and iii) think about new information gained as a result the reflection process. Thus, the results of this study (Study 1, Section 5.1) demonstrate the potential of data stories for helping students to identify misconceptions and errors, think about strategies they could use to address errors, detect aspects to improve in future scenarios, and reflect on their arousal levels.

Illustrative Vignettes about nurses' proxemic behaviour - Study 2. The four vignettes created in this study illustrate the modelling approaches presented in Section 4.2.2, using the data collected from Simulation 2 (Section 3.2.1.2). The vignettes associated with the three constructs described in sections 4.2.2.2, 4.2.2.3, and 4.2.2.4, with authentic teachers' questions (see 4.2.2.1), show how positioning data can help in telling data-informed stories about nurses' proxemic behaviours. The modelling outcomes were communicated by visualising: i) nurses' full-proximity networks and role-centred ego networks, using Social Network Analysis (SNA) (sec. 4.2.2.2); ii) nurses f-formations (sec. 4.2.2.3); and iii) Epistemic Network Analysis representations.

The vignettes provided exemplar insights from the teams who participated in simulation 2 (see Section 3.2.1.2). For instance, the use of patient-centred ego network helps students and teachers identify the presence of team members in the vicinity of the patient's interactional space. This insight helps team members to explore how well the patient was cared for, a key insight as patient should be central in the simulation practice. Similarly, full social networks can model proximity ties among nurses, indicating whether nurses were working together appropriately during the simulation (e.g., during the delegation phase, nurses are expected to be close to the team leader). F-formations, were used to identify if the nursing teams followed the official guidelines Australian Nursing & Midwifery Federation (2020) when preparing, administering or stopping medication. ENA representations provided meaning to individual team members' coordinates by modelling interpersonal ties based on proximity, as well as the particular locations adopted by individuals during the simulation. These data vignettes have illustrated the potential for automatically generating evidence about positioning strategies to address authentic questions that teachers have when monitoring, assessing and reflecting on nursing team simulations. Although these vignettes were not evaluated with students they could in future work enable the provision of automated feedback which could be used to spark reflection in a post-simulation debrief.

CHAPTER 6

EVALUATION OF MMLA INTERFACES WITH TEACHERS

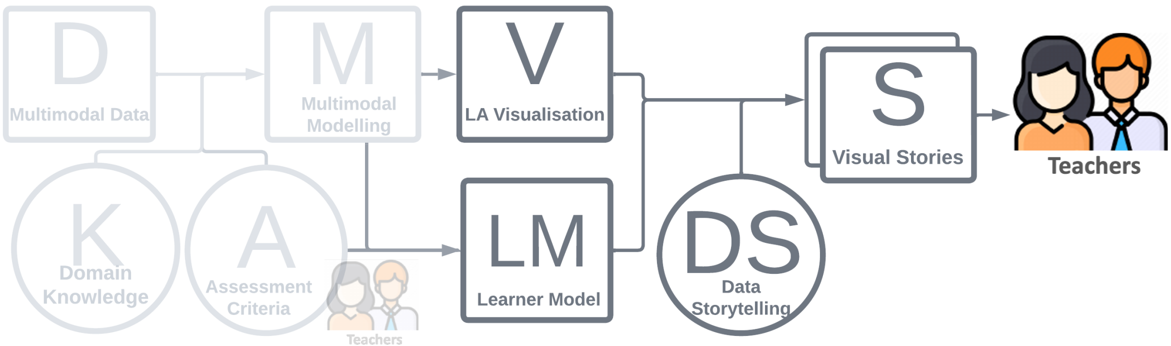


Figure 6.1: Crafting and evaluating MMLA interfaces with teachers

The third and fourth contributions of this thesis are described in detail in this chapter. This chapter focuses on the second component (Figure 5.1) of the conceptual model introduced in Chapter 3. Thus, the chapter presents a set of **multimodal visual interfaces** *(V-S)* designed based on the modelling approach described in Chapter 4. The framework enables the generation of two types of interfaces; interfaces that invite users to **explore** *(V)*, and interfaces that **explain** or communicate insights *(S)*. The framework introduces the learner model *(LM)* and data storytelling principles *(DS)* to enable overlay additional information or enhancement to the visualisation to guide interpretation. Section 5.1 presents the teachers' exploration of two explanatory MMLA interfaces. Section 5.2 provide four illustrative vignettes (exploratory MMLA interfaces) and exemplar insights about team dynamics from an evaluation with teachers.

Each of the studies presented in the following sections were designed with various protocols and research objectives with the intention to gain additional insights and perspectives from the evaluation. Figure 6.2, explains how studies 3 (Section 6.1), 4 Section 6.2), and 5 (Section 6.3) are addressing the main research questions of this thesis (Section 1.2). The identification of research objectives derived from the different qualitative studies are defined based on the study number (St3-St5) and the research objective number (Ro).

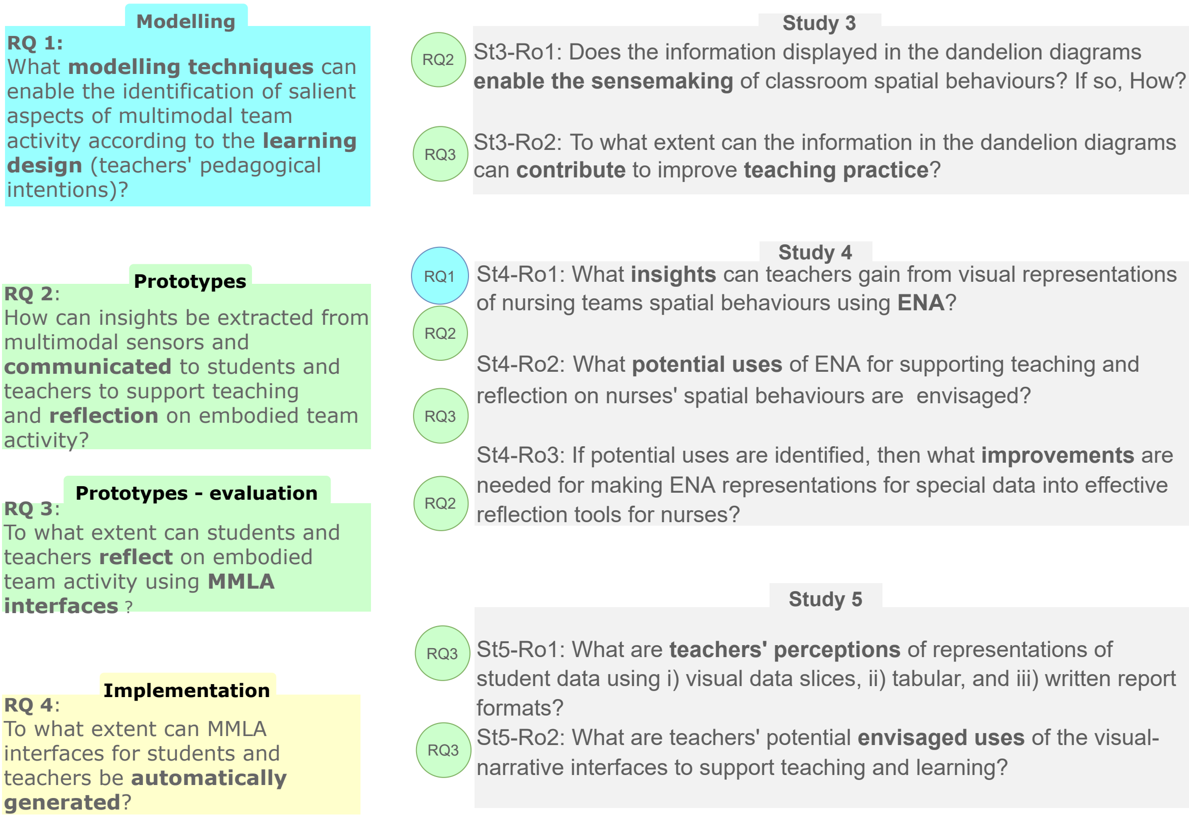


Figure 6.2: Mapping of thesis RQ and the objectives of studies 3-5

6.1 Study 3. Classroom Dandelion

The dandelion diagram is a positioning visualisation technique that aggregates positioning and orientation data, and thereby can depict both the whereabouts and heading directions of a person at different points in time. Dandelion diagrams also visualise the trajectory of people across the physical space, which can reveal their mobility patterns. Moreover, it uses color codes to communicate extra information, such as to differentiate the data points representing multiple actors in a collaborative process. The visualisation is agnostic to the underlying positioning tracking system. The input data for dandelion diagrams should be formatted continually in equal timeframes (e.g., 1 data point per second as described below). Each data point should include both *location* (i.e., x-y cartesian coordinates), and *orientation* (in radians or degrees)[1]. The dandelion diagrams feature the following four visualisation design components:

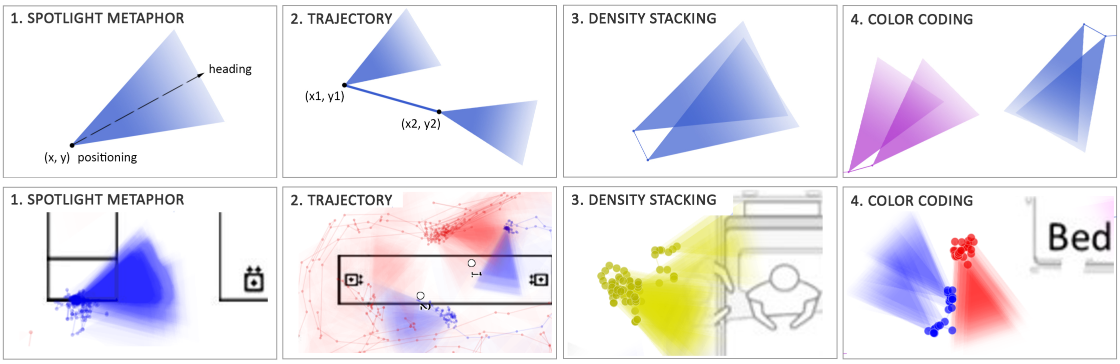


Figure 6.3: Top: Illustrations of four major design components of the Dandelion Diagram. Bottom: Embodiment of these design components in data visualisations using data from the two learning contexts targetted in this thesis.

1. *Spotlight metaphor.* The position and orientation of an actor in a given point of time are represented in a triangular shape, following the metaphor of "spotlight" (see Figure 6.3, 1). The farthest vertex indicates the x-y location of the actor, and its opposite side shows the direction the actor is facing to. This representation is commonly used in video games and navigation systems, and can intuitively depict an entity's location and orientation.

2. *Trajectory.* The trajectories of an actor are delineated by connecting every two consecutive position coordinates, which shows any change in location and orientation over the visualised period of time (see Figure 6.3, 2).

3. *Density stacking.* Based on the metaphor of the heatmap, each spotlight shape is semi-transparent, and multiple shapes can stack on one another to increase the

[1]This section content is based on the work presented in: **Gloria Milena Fernandez-Nieto**, Pengcheng An, Jian Zhao, Simon Buckingham Shum, and Roberto Martinez-Maldonado. 2022. Classroom Dandelions: Visualising Participant Position, Trajectory and Body Orientation Augments Teachers' Sensemaking. In CHI Conference on Human Factors in Computing Systems (CHI'22), April 29-May 5, 2022, New Orleans, LA, USA. ACM, New York, NY, USA, 17 pages, doi: 10.1145/3491102.3517736. 2022.

density of the colour, which reflects that the actor was cumulatively staying in a location (see Figure 6.3, 3).

4. *Colour coding.* The different colours of the spotlight shapes represent different actors in the practice context, to help the audience differentiate multiple roles in a collaborative process (Figure 6.3, 4). Alternatively, this colour coding feature can also be used to represent other types of information (e.g., in a prior iteration An et al. (2020), it was used to differentiate data points from different activities or periods of time).

6.1.1 Participants

The study presented in this section focuses on the perspectives of teachers involved in the design, delivery or evaluation of the two units of study described in Section 3.2. Two learning contexts used to evaluate the Classroom Dandelions are described in sections 3.2.2.1 and 3.2.1.2. In this study **Context A** make reference to the science laboratory 1 (see 3.2.2.1) and **Context B** to Simulation 2 (see 3.2.1.2).

For Context A, the unit coordinator (physics teacher 1 - PT1 - who designed the learning tasks and did not teach any class); a main teacher and a secondary teacher (PT2 and PT3, respectively, who taught a total of 14 classes in pairs) participated in the study. All of the participants were experienced teachers in each of their roles (males: 3, average years teaching: 11.3). For Context B, 5 nursing teachers (NT1-5) inspected students' positioning data (females: 4, average years teaching: 12.6). NT1-4 delivered the 11 simulation classes and NT5 is a nursing researcher who assesses nursing education programs at the hosting university.

6.1.2 Protocol

Inspired by theoretical foundations of information visualisation design (Meyer and Dykes, 2019; Sedlmair et al., 2012), and sensemaking with data visualisations (Campos et al., 2021; Card, 2004; Pirolli and Card, 2005), I used a set of dandelion diagrams as *evolving MMLA interfaces* to investigate the relationship between teachers and the positioning data, beyond learning about the specific visualisation technique. For this, the eight teachers in both learning contexts were interviewed with the purpose of documenting their perceptions of the dandelion diagrams about themselves (in Context A) and their students (in Context B). Each interview was recorded using an online video conferencing platform (i.e., Zoom) and had an approximate duration of 60 minutes. Before each interview, the four major design components of the dandelion diagrams were explained to the teachers using examples such as the ones depicted in Figure 6.3. They were also invited to ask clarification questions regarding the design of the visualisations during the rest of the interview. Following a semi-structured format, the interview had two parts:

Part 1. A think-aloud protocol was followed to document how teachers explored the positioning data in order to address research objective (Ro) 1 (sensemaking of spatial behaviours). These data were presented to teachers as digital indoor maps corresponding to *critical classroom events* of the classes relevant to them in two ways: i) by visualising x-y coordinates only, and ii) using dandelion diagrams. The x-y coordinates were presented in the form of regular heatmaps as in previous classroom studies reported in Section 2.3.3. The same data was made available through the dandelion diagrams. The maps were generated by normalising the positioning data to 1Hz to make them comparable.

The *critical classroom events* presented to teachers were selected based on previous work. For Context A, results from a previous qualitative study Martinez-Maldonado et al. (2020b) emphasised the importance of understanding how teachers interact with (i) classroom resources, (ii) students and (iii) each other in co-teaching scenarios, which is aligned to contemporary literature on the materiality of the classroom environment (Yeoman and Wilson, 2019). Each class was segmented into three phases by the unit coordinator. Phase 1 includes the main teacher of the class giving instructions (average duration 13 average=8 minutes). Phase 2 corresponds to the period in which all students start working on the experiment(s) in small teams (1.5 hours average=18 min). Phase 3 corresponds to the time when some teams complete their experiments and start leaving the class (33 average=22 min). The analysis of this paper focuses on Phase 2, which enables comparison across the classes considered. Phase 2 was further segmented into quartiles of the same duration. A total of five critical events were selected by the research team from the resulting quartiles, displaying events identified by the teachers who participated in the previous study as examples of potential interactions of teachers *with classroom resources* (2 instances), *with students* (one instance) and *between themselves* (2 instances). Details of these five critical events, and the extent to which each is representative of the dataset, are provided in the next section.

For Context B, the critical events were closely related to the learning design of the team activity created by the unit coordinator. High effective teams should go thrugh the following events: (i) perform an initial set of vital signs measurements; (ii) prepare/administer the intravenous fluid-IV antibiotics; (iii) perform another set of vital signs measurements after the patient complains of chest tightness; (iv) stop the IV antibiotic; (v) perform an ECG; and (vi) call the doctor after stopping the IV antibiotic. In this case, the research team randomly selected instances from this critical moments, focusing on events iv and v, which are the most critical events identified by the unit coordinator. All the teachers within each context inspected the same set of visualisations.

Part 2. Then, teachers were asked to respond to two main questions (and trigger sub-questions) to elicit their perceptions of the dandelion diagrams in terms of sensemaking (Ro1) and potential to support teaching practice (Ro2).

1. **St3-Ro1:** Does the information displayed in the dandelion diagrams enable the *sensemaking* of classroom spatial behaviours? If so, How? Trigger questions:
    - What differences does it make to add rotation/trajectory information to the positioning data maps, if any?
    - Could you name some examples from your experiences in which the heading information can help us identify what was happening in the classroom from the positioning data?
    - Can you envisage any risk in showing the diagrams with or without body orientation information?
2. **St3-Ro2:** To what extent can the information in the dandelion diagrams can contribute to improve *teaching practice*?
    - Do you think such visualisations can support or hinder teachers' reflections upon their practice? Why?
    - Do you think such dandelion diagrams could also support professionalisation of researchers/experts in doing their research or training teachers? Why and How?
    - How do you envisage teachers can use the diagrams to support reflection and teaching over time?

6.1.3 Analysis

The interviews were fully transcribed using a professional service. In qualitative studies, notions of generalisability and reliability are usually replaced by validity, rigour and attention to the quality in the research process (Creswell, 2014). To achieve this, I used data triangulation (various sessions with teachers in two contexts), triangulation of sources of evidence (think-aloud recordings, video data, and interview responses), and analysis triangulation (three researchers in the analysis process). Informed by best practices of qualitative research in HCI (McDonald et al., 2019), I analysed the interviews as follows.

Analysis of Part 1: Vignette analysis. This involved three researchers independently screening the video recordings of the interviews looking for how teachers talked about spatial behaviours based on the visualisations of x-y coordinates and the dandelion diagrams. Then the researchers discussed each particular moment, selecting vignettes that could potentially point to insights and contradictions across teachers inspecting the same visualisations. The researchers discussed their independent analyses to reach an agreement. Results from this analysis are reported in the next section by describing the critical classroom event, and a summary of teachers' perceptions on them. Footage from the classroom sessions was additionally used to confirm teachers' interpretations of the diagrams they explored.

Analysis of Part 2: Interview analysis. Given the direct alignment between the interview protocol and the research questions, statements of interest were jointly identified by three researchers. These were thematically coded (Braun and Clarke, 2012) by one researcher. Resulting coded statements were double-coded by other two researchers until full agreement was reached. Next, researchers had several discussions to identify and group emerging themes in alignment to the interview questions, which are reported in Section 6.1.6.

6.1.4 Vignettes from the co-teaching context (A)

6.1.4.1 Vignette A1: Proximity to Classroom Resources

This vignette focuses on two *critical classroom events* in which teachers were in close proximity to key classroom resources. Previous studies looking at classroom proximity traces have assumed that the teacher may be interacting with classroom resources if they are within arm's reach (Martinez-Maldonado et al., 2020b) (e.g., while giving an explanation to students). The two classroom maps on the left of Figure 6.4, illustrate two instances where the secondary teacher (red data points) is very close to the classroom whiteboard (see Points A and B) and the main teacher (in blue) is at a bench where there are no students (Point C) for an approximate period of 22.5 minutes.

When teachers explored the visualisations of *only x-y coordinates* (Figure 6.4, left), they confirmed the assumption is that close proximity to a classroom resource most

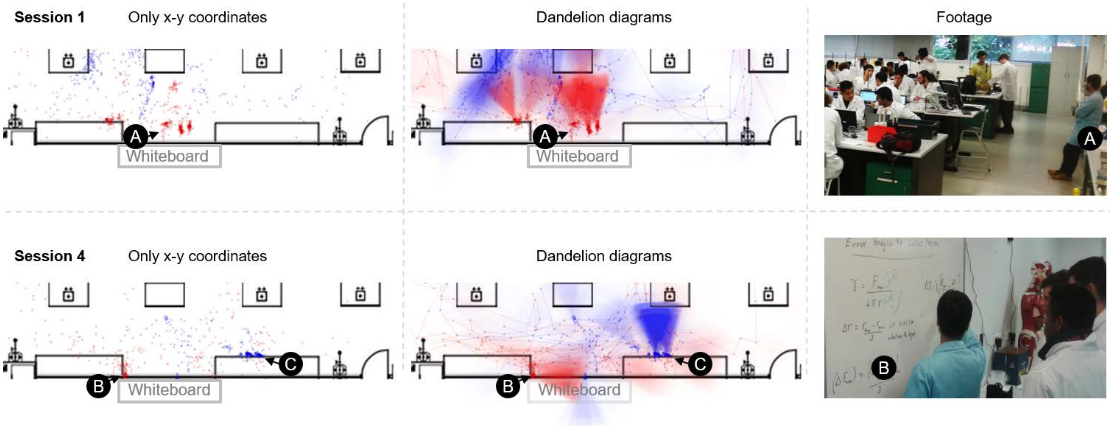


Figure 6.4: Visualisations over 22.5 mins. approximately, of the main teacher (blue) and the secondary teacher (red) in close proximity to the classroom whiteboard (see points A and B) and an auxiliary bench (C). Left: *x-y coordinates only*. Centre: *dandelion diagrams*. Right: *video footage* of the critical events.

commonly means that teachers may be using it. The unit coordinator (PT1) explained this as follows *"It is unusual to stand at the whiteboard not using the whiteboard unless there are students standing in front of him or her"*. The main teacher who was in the room in both sessions (PT2) also assumed that the auxiliary teacher in red (PT3) was *"probably explaining some theoretical points, using the whiteboard"*. Yet, PT3 confirmed that he *was scanning all the tables to see if all students were maintaining self regulation"* in session 1 (top), and *"just standing there and seeing if anyone needed help"* in session 4 (bottom). This teacher also explained that for Point C, that teacher was *"most likely using the bench to write on the laboratory sheet, indicating who did which experiment, and who attended"*.

However, when the same teachers looked at the dandelion diagrams (Figure 6.4, centre), all of them changed their minds about the spatial behaviours in both situations. For example, PT1 indicated that PT3 spent most of the time supervising and did not consider this behaviour adequate if extended for a long period of time (see Point A): *"the [red] teacher seems to be just watching the crowd. This should not really happen unless you use a whiteboard, which obviously is not the case"*. This was confirmed through a snapshot of the classroom footage (see Figure 6.4, top-right). Regarding the blue dandelion at Point C (Figure 6.4, bottom- centre), he suggested a particular behaviour where the 'spotlight' remains at a specific angle as follows: *"it's not like a swiveling over like watching people. So I suspect that teacher in blue is either idle or explaining something to students"*. PT2, after looking at his own data, explained: *"It seems I was just deeply looking to the class, because there are dense dots. It is near the bench, so I was probably explaining something* *to some group of students there gathered around me"*. PT3 confirmed that in Session 1 (Point A) he was *"facing the students, changing [his] orientation from left to right, trying to scan the whole classroom to see if there were any problems or safety regulations that needed attention"*. Yet, he corrected himself as he could see through the dandelion diagram that his body orientation in Point B was towards the whiteboard, explaining this as follows: *"I was writing on the whiteboard or explaining a formula written on the whiteboard to some student"* (see Figure 6.4, bottom-right).

In sum, the added information contained in the dandelion diagrams enabled teachers to understand the situation more in detail. This also demonstrates that it is not sufficient to simply use the proximity distance as a proxy to identify if certain classroom resources are being used by the teachers. Importantly, teachers could differentiate important spatial behaviours, such as scanning the classroom when the spotlight is *swiveling* versus the dense stacking indicating that the body was oriented towards a fixed point for longer periods of time; these look identical in an x-y heatmap. The next vignette focuses on spatial behaviours when teachers are close to students.

6.1.4.2 Vignette A2: Proximity to Students

This vignette is focused on a *critical classroom event* that is representative of 20-40 instances that occurred in each class (avg=30; std= 10.3): teachers standing in close proximity (within 1 metre) to teams of students working on their experiments. An instance was selected based on a previous study Martinez-Maldonado et al. (2020b) in which 3 clusters of teachers' positioning data points are surrounded by 4 students' experimental setups. Figure 6.5 depicts the selected instance visualised by plotting x-y coordinates (left) and dandelion diagrams (right). Teachers were asked to comment on what they thought PT3 (in red) was doing in close proximity to team 4 (Point X) and what PT2 (in blue) was doing near team 7 (clusters in Points Y and Z) for a period of 23 minutes approximately. This instance was selected because the body orientation data in Figure 6.5 (right) appears to help disambiguate which team of students is being attended to (i.e.„ although the teacher at Point Y is closer to team 7, his attention was in fact on team 6 working further away on the laboratory bench).

Similar to the vignette A1, when teachers explored the x-y data plot (Figure 6.5, left), they confirmed that the assumption is that being in close proximity to students means that they may be interacting with that team. PT1 explicitly explained this, by considering the details of the learning design (i.e., *"Because this is a very concentrated experiment, you have to be in close proximity to the students to talk to them"*), as follows:

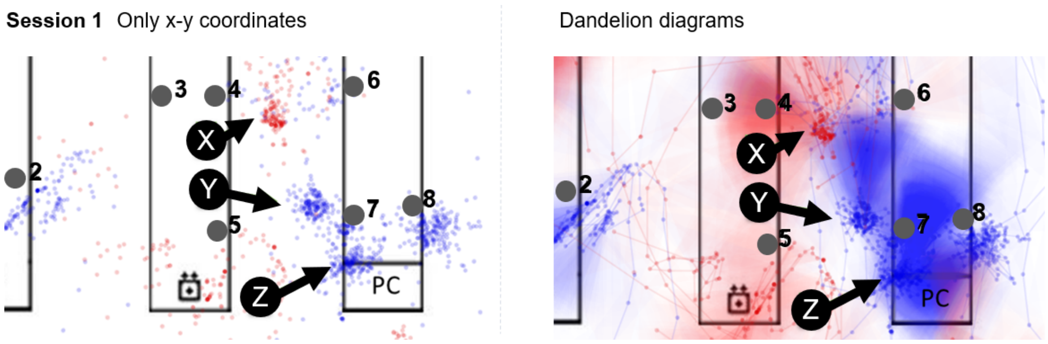


Figure 6.5: Visualisations over approximately 23 mins. of the main teacher (blue) and the secondary teacher (red) in close proximity to students' laboratory experiments (see points X, Y and Z). Left: *x-y coordinates only*. Right: *dandelion diagrams*.

"Just looking at the obvious proximity, we have teams 4 and 6 for Point X. For Point Y, we have team 7, and 5 possibly. This is the principal teacher so for Point Z may also be close to the PC checking information.". The other two teachers (whose data was represented in the diagrams) followed the same rationale (e.g., PT2 stated: *"At Point X I was attending team 4 because I was closer to them, and so on."*.

Yet, when teachers explored the dandelion diagrams (Figure 6.5, right), they had to revisit how they initially interpreted the data. For example, PT3 (in red) explained how the information about his body orientation and that of the other teacher, could help him confirm how they used the space to support students, as follows; *"From position X I was facing team 4 so I was like helping them. From Y, maybe [the other teacher] was [walking] from [position Z] and maybe was monitoring team 6 from afar"*. PT2 (in blue) also corrected himself and could describe the situation by looking at the dandelion diagrams representing his own data: *"So I said in X I was attending number 7. But my conclusion was wrong because I was oriented towards group 6. But it is too far away from this group unless there is some student from this team maybe asking some question. Looks like definitely from point Z I was explaining something to team 7. Awesome!"*

In sum, the added body orientation information enabled teachers to get a richer picture of the classroom situation they experienced. The next vignette explores two cases in which both teachers were providing close attention to only one team at a time, which I refer to as instances of co-teaching.

6.1.4.3 Vignette A3: Instances of Co-teaching and F-formations

This vignette is focused on two *critical classroom events* that are representative of the 3-10 instances that occurred in each class (avg=5.6; std= 4): both teachers standing in close proximity (within 1 metre) to the same team partially at the same time. Two instances were selected to illustrate two different f-formations displayed by the spatial behaviours of the teachers approximately 22.5 minutes apart. The classroom maps at the top of Figure 6.6 illustrate an instance of a face-to-face f-formation of both teachers near team 1, and the maps at the bottom illustrate a side-by-side formation Kendon (1976), where both teachers faced team 2. Both moments were extracted from the same class session to enable comparison.

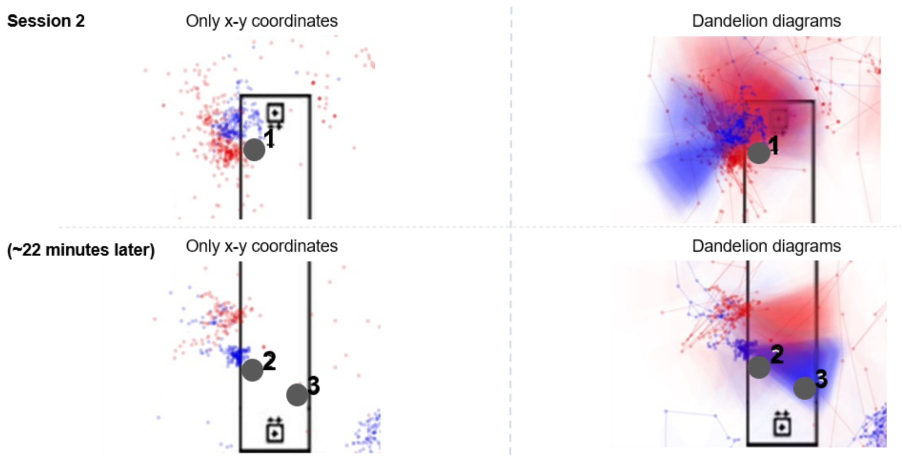


Figure 6.6: Visualisations approximately 22.5 mins. apart, of the main teacher (blue) and the secondary teacher (red) while both attending to the same team of students. Left: *x-y coordinates only*. Right: *dandelion diagrams face-to-face* (above) and *side-by-side* (bottom).

When teachers inspected the maps showing only x-y coordinates, the teachers did not have much evidence to reflect on. Both maps in Figure 6.6 (left) look similar. PT1 wondered if *"maybe the (red) teacher sort of sought some extra help from the (blue) teacher to support team 1"*. The other two teachers explained that having two teachers in close proximity to the same team may be a sign of students needing to *"solve a problem"* (PT2) or to *"debug the experiment or explaining to students what they needed to do"* (PT3). Yet, teachers did not explain much about their own team dynamics as a teaching team when attending the same group of students.

When looking at the dandelion diagrams (Figure 6.6, right), for both cases, they all reflected about the strategies that could potentially be illustrated through these examples in terms of socio-spatial behaviours. For example, PT1 explained that, in the case of the *face-to-face* f-formation (Figure 6.6, top-right), *"even though they may have attended the same team a few minutes apart, what it actually shows is that team 1 requires quite some attention because there's a fair concentration of both teachers, the red* *and the blue."* PT1 also explained that the body orientation gave him clues about the kinds of interactions that may have occurred: *"If the blue is talking to you, or the other way around, they would be facing each other. The main teacher might as well talk to one of the students who is engaged in that experiment because is facing to the other direction. The red one is engaged with these students at the bench. You see? The red is much, much more concentrated"*.

In the case of the *side-by-side* f-formation (Figure 6.6, bottom-right), the dandelion diagrams helped PT3 to recall exactly what happened, and he reflected on his collaboration strategies as follows: *"So, the clear difference is that in the [face-to-face formation] the main teacher commonly talks to the [secondary] teacher; and in the [side-by-side formation] we usually are giving instructions to students"*. PT3 also explained in more detail how the information about their body orientation helped him reflect on their social dynamics. He explained: *"face-to-face situations would be when we are both having a discussion so that we can both work on the same project. But the side-by-side was the situation when one tutor was actually explaining to the group and the other is standing there. So, you see? like the fans [spotlights] actually intersect"*.

The next set of vignettes illustrate new aspects yet to be discussed related to social dynamics, since it focuses on the Nursing teams of 5 students engaging with their teacher.

6.1.5 Vignettes from the teamwork context (B)

6.1.5.1 Vignette B1: Teacher Monitoring or Intervening

This vignette is focused on a *critical event* when one of the teams of nurses is attaching the 3-lead electrocardiogram device (ECG) to the simulated patient. During this procedure, nurses around the bed attach the ECG directly to the patient's chest. This vignette illustrates nurses' embodiment of this task in the classroom space for a period of one minute (one point per second). This instance was selected because the teacher was in close proximity to the team (see grey data points in Figure 6.7).

When nursing teachers (NT1-5) where asked to explain the situation using Figure 6.7 (left), four out of five interpreted that the role of the teacher was actively involved in helping and guiding nurses through the process (performing an *intervention*, for example, to provide feedback to the students) by focusing on Point A in the figure. For instance, NT3 mentioned that *"in this situation [nurses] commonly don't realise what is going to happen so they may need a little bit of guidance and reassurance from the teacher"*. NT5

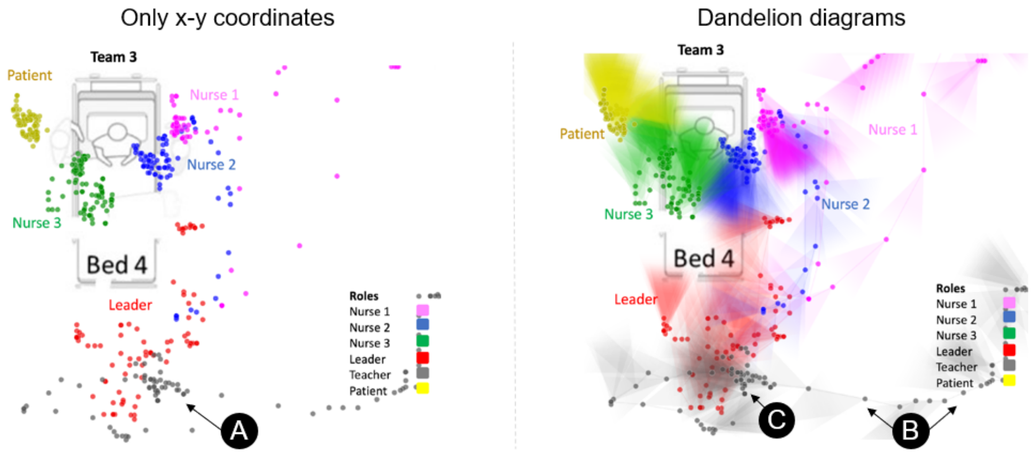


Figure 6.7: Visualisations over one minute of a teacher, and four student nurses playing different roles in a team simulation during a critical event: attaching an ECG device. Left: *x-y coordinates only*. Right: *dandelion diagrams.*

more strongly assumed that *"the teacher must have been explaining something specific or nurses asked for some kind of clarification"*. NT2 (whose data was being inspected) admitted that he could not remember what may have been happening without having further information.

However, when presented with Figure 6.7 (right), all the teachers changed their initial interpretations. Three out of five teachers pointed at the teacher's trajectory (Point B), which was interpreted as the teacher just *monitoring* the team and then moving away to the next bed. NT1 indicated *"I can see the teacher was moving away to the next bed [pointing at the trajectory]"*. NT2 this time commented *"So you can see that I was standing off to the side just looking, and I might have come over to briefly say something to the team leader [referring to Point C]. I didn't invade their space for some particular reason"*. This monitoring behaviour was confirmed through the video footage.

This vignette illustrates that visualising the trajectories and the body orientation of the teacher and students contributed to characterising how the teacher's classroom tasks (e.g., monitoring versus intervening students) were enacted. This may be helpful to train novice teachers for them to learn how to work in the classroom space.

6.1.5.2 Vignette B2: Spatial Team Dynamics

This vignette focuses on the spatial behaviours of two teams (3 and 5) during the most critical event of their team simulation: noticing the patient's adverse reaction to an antibiotic and stopping its intravenous (IV) administration (located at the top right hand side of the bed in all diagrams in Figure 6.8). In the simulation illustrated in this vignette, there is commonly one student (Nurse 3 in Team 3 and the team leader in Team 5) asking questions to the student enacting the voice of the patient (student sitting next to the bed whose data is depicted in yellow). According to the teachers, the spatial behaviours of Team 3 were more effective compared to Team 5.

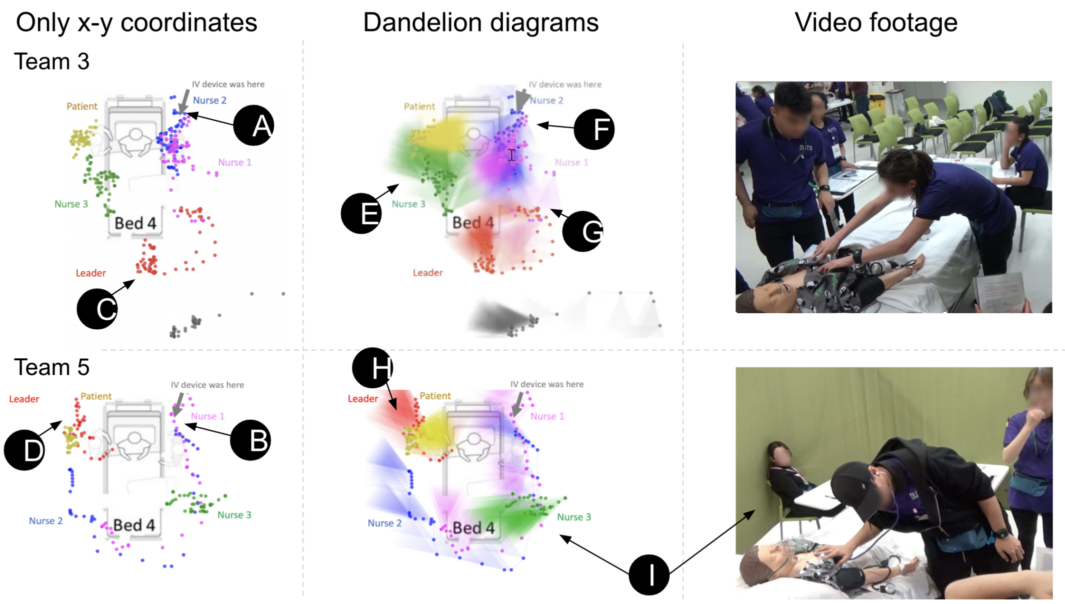


Figure 6.8: Visualisations over one minute of teams 3 (top) and 5 (bottom). Left: x-y coordinates only. Centre: dandelion diagrams. Right: footage of the critical event: stopping the IV fluid on time.

Based on the x-y coordinates (Figure 6.8, left), teachers did not perceive major differences between both teams. NT1, NT2, NT4 and NT5 explained, using similar words, that in both teams Nurses 1 and 2 performed the critical actions. Regarding Team 3, NT2 explained: *"...so, nurse 1 and nurse 2 have identified the incident. They realised that the adverse event probably is related to the antibiotic being delivered and therefore they are around the IV fluids"* (Point A in Figure 6.8, top-left). Similarly, NT5 explained the following for Team 5: *"I think that nurses 1 and 2 took a very active approach. And probably nurse 1 was who was stopping the IV fluid (signalling Point B)"*. The teachers also inferred some differences between the behaviours of students across both teams (e.g., asking questions to the patient or checking the IV machine) by noticing which students were close to each other or to the patient. For example, the team leaders in both groups behaved differently during the critical event. The team leader in Team 3 was effectively positioned according to her role. This was confirmed by NT4 as follows: *"It looks like the team leader of team 3 is at the base of the bed (Point C) from where she* *can notice if something is wrong with the patient"*. In contrast, in Team 5 the team leader was standing at an unexpected position during the adverse event: *"Oh look! [signalling Point D], it looks like the team leader is in a different position here, because usually they are at the end of the bed. So it looks like she/he is actually asking the patient"* - NT1.

In contrast, when teachers inspected the dandelion diagrams (Figure 6.8, centre), they identified certain spatial behaviours according to nurses' roles. For the case of Team 3 (top-centre), teachers used the body orientation of the students as a potential marker of interaction with the patient or with other team members. For example, NT4 said: *"everybody was facing the patient"*. NT1 further interpreted the information about body orientation in terms of team dynamics, as follows: *"You can see nurse 3 is focused on the patient (see Point E) and you can see nurses 1 and 2 have gone and stopped the IV antibiotic (Point F)"*. NT5 added: *"it is clearer here that the leader had a kind of communication with nurse 1" (signalling Point G)*. However, for the case of Team 5 (bottom-centre), teachers highlighted some potential areas for improvement. For example, although NT1 effectively recognised team leader's behaviour from his spatial traces (i.e., *"I can see here that the team leader may be doing the patient's assessment (signalling Point H")*, NT2 explained this may be related to some problems faced by this team as follows: *"I'm not happy with the team, I think the team leader should not be doing the procedure. It should be [nurse 1] or another nurse allocated to this task"*. Similarly, NT4 explained this in terms of a leadership problem, as follows: *"...the team leader is doing everything and not delegating. Nurse 3 has done nothing (see Point I). Nurses 2 and 3 seem to be waiting for direction and taking no initiative"*.

In sum, with the x-y heatmaps, teachers could not differentiate the two teams, but with the dandelion diagrams, they were able to identify key behaviours related to the *effective* execution of the team task (by Team 3), in comparison to a team showing potential *leadership and communication* issues (Team 5). The next section presents additional insights from the rest of the teacher interviews.

6.1.6 Thematic results

This subsection presents reflections externalised by teachers during the interviews. Questions asked were organised according to our two research questions under the themes: added value of the dandelion diagrams (Ro1) and envisaged integration into teaching practice (Ro2). A third theme emerged concerning potential challenges, risks and limitations.

6.1.6.1 Added Value of the Dandelion Diagrams

Two sub-themes in this regard were identified:

Advantages over representations of x-y coordinates. The first subtheme that emerged addresses the benefits of incorporating the representation of body orientation in the visualisation, in comparison to existing tools that only depict coordinates, e.g., a heatmap, or point cloud visualisation. As NT2 experienced, when there were only dots (x-y data) on the visualisation, the audience would not have complete insights into the situation: *"you're only going to see part of the information, you might not see everything."* In contrast, with the dandelion diagrams, *"when you see more of that directional information you can then put together what is actually happening"*. As stated by PT2, *"if only the dots [are shown], it just says that you spent more time in a certain location. But when you add a cone (body orientation) you will see exactly where the teacher was facing to and you could attach some explanation to those dots"*. Correspondingly, the teachers provided examples about how the dots-only visualisation can lead to more misinterpretations of the situation. For instance, PT3 explained: *"Teachers may not be facing the group at all. They may be facing the other way. So it may be easy to interpret that they were actually helping the group because they were somewhat close to them, but it was something else in reality"*. More examples could also be found in the vignette review sections (i.e., Subsection 6.1.4.2). Hence, as NT4 explained, *"the points [alone] can be misleading when you try to read them"*. A similar perception was shared by other teachers. This recurrent opinion confirmed that one major added value of the dandelion diagram is the aggregation of x-y position and body orientation, which help teachers to better interpret the classroom situations at critical moments.

Potential complement to video analysis. Another emerging subtheme concerns how the dandelion diagrams can serve as a complement or alternative to video analysis, to save teachers' time and efforts in professional training. In current contexts, video has been widely adopted as a learning tool to promote practitioners' reflection on their performance. For example, in a training session, teachers might need to watch and analyse classroom recordings of peers or their own, and discuss different aspects of practice. A similar approach has also been followed in the training of nurses. Despite being effective, this video analysis approach can be time-consuming for practitioners who already have busy routines. As experienced by our teachers, they suggested that the dandelion diagrams may serve as objective summaries or episodic indexes to help professionals intuitively analyse the depicted incidents, or quickly retrieve a relevant video clip. As NT5 explained, *"for me as a teacher, it will save a lot of time, because I* *won't need to go to the video as a source of information. I can just use this visualisation to explain what is going on, and I know that it will be accurate."* Correspondingly, another similar explanation was offered by PT2: *"it can work as a kind of summary for reflection because I don't think any teacher is going to go through the video recording of one hour or two hours to see how to do the teaching."* In addition, NT2 pinpointed another potential advantage of dandelion diagram as an alternative to the video approach, *"I think it gets away from the whole issue around privacy."*

6.1.6.2 Envisaged Integration into Teaching Practice

Under this theme I summarise the teachers' opinions on how the dandelion diagrams could be integrated into existing practice of professional learning and teaching. The first three subthemes address what the dandelion diagram could be potentially used for to facilitate the process of professionalisation, whereas the last subtheme summarises when and where professionals might use the visualisation as a tool.

Enabling evidence-based reflection. All teachers in both contexts believed that dandelion diagrams can support practitioners' reflection both individually and collectively, by summarising evidence from their own practice. For example, in the context of the physics labs, PT2 believed that the dandelion diagrams could serve as a reflection of how teachers have helped their students and motivate them to *"think about what needs to be changed before the next class"*. PT3 said that this visualisation tool, as a more objective data portrayal, could help them identify *"unconscious bias"* in their spatial performance and thereby *"promote learning better"*. In the nurse training context, NT2 believed that the visualisations could be used not only for *"self-reflection"*, but also for *"group reflection"*, in which a team of nursing students could reflect on *"what worked well and what didn't work well and what was going on here and what would they should do differently the next time"*. Similarly, NT5 envisaged that the dandelion diagrams could offer students clues to help them identify surprising or unwanted patterns, which provoked reflective questions such as *"what happened to me in this situation? Why? How can I improve it? What can I do to avoid it in the next simulation?"*.

Facilitating example-based learning. The teachers also highlighted the promises of dandelion diagrams in helping professionals learn from others' (good or bad) practices. For instance, NT2 considered them potentially beneficial in professional learning, if the visualisation technique is used to curate and compare examples of *'good'* and *'bad performance'*: *"here's some examples about [...] really good interaction using this visualisation, compared to another example where someone was facing in the opposite direction or away* *from the patient most of the time"*. NT3 proposed using anonymised visualisations for peer case analyses, e.g., to help nurse trainees to discuss opportunities for improvement from others' cases. Moreover, NT4 believed that the dandelion diagram visualisations could serve as concrete demonstrations for novice professionals to understand ideal practice patterns in certain situations: *"that is a good way to sort of introduce new teachers [. . .] what pattern we think is probably more suitable or the best to use in those various circumstances"*. She gave an example about how this would work better than only providing abstract instructions to novice professionals: *"the visualisation helps, certainly, because just talking about roaming around [is not enough]. What does roaming around mean? And, you know, the dandelion diagram can be used to illustrate that"*.

Augmenting feedback provided to students. Several teachers also pinpointed the potential of the dandelion diagram in terms of supporting instructors to generate more personalised, concrete feedback for students. For example, as NT1 explained: *"I think it would definitely give more support in regards to teachers giving student feedback."* She first argued that the visualisation tool could help teachers assess the performances of students in detail: *"you see the visualisation and you say, well, sometimes the students do what they're expected to do, and sometimes they don't, or sometimes they might be slow to react to what's happening"*. She also explained how such insights could help the teacher to give more pertinent feedback to the students: *"then you could base your feedback on that as well. You would actually say to them: I can see that you were in this direction, in this particular position, could you tell me a little bit more about what you were doing during that time?"* Moreover, NT2 pointed out that the dandelion diagrams could help teachers to generate relevant feedback in their own professional development: *"very useful information that could come of that [the dandelion diagram] for the teachers in terms of the feedback in debriefing in professional situations"*.

When to use the Dandelion Diagram. In terms of supporting students' reflection and instructors' feedback, the teachers recognised that the debriefing session (or a quick feedback session) after an episode of practice (i.e., a class), would be a good occasion to integrate the visualisation tool: e.g., *"when we're debriefing with the students. We could show them this visualisation" (NT3). NT2 emphasised that it would be particularly beneficial if the debriefing session took place in "relatively real time, after the scenario."* This would make sure that the practitioners still retain fresh memories about the practice session, *"so they would not need to see the video as well"*. In terms of supporting students' example-based learning, some ad-hoc training sessions were considered suitable occasions to utilise the dandelion diagrams, e.g., to curate examples of good and bad practices (NT2, PT1), and peer case studies (NT3). Moreover, PT2 suggested that the visualisation could also be integrated *"with a journal for teachers to reflect on the decisions they took"*. NT3 suggested using the visualisation as an awareness tool to help teachers *"in trying to maintain professionalism"*. PT1 mentioned the possibility of using the visualisation as a long-term intervention in professionalisation to see how it would promote iteration and evolution in professionals' practice. In addition, PT2 specifically proposed using the dandelion diagram together with video, in which the visuals could be *"useful to give a summary for certain periods"*, and help practitioners *"navigate to a specific video section and then figure out what happened there"*.

6.1.6.3 Challenges, Risks and Limitations

This theme collates the teachers' opinions about the potential challenges, risks and limitations of using the dandelion diagrams in practice, which has led to some detailed considerations in designing and implementing this new tool. The first challenge refers to how to represent absolute time duration in addition to the relative proportion of time in the visualisation. Each frame of the dandelion diagram depicts spatial data within a certain period of time, thereby, its color density represents the relative proportion of time rather than an absolute duration. Yet, as recognised by PT2, practitioners sometimes may be interested to know *"exactly how much time a blue cone represents"*. To resolve this need, I propose that future diagrams could add extra annotations to indicate the absolute duration of a specific data cluster. These annotations could be either auto-generated or based on manual labels.

PT3 raised another challenge which is about the over-saturation of the color density which might cause certain data points being *"washed out"*. This could be avoided by dynamically configuring the amount of color density based on the time-frame of the whole visualised session: i.e., the longer the session, the weaker each accumulation of color density will be. NT2 noted that the position of the tracking sensor on the body should also be carefully considered, to maximally guarantee the accuracy of the data in terms of body orientation: *"there's probably some [...] considerations about how to best position [...]"*. He envisioned that the sensor unit might be integrated into a smart vest or glasses in the near future.

Apart from these design challenges and considerations, NT3 also pointed out two concerns in implementing this visualisation in practice. The first is the potential *"embarrassment"* it might bring to teams of students who did not perform well, and this implies that extra thought would be needed regarding who to share the diagrams with and when to anonymise them. A second concern was about how effective this tool would be in regard to changing professional behaviours. While agreeing with the opportunities that can be created for professional learning and reflection, she also emphasised that more evidence would be needed to assess *"whether it would actually contribute to changing behaviours"*.

6.2 Study 4. Teacher-facing visualisation of students' spatial abilities

The qualitative study in this section used a retrospective reflection technique Hassenzahl and Ullrich (2007) to investigate the Nursing teachers' responses to the ENA diagrams of their students' activity[2]. The discourse analysis of five experts is sufficient to identify salient patterns and is effective for identifying most usability problems with MMLA interfaces Nielsen and Landauer (1993). Our study sought to address three research objectives (Ro):

- **St4-Ro1**: What insights can teachers gain from visual representations of nursing teams spatial behaviours using ENA?
- **St4-Ro2**: What potential uses of ENA for supporting teaching and reflection on nurses' spatial behaviours are envisaged?
- **St4-Ro3**: If potential uses are identified, then what improvements are needed for making ENA representations for special data into effective reflection tools for nurses?

6.2.1 Participants

Five teachers (T1-T5) were interviewed, each of whom had taught the simulation beforehand, to preserve the authenticity and value of the study (females: 4, male: 1, average years teaching: 12.6).

[2]Based on: **Gloria Milena Fernandez-Nieto**, Roberto Martinez-Maldonado, Kirsty Kitto,and Simon Buckingham Shum. 2021. Modelling Spatial Behaviours in Clinical Team Simulations using Epistemic Network Analysis: Methodology and Teacher Evaluation. In LAK21: 11th International Learning Analytics and Knowledge Conference (LAK21), April 12–16, 2021, Irvine, CA, USA.ACM,New York, NY, USA, 11 pages. doi: 10.1145/3448139.3448176

6.2.2 Protocol

The interviews were conducted as approximately 20 minute online video conferences via Zoom, structured as follows: (1) *Explanation of the ENA diagram*. The researcher showed a floorplan to explain the spaces of interest, and then showed the ENA for Team 1 (Figure 6.9a) to show how they corresponded to the node labels. It was emphasised that node locations bore no correspondence to the floorplan, the meanings of edges and thickness were explained as reflecting the number of transitions between locations, but no explanation was given about why nodes were positioned as they were (this was judged to be too complex to explain quickly). (2) *Interpretation of ENA (Ro1)*. Teachers were asked to think aloud while inspecting the connections presented in the three ENA diagrams shown in Figure 6.9. The prompt question was *"According to this diagram, can you explain how TEAM 1 transitions between the Spaces of Interest?"* (3) *Eliciting envisaged usage (Ros 2 and 3)*. After viewing the three ENA diagrams, teachers were then asked about the potential value of ENA for student and teacher reflections, and about any improvements to the ENA design that might help to support their teaching practice.

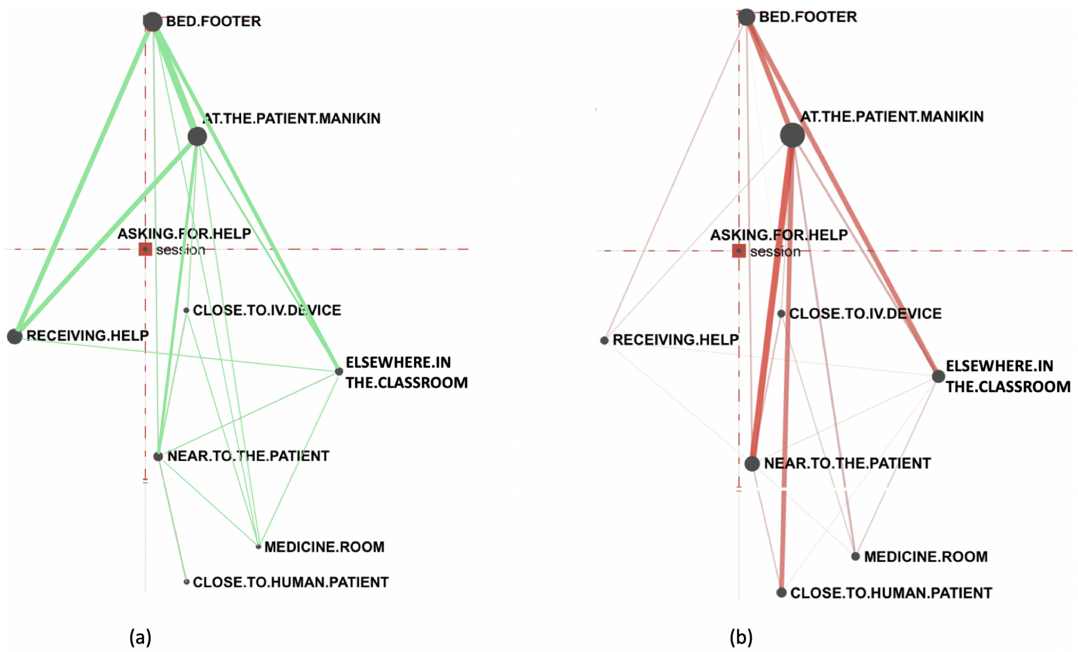


Figure 6.9: Presence of nurses in spaces of interest around the patient and in the classroom during the healthcare simulation using Epistemic Network Analysis for Team 1 (a) and Team 3 (b).

6.2.3 Analysis

The interviews were video-recorded, fully transcribed, and coded using NVivo. Two researchers were present in each session. I examined participants' statements and their actions exploring the MMLA interfaces. Following McDonald et al. (2019), and given the direct alignment between the study protocol and the analysis themes, statements of interest were jointly coded Braun and Clarke (2006) by two researchers according to the pre-set themes of the study protocol: (a) teachers' interpretations of ENA visual representations of spaces of interest; (b) anticipated usage strategies; and (c) opportunities to improve ENA to support reflection. Resulting coded statements were examined by the authors who had several discussions to select instances that illustrate the opportunities and concerns raised by the teachers.

6.2.4 MMLA exploratory interface: Presence of nurses in spaces of interest

The output of the modelling described in Section 4.2.2.4 was processed using the online ENA tool[3] for the duration of the simulation. In the resulting epistemic networks, each node represents the codes for *fixed, semi-fixed and dynamic spaces of interest*, and the activities of *asking for and receiving help*, and each edge represents transitions between two spaces of interest, possibly in conjunction with help seeking/receiving. The positioning of nodes does not correspond to actual positions on the floorplan (a key point to which I will return when I report teachers' responses). Instead, ENA automatically places the nodes in fixed positions to facilitate visual comparison of networks (for details of the algorithm, see Shaffer (2006)). From the 5 teams, I selected the ENA diagrams of *Team 1* (Figure 6.9a) and *Team 3* (Figure 6.9b) for our study, because they were the more contrasting teams (see Figure 6.10).

6.2.5 Results

This section presents the results of the analysis organised around the three questions motivating these studies.

[3] http://www.epistemicnetwork.org/

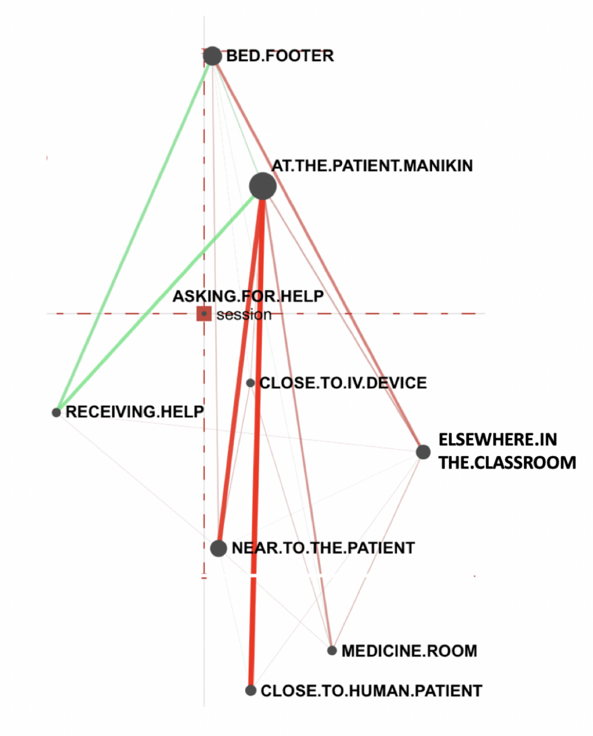


Figure 6.10: The *comparison* network shows the difference between teams: green edges where Team 1 is stronger, red edges where Team 3 is stronger.

6.2.5.1 Ro1: Teachers' Interpretations of the ENA Diagrams

Strong connections. Four of the five teachers were immediately able to start interpreting the ENA diagrams of both teams (see Figure 6.9, a and b respectively) focusing on the visually salient, strongest connections (thicker edges). However one teacher found them very confusing and could not volunteer any reading of them. When interpreting the ENA diagram for team 1, teachers first mentioned the strong connection between the nodes *bed footer* and *at the patient manikin*. According to teachers, the meaning of this connection is associated with the patient-care construct (e.g., students assessing the patient vital signs). By focusing on the other two edges forming a triangle with the node *receiving help*, one of the teachers explained that *the most common behaviour of nurses in team 1 was moving from the head of the bed to the footer, because they [focused on] assessing the patient, and then they were receiving help from the teacher, probably because they needed guidance to achieve the task* (teacher T1). For team 3, teachers highlighted the connections between the nodes *at the patient manikin*, *near to the patient* and *close to the human patient*. For example, teacher *T2* described that *these students seem to have started at the patient manikin more than anything else, then going close to* *the human patient, students did interact with the patient, with the person and the actual manikin*. Similarly, teacher *T4* confirmed this as follows: *it looks like team 3 was doing more of communication with the actual human patient as well*. Regarding other strong connections, teachers explained the meaning of nurses using those spaces, for example, nurses being *elsewhere in the classroom* suggested that nurses may have gone to find additional help (e.g., books), or the ECG device to assess the patient.

In sum, all teachers agreed that team 1 was receiving significant help from the teacher, which for this simulation was not expected, because this is an immersive simulation and the students were meant to be addressing critical incidents independently as a team. Team 3 was much more focused on the patient, as expected. This suggests that teachers associated strong connections to the predominant spatial behaviours of students.

Weak and missing connections. Teachers also interpreted thinner edges. For example, regarding the node *close to IV devices* in team 1, teacher *T1* explained that *the [team members] do not need to be there very often [in that space], they are there just for preparing the medicine and then they give the medicine so that is fine, instead they need to be closer to the patient*. Also, teacher *T5* explained that this weak connection occurred *probably because students first tried to figure out what was going on with the patient*, which according to this teacher is the explanation of the presence of some connection to the node *receiving help*.

Regarding identified missing connections, three out of the five teachers agreed that team 1 was generally not as close to the patient as was expected for this simulation. For example, the teacher *T3* explained that *there was not a lot of contact with the human patient and it is a procedural problem of team 1* (note just one thin edge connected to the node *close to human patient*. However, teacher *T1* argued that, although not ideal, team 1 could still assess the human patient because *students can still talk to the human patient from the other side of the bed. So they may have done a lot assessment on the actual manikin and then maybe just talked to the human patient from the other side of the bed*. The principal insight from the ENA representation for team 3, that all teachers agreed with, was that students were generally far from the teacher, neither asking nor receiving help. For instance, teacher *T1* explained that *this team was autonomous because they look like they asked less for help, even when they may have received some help, it seems they did not depend on it*. Teacher T5 described this behaviour from the spatial traces as follows: *probably this team was in a more advance level of expertise or was more confident with the work they were doing*. This suggests that, based on missing connections, teachers were also able to identify the nurses' lack of presence in spaces of interest, which pointed to qualities and also potential areas of improvement for the teams.

Comparing networks. Four out of five teachers suggested that this visual comparison (Figure 6.10) confirmed what they interpreted from the individual ENA representations. For example, teacher *T4*, suggested that *this [comparison visualisation] just reinforces what I was talking about before, there is a correction to be made for team 1 or a couple of corrections in terms of performance. More interaction to the patient is needed and they should avoid receiving too much help from the teacher. Whereas, team 3, was much more engaged with the patient*. Likewise, teacher *T5* reflected on this comparison, suggesting that *students in team 1, which is the green one, were more keen to ask for more help from the teacher than team 3, which was more independent, and tried to figure it out by themselves what to do*. Finally, only one of the teachers interpreted other connections apart from the more prominent edges. Teacher *T2* explained that *Team 3 went to the medication room a little bit more than the other team* and associated this to the more independent and active behaviours of team 3.

6.2.5.2 Ro2: Anticipated Pedagogical Uses

Regarding the potential use of ENA representations for supporting their teaching practice, four out of five of the teachers agreed that this tool could be very useful for nursing students to reflect on aspects like patient-care and team autonomy. For example, *T2* suggested that this tool could be used *during the debrief session to focus on teams that might have required specific interventions, such as team 1*.

Teacher *T4* also highlighted the potential of using the ENA representations for teachers to reflect on their own practice. This teacher explained how she focused on the extent to which she provided help to the students, as follows: *for me as a teacher if I am doing an immersive simulation, I am expected to let students to figure out the situations or try to address the simulation scenario by their own without my help*.

Moreover, *T1* explained that ENA representations can be very useful for teachers because they normally want to compare teams at a glance, *it is good that you can see the comparison because then you can see the differences among different teams*. Additionally, teacher *T2* mentioned that whether or not students receive or request help *can also indicate that they had to receive a kind of additional support or instruction to address the simulation, it might suggest possible changes in the learning design*.

In sum, teachers recognised the contribution of ENA diagrams to: identify teams' spatial behaviours, compare teams, interrogate their own practice (regarding to what extent they affected the immerse character of the sim), and to revise the learning design.

6.2.5.3 Ro3: Improvements to ENA Diagrams

During the interviews teachers expressed concerns about the complexity of the visual representations to interpret. For example, *T3* stated that he recognised the value of the tool for reflection but *it is a bit difficult to interpret, and there should be some clear guidelines to go to the clinical staff and students for them to understand what the visualisation means*. In fact, I acknowledge that an accurate explanation of the ENA representation (codes and connections) is needed to avoid teachers' misinterpretations, specifically regarding the position of codes and its independence with the actual floor position. This because, all teachers confused the node positions with a floor position.

A number of improvements were suggested, which reflected the distinctions being made between codes (nodes). Three teachers recommended simplifying the ENA representations by combining the codes for the *patient-manikin* and the *human role-playing the patient* (which counter-intuitively, were not next to each other in the diagram): *the patient manikin and the person playing the role of the patient represent the same entity for the simulation, both might refer to patient-care (T1)*; *even when (the human patient and the manikin) are located at different spatial data points, it would be worthwhile to combine them because that's the composite (T3)*.

However, in another instance, the nodes were not making an important distinction. *T3* suggested splitting the code *near to the patient* into two different codes: *there is a left side to the patient and a right side to the patient. There will be different procedures being performed at each side. I think it may be worthwhile to consider separating those ones out* [...] *the right hand side of the patient will be predominantly where students will be doing clinical assessment, checking vital signs, and talking to the patient. Whereas the left side in this scenario is where the IV-device is, having both might bring additional insights about the nurses behaviours*.

Finally, other recommendations were related to the inclusion of additional elements to support interpretations. For example, the *T3* suggested that the ENA representations *could just have some legend down the side or some explanatory notes linked to the edges* to explain the meaning of the connections. Alternatively teachers requested more contextual information regarding what was happening during a particular period of time, for example, when critical clinical procedures were occurring, or focused the analysis pf differentiated spatial behaviours according to the specific roles enacted by students such as the *team leader*.

6.3 Study 5. Teachers' facing MMLA Visual Narrative Interfaces - LAK

A qualitative validation study, using a retrospective reflection technique Hassenzahl and Ullrich (2007), was conducted to investigate the Nursing teachers' responses to 3 visual-narrative interfaces [4]. This validation study sought to address two research objectives:

- **St5-Ro1**: What are teachers' perceptions of representations of student data using i) visual data slices, ii) tabular, and iii) written report formats?
- **St5-Ro2**: What are teachers' potential envisaged uses of the visual-narrative interfaces to support teaching and learning?

6.3.1 Participants

Four teachers (T1-T4) were interviewed (3 females, 1 males, years of experience as register nurses from 40 to 9, avg=21, std=14.6). Each of them belong to the teaching team delivering the unit, to preserve the authenticity and value of the study.

6.3.2 Protocol

To address the research objectives, a **think aloud protocol** was defined. The interviews were conducted in the form of 45-minutes, individual interviews (3 in person and 1 via Zoom). The interview was structured in two parts as follows.

i) Exploration of teachers' reactions to the three different visual-narrative interfaces: teachers were asked to navigate through each interface while thinking aloud. Teachers were asked to compared two different groups using the data slices and compared 5 groups while inspecting the other two visual-narrative interfaces. For the case of interface 1, teacher's navigated first through the timeline (without data stories) to then explore each of the 5 slices.

ii) Open questions about the alternative uses of the visual-narrative interfaces. Triggering questions were used to investigate how teachers envisaged the use of each

[4]Section based on: **Gloria Milena Fernandez-Nieto**, Simon Buckingham Shum, Kirsty Kitto, and Roberto Martinez-Maldonado. 2022. Beyond the Learning Analytics Dashboard: Alternative Ways to Communicate Student Data Insights Combining Visualisation, Narrative and Storytelling. In LAK22: 12 th International Learning Analytics and Knowledge Conference (LAK22), March 21 –25, 2022, Online, USA.ACM, New York,NY, USA, 16 pages. doi: 10.1145/3506860.3506895

MMLA interface to support their teaching practice. The two questions asked were: 1) what would you do with each of these three representations?, and 2) when and for what purpose each representation would help in your teaching or students' learning?

6.3.3 Analysis

The interviews were video-recorded, fully transcribed, and coded using NVivo. I examined participants' statements and their actions exploring the visual-narrative interfaces. Following McDonald et al. (2019), and given the direct alignment between the study protocol and the analysis themes, statements of interest were jointly coded Braun and Clarke (2006) by two researchers according to the pre-set themes of the study protocol: (a) teachers perceptions on different visual-narrative interfaces while exploring them; and (b) open questions regarding the potential uses of each visual-narrative interfaces. Resulting coded statements were examined by the authors who had several discussions to select instances that illustrate the opportunities and concerns raised by the teachers.

6.3.4 MMLA exploratory interface: Visual-Narrative Interfaces

Three visual-narrative interfaces were designed to communicate information about the group(s) and class outcomes, namely, i) visual data slices, ii) tabular visualisations and iii) written reports.

6.3.4.1 Visual-Narrative Interface 1 - visual data slices

The first interface was designed based on the *timeline* metaphor, which is commonly used to represent temporal relations in student actions Riel et al. (2018), and the notion of data slices reviewed above Chen et al. (2019). Five data slices were created based on the interactive slideshow metaphor. Activity logs were grouped into four slices, related to the critical actions modelled by the rule-base algorithms described in the previous section. Critical actions, that involved similar activities, were aggregated in the same slice to avoid a big number of them. That way, critical actions related to vital signs validation (at different moments during the simulation) were condensed in the same slice, as well as IV antibiotic actions (ether administer or stop it).

Figure 6.11 presents an example of data slices that communicate nurses errors. It includes visual enhancements such as *enclosing areas*, shaded areas and boxes, to emphasise where the error was detected or to explain what was the error (A); *changing colour*, to guide attention to certain points or messages such as errors (yellow) or correct actions

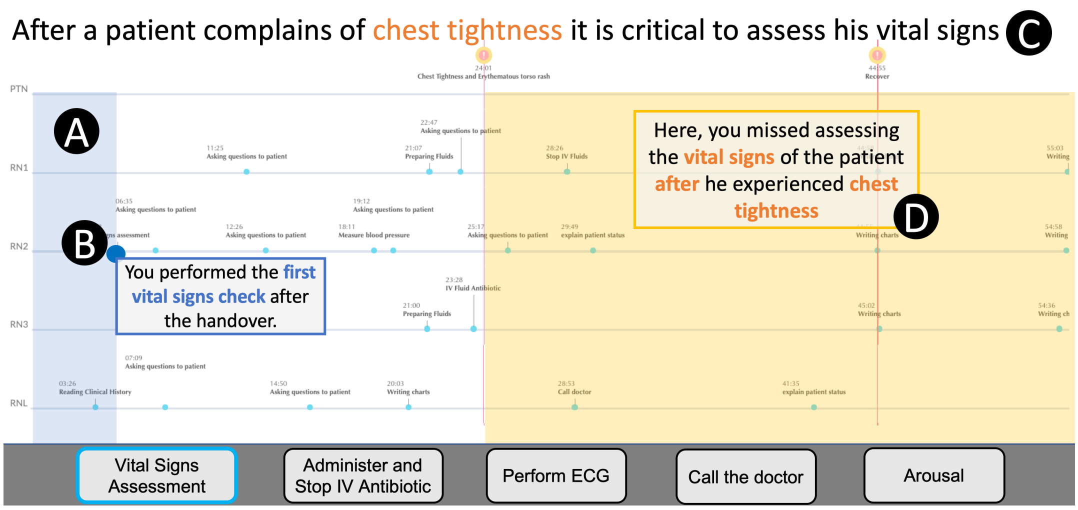


Figure 6.11: Data slice example, communicating that the group of students skipped some actions, including: A) shaded areas to group relevant data points according to the learning task; B) annotated data points; C) a summary of the main story of the data slice; and D) text narrative explaining the main issue in student's performance.

(blue) (B) (principle 1 in Section 2.4); and *text annotations* such a *title* that summarises the take-away message of the slice (C), or additional *explanations* for particular data points (D) (principle 3 in Section 2.4). Gestalt principles (Ali and Peebles, 2013; Knaflic, 2017) of proximity and similarity were used to group similar elements of the messages,

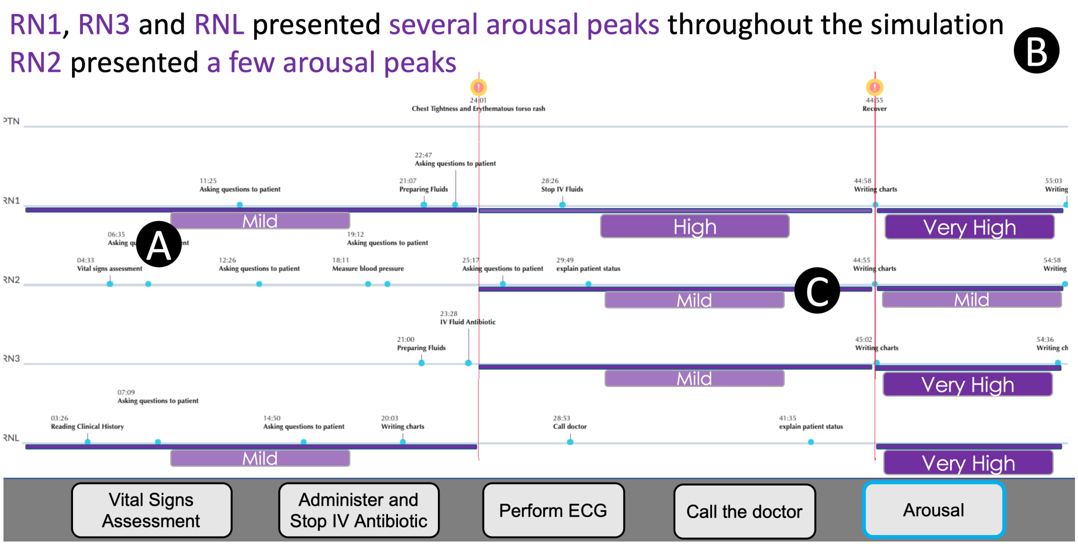


Figure 6.12: Data slice example, communicating nurses arousal peak levels using text labels, including: A) intensity of colour to visualise a range of values from very low (*lighter*) to very high (brighter); B) text indicating the levels experienced by nurses; and C) explanatory labels indicating the arousal levels per phase.

for example using the same color for shaded areas, messages and highlighted points or including annotations in close proximity to points of interest (principle 2 in Section 2.4).

One additional slice was included to communicate student's arousal levels. Figure 6.12 was build to communicate the different levels of arousal peaks that nurses experience during the simulation. This visual-narrative interface includes *changes in colour* (A), indicating levels from very low to very high levels of arousal peaks (principle 1 in Section 2.4). The use of *text* in this interface helps to emphasise the intended message to be communicated (principle 3 in Section 2.4). For example, it explicitly flags those roles that were highly aroused in the title (B), and at the level of the phase of the simulation (C).

	Group 1	Group 2	Group 3	Group 4	Group 5
Initial Vital Signs Assessment A (Asses **after** handover)	✓	✓	✓	✓	✓
IV Antibiotics Administration (Administer IV antibiotic **after** handover)	✓	✓	✓	✓	✓
Second Vital Signs Assessment (Asses **after** chest tightness)	- Miss assessing vital signs **after chest tightness**	✓	✓	- Miss assessing vital signs **after chest tightness**	- Miss assessing vital signs **after chest tightness** C
Stopping IV Antibiotics (Stop IV antibiotic **after** erythematous torso rash) AND (Stop IV antibiotic **in less than** 5 min)	✓	✓	✓	✓	✓
Perform ECG (Perform ECG **after** chest tightness)	✗	✓	✓	✗	✓
Call the doctor (Call the doctor **after** stopping IV) AND (Call the doctor **in less than** 5 min)	✓	✗	- The group called the doctor **6 minutes late** B	✓	✓
Arousal D (stress or cognitive load, count of arousal peaks)	**NA**	**RN1: 7** **RN2: 25** **RN3: 38** **TL: 120**	**RN1: 182** **RN2: 22** **RN3: 156** **TL: 110**	**RN1: 92** **RN2: 32** **RN3: 69** **TL: 50**	**RN1: 351** **RN2: 92** **RN3: 42** **TL: 505**

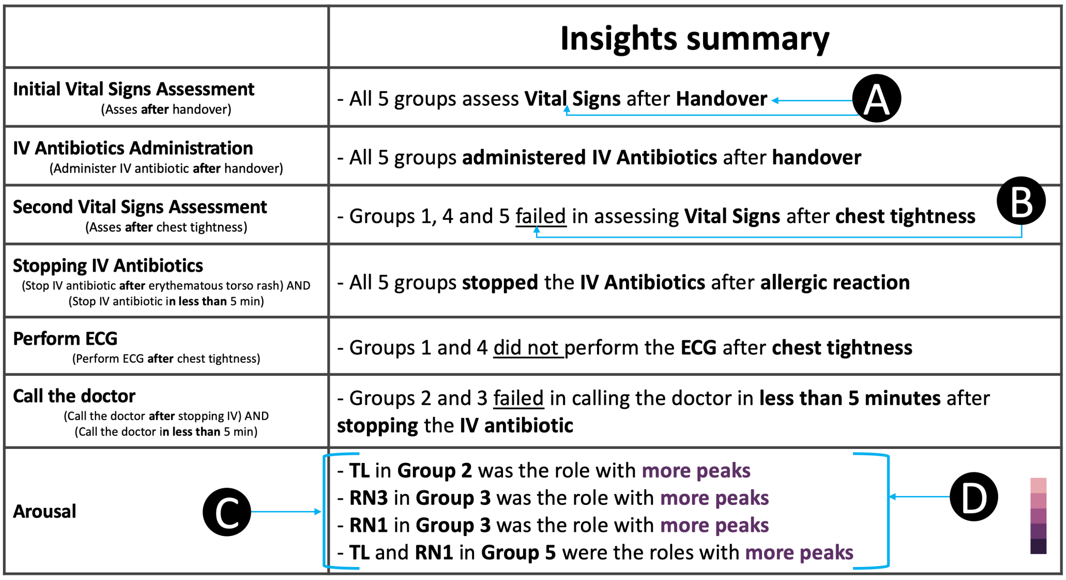
Figure 6.13: Example tabular visualisation (heatmap metaphor) including: A) rules' textual explanation; B) textual complementary explanations; C) changes in colour to present errors in blue and assertive responses in blue; and D) levels of arousal peaks using a sequential scheme of colour.

6.3.4.2 Visual-Narrative Interface 2 - tabular visualisation

The purpose of this interface is to visually report the outcomes of all groups that participated in one simulation (normally 5). The story of the whole class is communicated using a tabular structure based on a tabular heatmap approach (Knaflic, 2017). That way, the table mixes the detail of groups and critical actions while also making use of visual cues to guide teacher's attention. The first *column* of the table in Figure 6.13, introduces the learning intentions to provoke class discussion in terms of the simulation performance or arousal experienced by the students. The other 5 columns aggregate specific outcomes for group 1 to 5. *Rows* in the table, represents each of the critical actions (6 as per the learning task) and one additional row to present the groups' arousal peak levels. Unfortunately, the physiological data of group 1 was not captured because of some issues with the sensing technology.

In this case, *textual explanations* (A), were included to explicitly explain the rules used to generate groups' outcomes. For example, the rule used to generate values in row 4 was presented as follows: *"(Stop IV antibiotic after erythematous torso rash)* ***AND*** *(Stop IV antibiotic in less than 5 min)"*. Additional explanations were included in certain cells when actions were partially correct. For instance, group 3 performed an action but their time response was not appropriate: *"The group called the doctor* ***6 minutes late****"* (B). *Changes in colour*, were used to either emphasise errors (yellow) or assertive responses (blue) regarding critical actions (C). In this case, the levels of arousal were counted and shown explicitly through the interface (D).

6.3.4.3 Visual-Narrative Interface 3 - written report



	Insights summary
Initial Vital Signs Assessment (Asses **after** handover)	- All 5 groups assess **Vital Signs** after **Handover**
IV Antibiotics Administration (Administer IV antibiotic **after** handover)	- All 5 groups **administered IV Antibiotics** after **handover**
Second Vital Signs Assessment (Asses **after** chest tightness)	- Groups 1, 4 and 5 failed in assessing **Vital Signs** after **chest tightness**
Stopping IV Antibiotics (Stop IV antibiotic **after** erythematous torso rash) AND (Stop IV antibiotic **in less than** 5 min)	- All 5 groups **stopped** the **IV Antibiotics** after **allergic reaction**
Perform ECG (Perform ECG **after** chest tightness)	- Groups 1 and 4 did not perform the **ECG** after **chest tightness**
Call the doctor (Call the doctor **after** stopping IV) AND (Call the doctor **in less than** 5 min)	- Groups 2 and 3 failed in calling the doctor in **less than 5 minutes** after **stopping** the **IV antibiotic**
Arousal	- **TL** in **Group 2** was the role with **more peaks** - **RN3** in **Group 3** was the role with **more peaks** - **RN1** in **Group 3** was the role with **more peaks** - **TL** and **RN1** in **Group 5** were the roles with **more peaks**

Figure 6.14: Example Written Report including: A) bold, to emphasise the critical actions; B) underscored, to indicate common errors; C) bold, for arousal to set roles and group number; and D) bold and colour sequential scheme to indicate nurses' arousal peaks

This written reports summarise the key insights concerning critical actions and arousal peaks. The story which summarises the main insights from the class is presented using a textual *narrative* representation. That way, the straightforward key messages supports teachers' inference-making of what happened during the simulation (principle 3 in Section 2.4). The report was organised similarly to the tabular visualisation presented above, with columns and rows referring to groups of students and educational constructs, respectively. Pre-attentive attributes (see Section 2.4) were also used in the form of text formatting with the aim of directing the audience attention (Knaflic, 2017) (principle 1 in Section 2.4). Figure 6.14 presents an example of a resulting written report as a result of our modelling. *Bold* format (A), was used to present text that described critical actions. *Underscored text* (B) emphasises common errors that groups made during the simulation. In regards to arousal peaks, roles and group members were grouped using text in bold (C). Finally, the summary gives and emphasis to the roles that were most highly aroused during the simulation. Because of that, color in text was used (D) to present students, who experienced the most arousal peaks.

6.3.5 Results

This section reports the findings of the individual interviews. Results are presented according to the research objectives.

6.3.5.1 Ro1. Teacher's perceptions to visual-narrative interfaces

After exploring the visual-narrative interfaces these were the more relevant comments from teachers, which illustrate their perceptions and reactions for each interface.

Visual data slices When exploring interface 1 (with no enhancements, only the timeline), teachers tend to miss useful information regarding the critical incidents. For example, groups are expected to assess vital signs at the beginning and after the patient deterioration, all teachers were able to identify (pointing) in the visual representation the critical incident *Vital signs assessment*, but, none of them was able to identify that the second assessment was missed. Their interpretations changed after looking at the first *data slice*, as they realised that the groups they were assessing actually had skipped the second set of vital signs validations (e.g., see Figure 6.11). Using the data slice, teachers' interpretations were more complete and were more aligned to their pedagogical intentions. For example, T4 explained that *"group 1 took the initial vital signs earlier than group 4"* and then she realised that *"nurses haven't actually performed the second set of vital signs"*, which she *"did not pick up earlier"* exploring the timeline with no enhancements. Likewise, for group T3, indicated that *"nurses did the first check and then after the patient deteriorated or it changed in condition, they should do another vital sign check, which is what we teach them and what we would expect them to do"*.

This illustrates the potential for this visual-narrative interface to **communicate one specific message at a time**. This way, teachers can assess how well the team did in specific aspects of students' task.

With regard to the arousal data slice (see Figure 6.12), teachers interpreted the labels that indicated 'high' and 'very high' levels of arousal with those moments when nurses were performing various tasks. For example, T1 explained that *"the RN1 was actually doing a lot of stuff [actions] and so was the team leader, which makes sense, why they have high arousal levels"*. T2 followed a similar interpretation strategy, *"trying to link the high and very high arousal levels with what nurses were actually doing"*. Based on this he indicated that *"RN3 for example was doing an ECG. They've done that really quickly so they've had a fabulous response and RN3' arousal fluctuate to nothing over the beginning of the simulation to a very high arousal in the middle and the end of the sim"*. T3, also indicated that high arousal levels can be indicators of nurses feeling more *"involved"*, *"challenged"*, or *"engaged"*, which was also related to the number of actions that students were performing. In contrast, the mild and low arousal values were associated with how *"confident"* (T1, T2), *"calm"* (T4), or *"disengaged"* (T2, T3) a nurse can be. T4 explained that using this data slice teachers can *"talk about how to prevent stress so nurses do not make mistakes in a real situation where stress can affect group responses"*. In sum, in order to **contextualise** the individual levels of stress, teachers used both the description of actions and the arousal peak labels to justify their interpretations.

More specifically, teachers reacted positively to the *highlighted* elements and the *narrative* included on each of the data slices. For instance, T1 explained that the messages looks *"great, they are really clear, straightaway, the message stated that nurses did miss the vital signs, I did see that correctly"*. T2 indicated that he *"liked the tailored messaging in terms of the time sequence"* and suggested that *"the time responses would be quite interesting to also be included as a part of the message"*. Likewise, T4 identified the value of messages and also recommended that they can be complemented adding extra information that can be useful feedback for students. For instance, T4 thought that performing an ECG group G3 *"could have been performed faster because if a nurse has gone in and stopped the antibiotic fluids they can straight away talk to another nurse to get the ECG"*.

T3, explained that textual messages were *"very clear, nurses administered the antibiotic and then after the chest tightness and the rash they stopped the fluid, so I think that's a good way of showing that that's what they're meant to do and they did it fairly quickly, within less than five minutes. So, that's really good"*. Additional elements, such colour were commented by teachers, for example, T1 explained that she *"liked how the visual elements are matching up, and the colours used [for the text also] match up, so it's really easy to see"*. The *shaded areas* were correctly associated with the group time responses, for example while comparing groups, T1 indicated that *"G1 called the doctor quickly, more quickly than G4"*. The same comparison happened while all teachers explored quick reactions of groups to stop the IV antibiotic. The similarity gestalt principle was evidenced here when teachers **read/group elements with the same colour as part of the same message**. For example, the yellow shaded area and the highlighted words in yellow colour were associated as part of the same message (a group error).

Generally, teachers interpreted key messages using each of the data slices. The elements to capture teacher's attention worked as expected, as per their explanations, they interpreted correctly when groups made errors or asserts. I validated teachers' quotes rightness according to the group outcomes and videos. Additionally, the visual enhancements were mostly well perceived and helped teachers to identify the indented goal for each data slice. Messages, were perceived as useful, however, some suggestions regarding the specificity (e.g., include the time response) are evidenced by teachers.

Tabular visualisations The main reaction of all teachers about this interface was that it was particularly useful to compare different **group outcomes from the whole class at a glance**. For instance, T3 compared and reflected that *"two of the groups [G1 and G4] did not perform an ECG, which is quite a big error"* and then she realised that *"two of the other groups [G2 and G3] did not escalate things by calling to the doctor"*. Besides group comparison, teachers mentioned that individual interventions could also be planned based on this interface. For instance, T4 explained that *"she might be doing interventions per group to talk about specific points highlighted in the table"*. However, she also mentioned that when many groups missed something *"it is important that the teacher reinforces that for all the class and not individually [to a particular group]"* (T4).

The visual elements used in this interface, such as colour and icons, captured teachers' attention. For example, T1 indicated that *"with the table and the colours you can see visually, straightaway the class summary"*. For the case of T4, her attention was focused on the missing actions (errors) and indicated that *"because every group had yellow parts that is something that need to be debriefed"*. Additionally, T4 mentioned that *"ticks and crosses"* can be used as a *"checklist"* to rapidly assess *"what have students done or not done"*. Two teachers (T2 and T4), were interested to see the time response to complement the report, which can be a way to assess if the rules used should be changed or calibrated.

Yet, all the teachers needed additional explanations regarding the meaning of the numbers presented in the last row of the tabular representation (arousal peaks count), because at first they did not understand their meaning. After a short explanation (from the researcher guiding the interview), all teachers tried to make sense of the values. At this point, teachers indicated that having that information could be of interest and can help them to **trigger interactions with students who are under stress**. For example, T1' interpretations suggested that *"in G5, RN1 and the team leader were very stressed. Well, at least these are the highest numbers in the table, probably followed closely by G3"*. Likewise, T3 mentioned that based on the information presented *"G5 needs a lot of support if the team leader is striking 500 times in one simulation hour"*. T3 used the data to reflect about her practice in regard to the higher number of peaks that students had, she said that having these data *"would make [her] interested if maybe [she] was making the students stressed as their teacher"*. Additionally, some suggestions were mentioned by teachers such *"the use of labels or something more meaningful that communicates levels of stress would be ideal"* (T2), he even suggested the use of *"emojis"*.

In sum, teachers perceived the value in the use of this interface for the particular purpose of comparing groups at a glance. In addition, they proposed additional modifications and improvements to the way that certain data, such as arousal levels, were rendered.

Written summary report All teachers appreciated the value of text in this interface to create a summary of the simulation outcomes. For instance, T1 stated that *"with the table you can straightaway see exactly what's going on in a nice table, using text like a report or a summary"*. Also, T2 mentioned that *"interface 3 is a really nice written assessment and I see the value of having text in the report"*. He also stated that he would like this interface *"as the classroom dashboard"* summarising the key insights. Two of the teachers (T3 and T4), indicated that visual-narrative interfaces 2 and 3 were really similar. For instance, T4 explained that *"the last two interfaces were pretty much the same but because she didn't see the difference between them two"*. Teachers who saw no differences between visual-narrative interfaces also mentioned their preferences for interface 2, because they *"like visuals a lot better"* (T3) and because interface 3 *"has a lot more reading"* (T3) and because *"is more textual"* (T4), interface 2 was preferred because as it brings at a glance a visual picture of all group outcomes.

6.3.5.2 Ro2. Envisaged uses of visual-narrative interfaces

Use of visual-narrative interfaces in the classroom Generally, teachers expressed willingness to use all the visual-narrative interfaces to support their classes. For example, T2 mentioned that *"having this visual reminders [referring to all the visual-narrative interfaces] as a **debriefing** is really important. So, from the teacher there might be comments about the general group outcomes, but to focus on the missing aspects. This, in order to have the assessed evidence to focus on for a trigger point for debriefing, is fabulous"*. All teacher agreed that interface 1 can be useful to provide more specific **feedback**, T2 explained that she would be able to provide *"much more depth and much more ability to have trigger material for the debriefing"*. T3 thinks that all of them can be used in different moments during the class *"as soon as the class finishes would be very useful for teachers to use MMLA interface 2 and 3 (tabular visualisation and written report, respectively), to show teachers where nurses did miss out on the essential tasks. Then MMLA interface 1 (data slices) would be very beneficial to use in the debrief but probably before you ask students to reflect back on whether they did what they think they did. After that, hit all of the essential criteria with MMLA interface 1, and then bringing up MMLA interfaces 2 and 3 on the screen to show students if they actually did satisfy these criteria"*.

For visual-narrative interfaces 2 and 3, teachers can see more opportunities to use them to support teachers or other academic staff to validate whole class outcomes and **inform** the learning design. T3, indicated that interface 3 *"would probably be good for the subject coordinator or for their actual report"*. Interface 3 *"is really useful from a research and from a teacher evaluation perspective but not for the students"*. This suggests that teachers may find visual-narrative interfaces 2 and 3 useful to **compare performance** of different groups. Interface 1 is not adequate for that purpose in class, as explained by T4, as follows: *"it's really hard to see as a whole, as teachers we have got five groups happening at the same time and exploring slices of all of them would be difficult"*.

Regarding the data slice about levels of arousal peaks, teachers indicated that it can be used in the class by the teacher and individually. For teachers it would be useful *"to continue to reinforce how to manage stress"* (T2) or to identify *"moments where nurses are feeling more stress"* (T3). However, there is some concern regarding how to communicate such data, T1 indicated that *"it all depends on how the teacher explains [arousal and errors], so you've got to showcase your [students] their strengths and if they do have some weaknesses, you need to present it in a way that it isn't condescending. You've got to present it in a way that [students] can learn from it and not be too critical about it"*.

Uses for assessment or as an additional assignment All teachers agreed that the visual-narrative interfaces can be used as an additional assignment for students to **reflect** on what they did during the simulations. T2, came up with an idea to use the visual-narrative interfaces in which *"students could indicate the role they play (e.g., RN1) so then they [team and individual students] can reflect about the actions that the team performed and the individual role in the scenario. So, then that could form part of an **assessment**, the tricky thing is bringing it up to scale"*. Likewise, T3, mentioned that the visual-narrative interfaces can be used for assessment and reflected about her particular case *"I see some of the students might be on the phone not really engaging with the class. So this is a good set of visuals that can be used to improve the way students engage with the class and learn from their outcomes"*.

Also T2, explained that the visual-narrative interfaces can be used for **group assessment** because *"visual-narrative interfaces are linear and show also independent, individual, data. So, this gives a really good indication of what people are doing"*, but T2' main concern is that visual-narrative interfaces *"don't give you the indication of the interaction with each other"*, referring to interaction as *both verbal and non-verbal communication*.

6.3.5.3 Summary of results Study 5

Revisiting the research objectives (Section 6.2) I can summarise the outcomes from the teacher interviews as follows (see Figure 6.15). Regarding Ro1, teachers used the visual-narrative interfaces to gain different levels of understanding of the class, group and individual performance. For example, the visual-narrative interface 1 (the data slices), communicated key information about one specific group's outcomes at a time. The combination of narrative and visuals in the same interface enabled the interpretation of the meaning of several data points. Yet, this interface is rich and requires moving from one slice to the other.

Visual-narrative interface	Communication medium	Visual enhancements' perceptions
Visual data slices	**Visuals** enhanced with **narrative**	Opportunities: *Tailored messages* describing the meaning of data points. Colour helped teachers to *group data points meaningfully*. Challenges: Effective to navigate details but not for teachers to gain a *wide view of the classroom activity*.
Tabular visualisations	A table combining **narrative** and **iconic** visuals	Opportunities: Icons, colours and text support *quickly identification of points* that need attention. Challenges: Effective to navigate through the class activity but details of *individual team activity get lost*.
Written report	**Narrative** only	Opportunities: *Text highlighting and de-emphasising elements* of the class outcomes. Can inform the learning activity design. Challenges: Textual information will *require additional time to make sense* of the content.

Figure 6.15: Summary of results

The visual-narrative interface 2 visually integrates outcomes from various groups in a whole class at a glance. It uses a tabular format which is an acceptable data visualisation technique (Knaflic, 2017) without being a chart. It combines icons and text. Although a great extent of detail is lost compared to the data slices, it provides a quick overview of the state of the class that could be used in the debrief by the teacher to lead reflection at a class level.

The visual-narrative interface 3 summarised key insights regarding the class outcomes using text only. Although teachers see value in generating this class of report to document the outcomes of a class, its usage may be more limited to documenting and generating resources for teachers' own reflection rather than being used to orchestrate reflection sessions with students during the debrief.

In terms of Ro2, teachers complemented their own reflections by suggesting specific opportunities to use the different visual representations. The data slices were seen as useful to provide detailed feedback to a specific group, or to be given to students as input to complete a reflective (possibly assessed) task on their own. The tabular visualisation and the written report could be used to compare group performance, with the former having the potential to be used as the main reflective aiding interface during a debrief with students post-simulation.

6.4 Summary

6.4.0.1 Teacher-facing *Explanatory* MMLA interfaces

Although teachers are not necessarily data experts, they are experts on their specific domain (e.g., simulation about an adverse reaction to medication). As experts in their educational contexts, teachers were asked to search for insight about their scenarios using exploratory MMLA interfaces.

Classroom Dandelions - Study 3. Teachers from context A (science laboratory 3.2.2.1) and context B (nursing Simulation 2 3.2.1.2) were asked to explore their positioning data in the form of positioning heatmaps and Classroom Dandelions (see Section 6.1). Dandelion Diagrams integrate participant location, trajectory and body orientation over a variable period. The study showed how heatmaps that display only teacher/student location led to misinterpretations that were resolved by overlaying Dandelion Diagrams.

The added information contained in the five classroom dandelions enabled teachers to understand the learning situation in more detail. Science teachers' could reflect about aspects of their practices in the laboratory such: i) proximity to classroom resources (e.g., boards) and uses of them, ii) proximity to students and interactions with them, and iii) instances of co-teaching and f-formations. Nursing teacher's reflections about nurses' simulations include: i) thought about their roles during the practice (e.g., monitoring vs intervening) and ii) identification of team dynamics. The study found that Dandelion Diagrams assisted teachers' sensemaking, but also can raise ethical risks of over-interpretation (sec. 6.1).

ENA representations - Study 4.

Teachers in Simulation 2 (sec. 3.2.1.2) were able to explore ENA representations and gain insights about nurses' uses of spaces of interest. Although ENA representations were not intuitive for teachers to interpret at first, after detailed explanations of the MMLA interfaces, they became familiar with the representation and how to make sense of it. According to the teachers, the ENA representation were useful i) to see the differences among teams and ii) to support instructors to address the simulations (e.g., letting students to figure out situation by their own without much help).

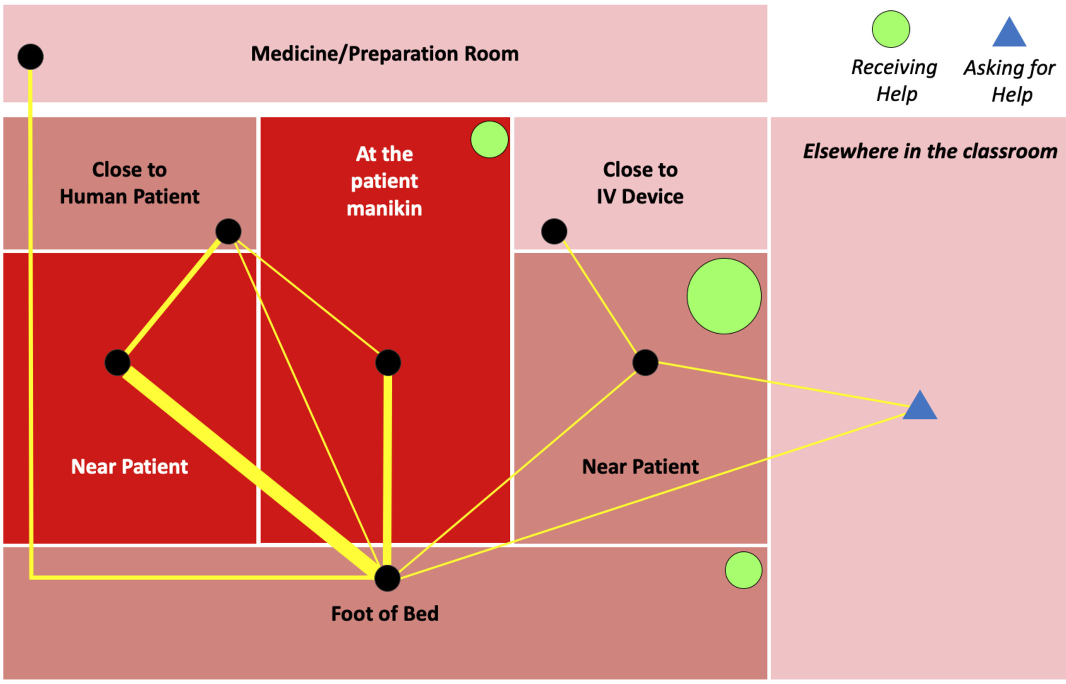


Figure 6.16: Design concept for visual feedback to teachers and students. A map of the simulation ward is overlaid with annotations from the ENA modelling, translating node/edge weights into a heatmap, movement trajectories, and icons for different activities (e.g., circle or triangle).

This study also showed that much of the valuable information in ENA diagrams can be decoupled from the particular network visualisation generated by the current tool. ENA can be used to enrich other spatial visualisations, that make more sense to teachers and students, such as the familiar floorplan of the simulation ward. Therefore, the design thinking has returned to the kind of visualisation developed by Echeverria et al. (2018a), but annotated with the information from the spaces of interest-based ENA modelling that teachers gave such positive feedback on. Figure 6.16 shows an interface mockup, in which ENA edge size is translated into edges overlaid directly onto a floorplan, ENA node size is translated into colour saturation to create a heatmap, which at a glance shows which spaces of interest were most frequently occupied, and asking/receiving help nodes are translated into differentiated icons of different sizes in the relevant location. The modelling could include other weighted activity-based icons, and their weighted interconnections, with controls to view/hide different layers. This map still facilitates comparisons between teams, but might be much more intuitive than ENA's abstract representations, designing out the possibility of the interpretive confusions that the teachers showed.

6.4.0.2 Teacher-facing *Exploratory* MMLA interfaces

Visual Narrative Interfaces - Study 5.

In this study, teachers, who taught Simulation 2 (3.2.1.2), evaluated explanatory visualisations. Based on foundations in data storytelling, three visual-narrative interfaces were designed with teachers: i) visual data slices, ii) a tabular visualisation, and iii) a written report. Each visual narrative interface supported teacher reflections at an individual, team, and class level. The data slices, provided team and individual outcomes of specific learning outcomes, results indicated that they are effective for students to navigate details but not for teachers to gain a wide view of the classroom activity. Opposite to that, the tabular visualisation supported teachers to quickly identify team outcomes but detail of individual teams activity got lost. Finally, the written report provided a summary of the class outcome, that can inform the learning design, it will require additional time to make sense of the content.

In sum, results from this study suggest that alternatives to LA dashboards, such visual-narrative interfaces, can be considered as effective tools to support teachers' reflection, and that LA designers should identify the representation type that best fits teachers' needs.

CHAPTER 7

TOWARDS FULL AUTOMATION: FUNCTIONAL ARCHITECTURE AND REFERENCE IMPLEMENTATION

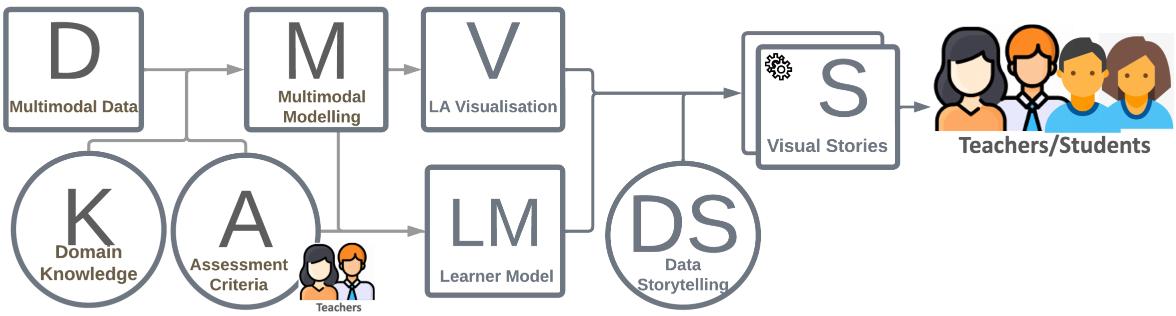


Figure 7.1: Automating data storytelling

The fifth and sixth contributions of this thesis are described in detail in this chapter, which addresses the question *RQ4: To what extent can MMLA interfaces for students and teachers be automatically generated?* The modelling techniques described in Chapter 4 and the evaluation of the MMLA interfaces described in Chapters 5 and 6 demonstrated the potential of the conceptual framework (Figure 7.1) as a guide for the design of MMLA interfaces that support student reflection on their embodied team activity. The focus of this chapter is to explain and document a **Functional Architecture** (I called YarnSense) and a **Reference Implementation** to demonstrate that such MMLA interfaces are feasible to be implemented. More specifically, this chapter illustrates how (i) the system can be implemented in a scalable architecture, and (ii) the workflow can be automated from *feedback design* through to *delivery* as data stories.

Section 7.1 describes the four components that define the Functional Architecture to implement MMLA interfaces, and Section 7.2 details the implementation of the MMLA interfaces. These contributions constitute a proof of concept that the approach developed in this thesis, as summarised in Figure 7.1 is technically feasible.

7.1 YarnSense: Functional Architecture

This section presents YarnSense (see Figure 7.3), an architecture to automatically generate MMLA interfaces enhanced with DS principles (student data stories for now on). The student data stories guide students' interpretations to reflect on their embodied team activities. As shown in Figure 7.2, the conceptual framework works as a reference to map from abstract concepts to components of the Functional Architecture. The Functional Architecture consists of 4 components and is described below.

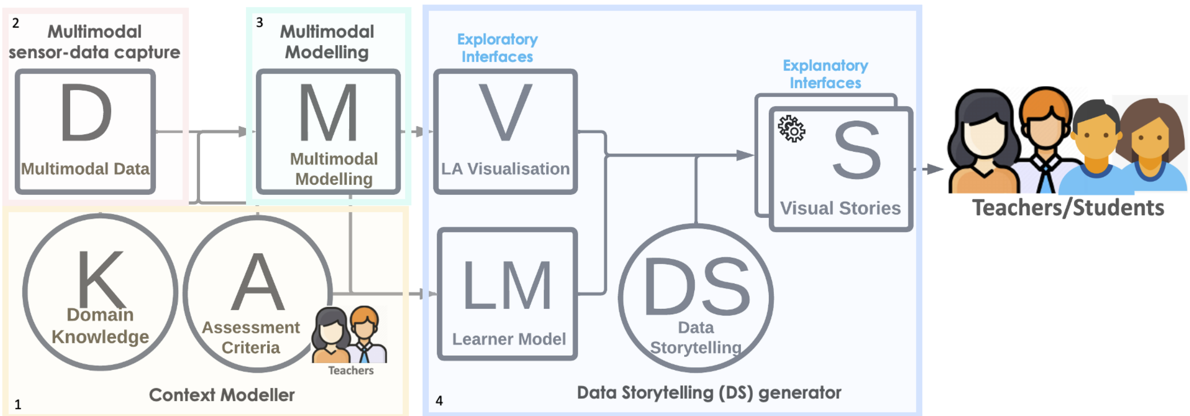


Figure 7.2: Conceptual Framework used as a reference to map the functional architecture in four components

The following sections explain each of the four components in more detail.

7.1.1 Context modeller

The context modeler (see Figure 7.3, 1) provides a set of user interfaces (UIs) for an expert in the embodied team activity (e.g., a teacher or researcher) to capture *pedagogical intentions* in a way that is readable to the system. Embodied team activities (e.g., nursing simulations) can commonly be specified based on the learning design. Such learning

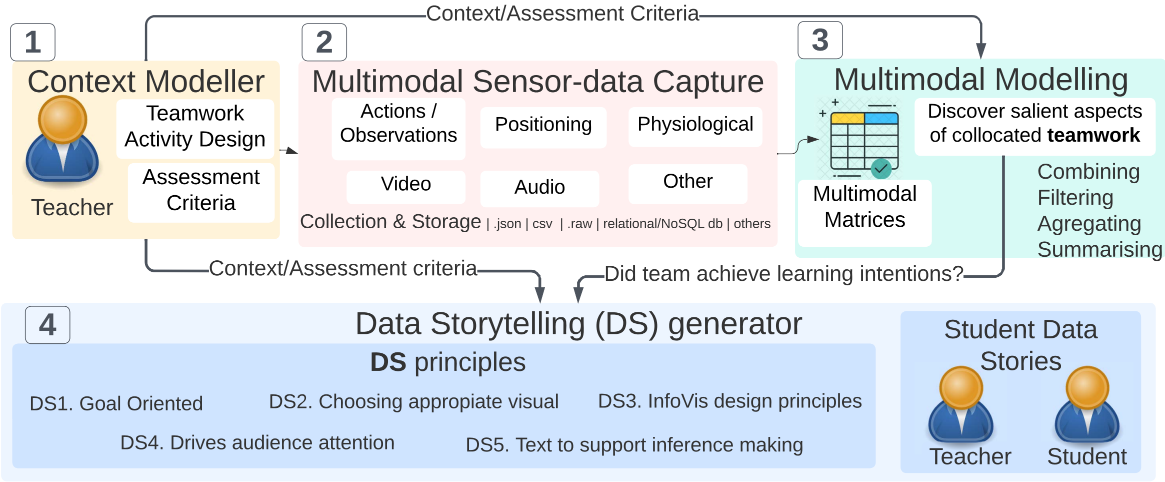


Figure 7.3: Main components of the YarnSense Functional Architecture: context modeller, multimodal sensor-data capture, multimdoal modeller, and data storytelling (DS) generator

design specifications can serve as an input for this first component. For example, the following information could be specified as **parameters** to contextualise the embodied team activity (K, Domain Knowledge in Figure 7.2):

- The list of *actions* of interest which are expected to occur during the activity or which will provide context in the analysis, such as critical moments or milestones during the embodied activity.
- Information about the physical resources available in the classroom settings. For example, the manikin, trolleys, or *sensor* data to be collected for the activity (e.g., indoor positioning data, audio, or arousal peaks).
- *Objects* including roles of team members and the devices that will be observed during the activity.

In addition, as presented in Chapter 4 the teachers' pedagogical intentions can be elicited and specified in a way that the context modeller can interpret them (A, Assessment Criteria in Figure 7.2). Both the parameter data for the embodied team activity and the assessment criteria should be accessible through the learning design of the embodied activity. Ideally, contextual information is provided prior to the embodied learning activity.

7.1.2 Multimodal sensor-data capture

The multimodal sensor-data capture component (see Figure 7.3, 2) collects embodied learning activity data captured both by *wearable sensors* and *humans-as-sensors*. Capturing multiple sources of data can potentially provide a wider understanding of the embodied team activity. As explained in the background section, sensors can capture vast amounts of interaction data from physical spaces at scale. Although sensor data can be analysed to generate immediate outcomes they commonly lack of of a higher-order meaning (Wise et al., 2021). On the other hand, while humans are able to observe physical and social events and give meaning to certain team actions they are not as accurate and precise as sensors can be (Wang et al., 2014). The second component of this architecture facilitates a collaboration of sensors and humans by collecting and storing different modes of physical interactions including human observations, physiological data, indoor positioning data, audio, and video. This participatory sensing guarantee that humans collect observations to enrich the sensor provision of continuous data.

7.1.2.1 Technical considerations

Technical requirements can be accommodated for the data collection according to the embodied activity. Important technical aspects for further implementations of component 2 are described as follows:

- It should be able to adapt and respond to unexpected events, such as failure of sensors. One way to achieve that is to reduce interdependencies (Pautasso and Wilde, 2009). Loosely coupling architectures can provide that (Kaye, 2003). For the second component of this architecture, sensor data are ideally captured in a parallel, independent, and loosely-coupled manner (e.g., using different applications per modalities). Thus, each modality can be cleaned and stored in a variety of formats (e.g., json, csv, mp4, relational databases, logs, or others).

- It should provide a way to discover, collect, and synchronise different sources of data automatically. To synchronise the data captured from multiple sensors, it is crucial to be able to relate all data capture times with each other. An strategy to synchronise between sensors (captured with separate clocks), is to centrally monitor the timings of all sensors (Lichtenauer et al., 2011).

- If the context requires real-time processing, then a streaming strategy is more suitable. That means that data can be collected, stored,and analysed as they are streaming or generated. If real-time is not needed, data can be collected and centralised after the observed activity (batch solution). A mixed strategy (e.g., batch and real-time) can also be implemented, as some modalities might require to be collected as they are being streamed, while others can be collected in batch and then uploaded to the storage (local or cloud) system for further analysis.

7.1.3 Multimodal modelling

The modelling techniques explored in Chapter 4 provide context for the automated analysis of the multimodal data. Component 3 in Figure 7.3 enriched the low-level data captured by sensors and humans in component 2 using the contextual data parameterised by the context modeller (component 1). This component aims to incorporate qualitative aspects that describes meaning to low-level data. Previous research has demonstrated how quantitative ethnography (QE) can be used to translate from low-level data to higher-order levels of meaning (e.g., (Echeverria et al., 2019)). Thus, component 3 implements a QE approach(Shaffer, 2017) and the notion of the Multimodal Matrix (MM) (Buckingham Shum et al., 2019a). In the MM data structure each data modality m is coded into n columns which are called *multimodal observations*. For sensor data, each row can represent a time window (e.g., one second) of the embodied team activity. Using the MM, the low-level data can be interrogated against the teachers' pedagogical intentions. Thus, the system will convert low-level data into meaningful information to validate if for example, a certain combination of logged events can be an indicator of students performing a correct procedure.

In this component, the multimodal matrices are interrogated against the teacher's pedagogical intentions (introduced in component 1) to generate a *Learner Model (LM)*. The LM is an structured representation of student performance, misconceptions or difficulties. The Learner Model validates the team's and/or student's data, using the contextual data to assess if the team achieved the pedagogical intentions defined by the teachers (by filtering, combining, summarising, or aggregating the MM, see component 3 in Figure 7.3).

The Learner Model can be student-centred tracking individual activity and competency curves over time (e.g., the model could highlight how students performed at specific moments ($time = 1$)). For example, it can be used to assess whether each team member physiological signals are indicative of higher stress levels. However, it is not limited to individual learners; it could also be a vector that contains information about how a team or different learners achieved certain learning goals / intentions. For instance, using the low-level *x, y* positioning coordinates, the systems will be able to indicate the spatial orientation among team members (e.g., face to face while doing a procedure). Using this model, it is possible to differentiate between individual and team achievements and behaviours. Although other modelling techniques (besides QE) can be implemented, a LM needs to be defined (e.g., a way to validate if the team did as expected).

Outcomes from this component can be in the form of *exploratory* user interfaces designed for experts to analyse and gain meaning of visualisations (Echeverria et al., 2018a), commonly without an specific educational goal in mind. Nevertheless, the following component is provided to support a variety of users expertise in data visualisation analysis.

7.1.4 Data storytelling generator

The data storytelling generator (Figure 7.3, 4) uses the inputs and outcomes from the first 3 components (top components in Figure 7.3). This component aim to support less technical team members and teachers/coaches by generating *explanatory* visualisations. That way, visualisations presented to end-users of the system emphasises insights rather than data points (Knaflic, 2017). Thus, this component automatically provide **student data stories** to guide or support the interpretation of end-users (e.g. teachers or students).

The teachers' pedagogical intentions provided for team performance experts (that is, teachers who define what *high-performance teamwork* looks like for a particular scenario) are then used to generate the student data stories. The outcomes of the modelling (component 3) are enhanced and rendered using data storytelling principles (described in Section 2.4) such as: including title, highlighting important elements, using a colour scheme (e.g., blue for correct performance), removing noise and distractions, and focusing attention on specific aspects relevant for the learner model.

Thus, student data stories are communicated using visualisations and narrative to provide individual/team outcomes according to the Learner Model (described above). For example, a story can highlight key data points of interest for the embodied team activity (e.g., critical event/action) and add descriptions to inform individuals/team about errors, stress levels, or team interactions that needs to be reinforced. The student data stories are illustrated in the reference implementation (Section 7.2).

7.2 Reference Implementation: prioritisation of beds Simulation

The Functional Architecture described above was implemented for Healthcare Simulation 3 (Section 3.2.1.3), focused on providing care to four patients and prioritising the care of each bed as a team. This section explains the technical considerations for the implementation of each component of the architecture. The components in this implementation are meant to be used by a non-technical person. Thus, educators and researchers can use this implementation to provide parameters for the data collection and analytics modules to generate data stories.

7.2.1 Context modeller

A custom web application was developed[1] to help experts incorporate relevant contextual information about the embodied team activity (Simulation 3 3.2.1.3). Figure 7.4 presents four user interfaces (UI) to specify the context of the embodied team activity. The first interface (a in Figure 7.4) defines the list of reference **actions** (e.g.,"Dressing change") that nurses are supposed to perform according to the educational context (presented in Section 3.2.1.3). In the second interface (b in Figure 7.4) the **sensors** used to collect evidence of embodied team activity (e.g., "microphone") are specified. The third interface (c in Figure 7.4) allows the specification of **objects/roles**, to indicate who is participating in the embodied activity. The last interface (d in Figure 7.4) registers information about the assignment of sensors to specific objects (e.g., an ECG device) and roles (e.g., "nurse2"), so it is possible to see what/who is wearing what. The tool to specify the initial parameters was developed using the Express Node.js framework and hosted on an Amazon Elastic Compute Cloud instance (Amazon EC2).

Using another view of the same web application (see Figure 7.5, A), teachers' pedagogical intentions were captured as the **assessment criteria** for the embodied team activity and structured in the form of **rules**.

The teachers who devised this simulation defined five rules to assess the expected **actions** of highly effective teams for Simulation 3 (see Section 3.2.1.3). For example, some rules were defined by teachers with the intention of providing the following.

- Feedback based on the sequence of actions: **sequence rules** (e.g., provide oxygen after the patient respiratory depression);

[1]https://github.com/Teamwork-Analytics/obs-rules

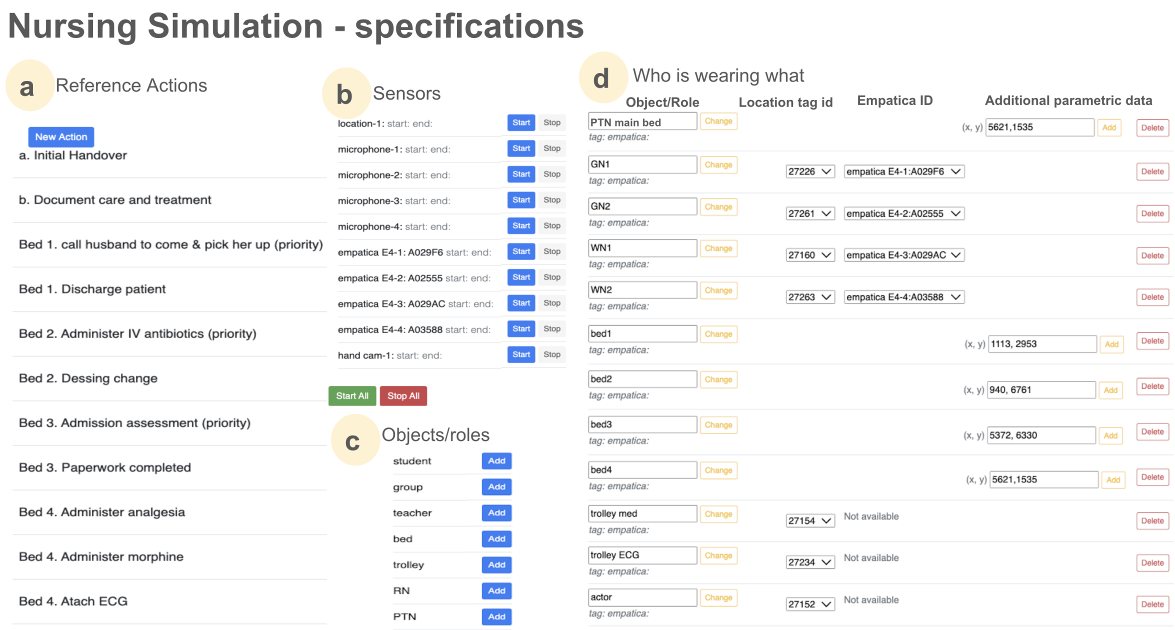


Figure 7.4: The nursing simulation specifications tool include: a) critical actions from the learning design to be manually logged by a human observer during or after the team activity, b) sensors to be used for the data collection, c) a list of objects and people (roles) being tracked, and d) a list indicating who is wearing what sensor.

- Feedback base on timeliness of actions: **timeliness rules** (e.g., stop the IV device in less than 5 minutes);
- Feedback based on the frequency of actions: **frequency rules** (e.g., validate vital signs every 5 minutes); and
- Feedback based on interpersonal proximity: **proximity rules**. By including the interpersonal proximity rules, teachers were able to understand the student's spatial behaviours.

Using the interface, teachers can modify (add, remove, or edit) the rules according to the needs of the embodied team activity. When creating the rules, teachers specify (see Figure 7.5: i) the type of rule to be created; ii) a short and self-explanatory name of the rule; iii) the reference action(s), which are of interest to the observer (e.g. from patient respiratory depression to patient recovery); iv) additional rule inputs, such as time considerations, roles, or devices (e.g., was the team leader in close proximity after the patient deterioration?); and v) feedback messages, for correct and incorrect actions.

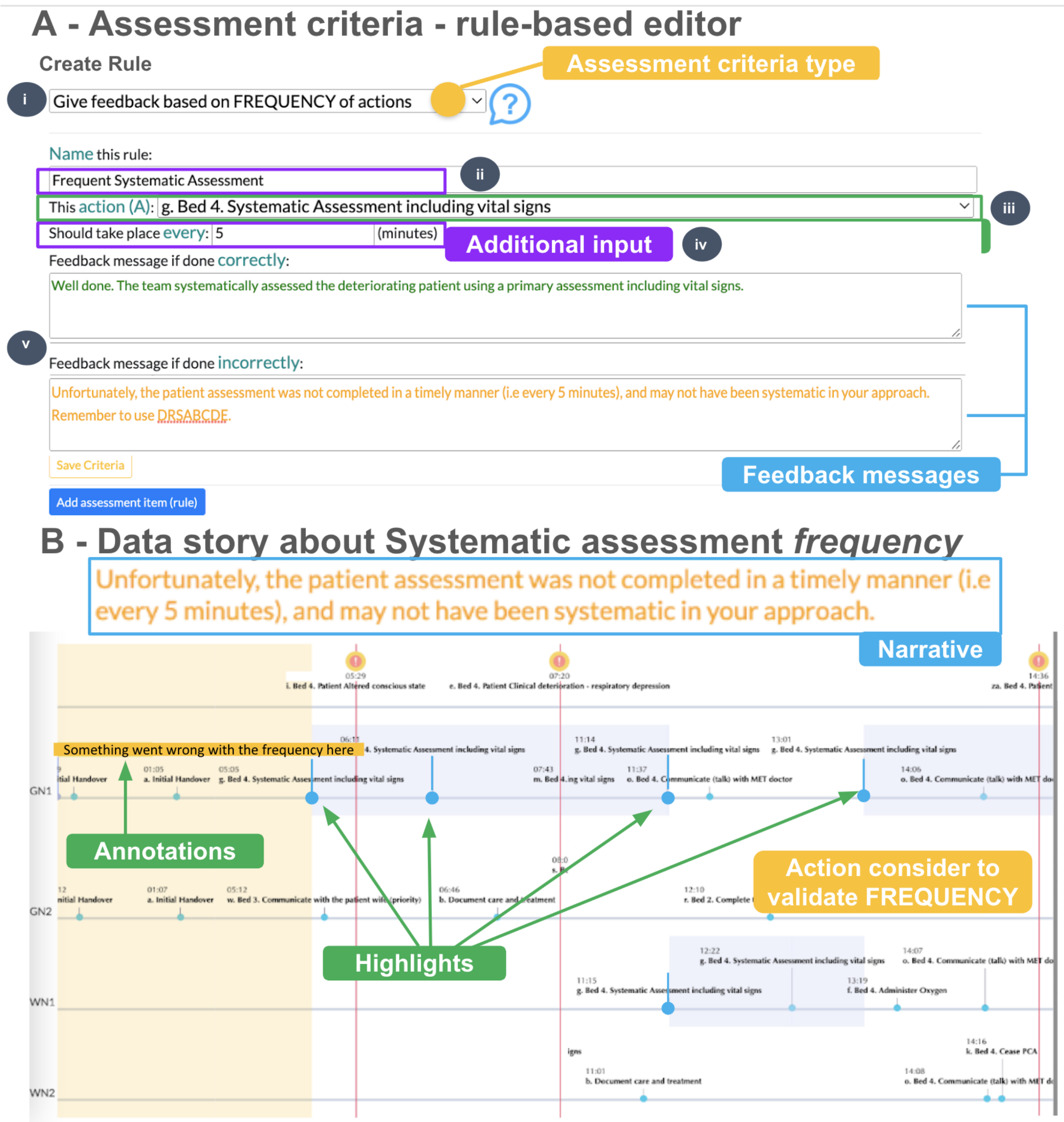


Figure 7.5: A. User interface for teachers to define the assessment criteria in the form of rules, and B. Data story driven by the rules predefined by teachers

7.2.2 Multimodal sensor-data capture

For this implementation all the multimodal data was collected locally and stored in the designated local directories and databases. Different modalities were captured in parallel. This component was implemented using the Spring-Cloud framework [2]. Each

[2]https://spring.io/projects/spring-cloud

component of data collection was developed as a web service, with a Eureka[3] server managing all of them. Using this structure, different services can be easily scaled up to handle large-set data processing if required. In this way, each service is an individual component and can run separately.

A custom web application service [4] was used to collect **audio**. This application distinguishes speakers by using different channels of the audio interface device. Then, the collected audio data are stored in the local machine with appropriate labels and timestamps (such as its *sessionId*, start timestamp, and stop timestamp). In parallel, **video** data was captured using two sources: i) a custom video web application service and ii) the simulation room camera. This is because the simulation room always has the camera on and I cannot connect the camera to our service. The video application uses the Jabra PanaCast[5] 180 degree wide angle camera to record data that can see the entire simulation room. The video data also contains the audio for all conversations in the room and are saved locally on our service running computer.

Positioning and **physiological** data were also collected locally, using a custom Python script and a C# application. Since the Empatica server (E4 streaming server[6]) was unstable, I removed the automatic physiological data collection service and chose to use the manual way to record Empatica data. The Empatica wristbands were turned on before the baseline recording starts, in order to synchronise all the simulation data. Finally, **actions**, are captured via a user interface *Start All*/*Stop All* buttons, and stored in a relational database (MySQL). The sensor data collection is initialised and stopped through the user interface (see Figure 7.4c). The *Start All* button performs three actions, (1) stop the recording audio of the baseline, (2) capture the simulation start timestamp for all services, and (3) notify all services to start capturing data.

Table 3.1 (in Chapter 3) summarises the different sensor data collected for this study. The formats used to store the different data modalities are presented in column 3 of Table 3.1. Tailored scripts were implemented and used to store sensor data in defined data formats.

[3]https://github.com/Netflix/eureka
[4]https://github.com/Teamwork-Analytics/multimodal-audio
[5]https://www.jabra.com.au/business/video-conferencing/jabra-panacast
[6]https://developer.empatica.com/windows-streaming-server.html

7.2.3 Multimodal matrix modelling

For this study, I instantiated a multimodal matrix using the critical actions logged by human observers and the positioning data captured by the positioning sensors. In this case, the columns in the matrix correspond to all the actions that team members are expected to perform, and the rows indicate the time (one row per second) when each action was performed by a particular team member (see the modelling technique in Chapter 4.2.1 and Figure 4.6). The remaining columns of the matrix were defined to model low-level positioning data into higher-order proximity constructs. Taking into account the construct *co-presence in interactional spaces* (explained in Chapter 4.2.2), it was possible to validate the proportion of time nurses spent in close proximity to 1) the four patients (beds 2-3); 2) the patient in the primary bed (bed 1); and 3) other members of the team (Graduate Nurse 1 and 2 -GN and Ward Graduate Nurse 1 and 2 -WN). Here, the columns of the matrix corresponded to the proximity detected between team members and patients (see construct 4.2.2.2); and the rows captured the time when this co-presence was detected (see Figure 7.6).

LOW-LEVEL POSITIONING DATA					CO-PRESENCE IN INTERPERSONAL SPACES							
phase	Time	Role	x	y	**Proximity to GN1**	**Proximity to GN2**	**Proximity to WN1**	**Proximity to WN2**	**Proximity to bed 1**	**Proximity to bed 2**	**Proximity bed 3**	**Proximity bed 4**
PH1	sec. 1	GN1	6132	2471	NA	Social	Intimate	Social	Intimate	Public	Public	Public
PH1	sec. 1	GN2	1481	6101	Social	NA	Intimate	Public	Intimate	Public	Public	Public
PH1	sec. 1	WN1	…		intimate	Intimate	NA	Social	Public	Public	Public	Intimate
PH1	sec. 1	WN2			Social	Public	Social	NA	Social	Public	Public	Intimate
PH1	sec. 2	GN1			Public	Intimate	Public	Intimate	Intimate	Public	Public	Public
PH1	sec. 2	…	…	…	…	…	…	…	…	…	…	

Figure 7.6: Portion of a MM, showing nurses co-presence in interactional spaces. Interpersonal Distances to team members (GN1, GN2, WN1, WN2), primary bed (bed 1), and other patients (bed 2-3)

The modelling techniques and the learner model definition consist of custom applications developed in different programming languages such as Python or Java. For this reference implementation the tailored scripts were implemented in Python.

7.2.4 Data storytelling generator

The rules created by the teacher using the context modeller assessed the order, timeliness and/or frequency of the actions performed by teams. The *Learner Model* was extracted from the matrices by interrogating them using *rules*. Part of the matrix is used to identify

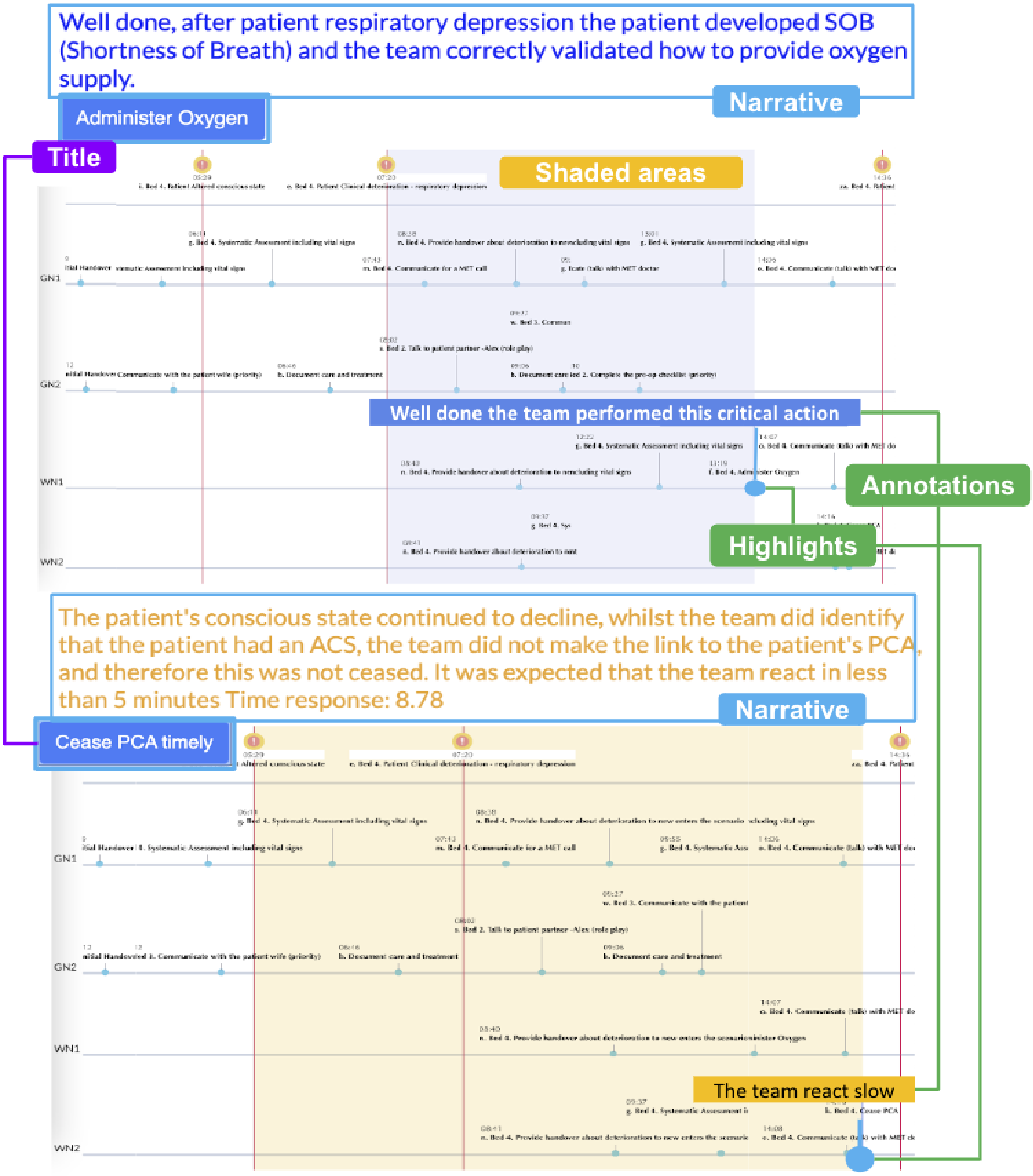


Figure 7.7: Student data stories focused on actions. These stories include DS principles such as title, narrative, annotations, shaded areas, and highlighted elements. Left story: showing stories about timely response, the story indicates that the team reacted slow in ceasing the PCA. Right story: showing the correct sequence of actions, the story indicates that the team identified the patient respiratory depression cause

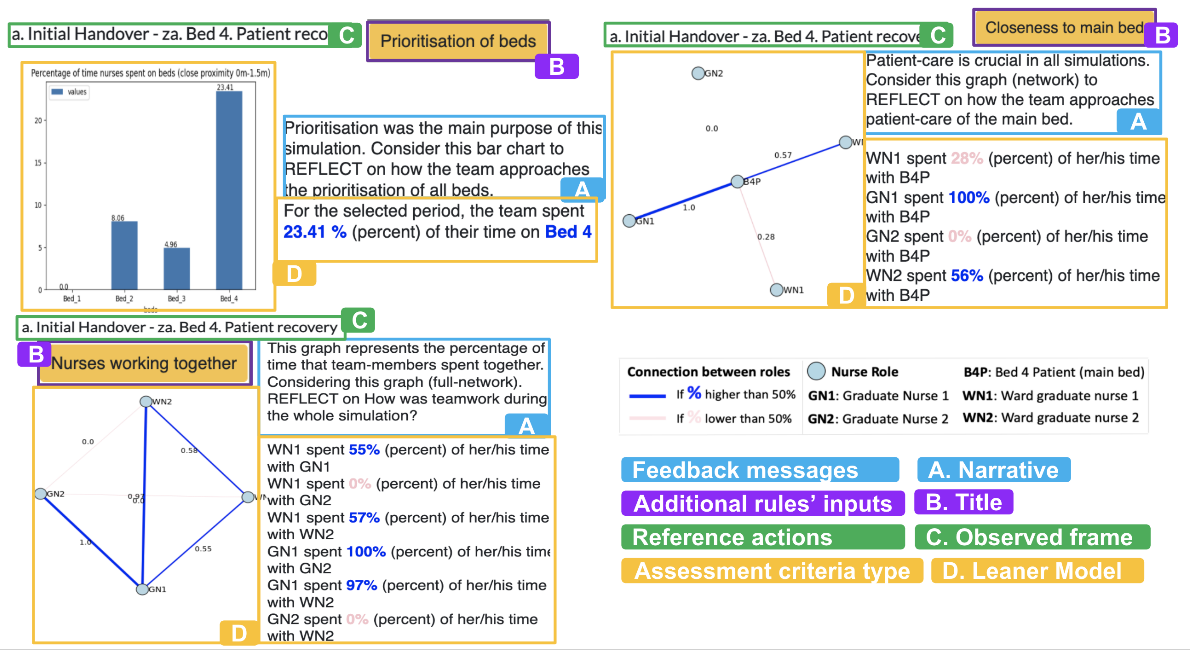


Figure 7.8: Students' data stories focused on positioning data. The narrative and titles are extracted from the assessment criteria created in the context modeller component. The reference actions and the type of intended feedback are used to drive the analysis to generate the Learner Model. The reference actions selected by teachers in the context modeller serve as a filter for the sensor data to focus the analyses on specific moment of the embodied activity. Then based on the assessment criteria, the learner is generated and the outcome is presented in the form of a visual graph (e.g., bar chart), narrative, and statistics. This figure shows three stories: the proportion of time that nurses spent on the four beds (top left), the proportion of time that nurses spent in close proximity with the patient on the main bed (top right), and the proportion of time that nurses spent working together (bottom left).

team successes and errors based on the critical actions recorded. Three types of errors were automatically identified for this simulation. **Sequence errors** are flagged if the team performed critical actions in the wrong order, for example, if they forgot to perform a vital signs assessment. **Timeliness errors** are identified if team members reacted slowly according to healthcare guidelines and teacher experience, for example, if it took more than 5 minutes to stop the PCA device (patient-controlled anaesthesia). An error related to **frequency of actions** is defined by calculating the timestamp difference between two critical actions that are meant to be repeated, for example, evaluating the patient's vital signs every 5 minutes. Teachers can configure the kind of feedback to be generated for each case, as illustrated in Figure 7.5.

The MMLA interfaces (outcome of the Multimodal Modelling) include: visualisation (e.g., bar charts, social networks); explanatory narrative (e.g., messages such: well done, nurses spent most of their time in bed 4, which was the priority); statistics (e.g., % of time that nurses spent on bed 4); or data summaries (e.g., aggregated values of nurses close proximity). From the teachers' pedagogical intentions, the following information is used to enhance the MMLA interfaces.

- Adding feedback messages to explain the team outcomes using narratives (blue in Figure 7.8)
- Adding additional inputs such as time, roles, or a title to provide additional context to the story (purple in Figure 7.8)
- Using the reference actions as a time frame window to filter the matrices and provide insights about specific moments of the simulation (green in Figure 7.8)

Part B in Figure 7.5, shows an automated Student data story rendered based on one of the teachers' pedagogical intentions. The story, emphasises that for this team the vital signs assessment was not systematic (e.g., every 5 minutes). Figure 7.7 presents two examples of stories showing errors (left) and actions performed correctly (right). The stories include **DS enhancements** such as highlighted points, title, narrative, colour (orange for errors and blue for correct performance), and annotations. Figure 7.8, presents other 3 examples of automated data stories: a) a bar chart indicating the percentage of time that nurses spent in close proximity to different beds (top left in Figure 7.8); b) a representation of the role-centred ego network, showing the percentage of time that nurses spent in close proximity to the patient (top right in Figure 7.8); and c) a full proximity network, depicting physical proximity among all team members. In sum, for this reference implementation, 3 types of MMLA interfaces with DS principles were automated:

- MMLA interfaces based on Study 1 (Section 5.1) communicating stories about team performance in the sequence, frequency and timeliness of actions. Due to technological limitations, the timeline to communicate physiological data was not implemented.
- MMLA interfaces based on Study 2 (Section 5.2) communicating stories about the proximity of nurses to i) main bed, ii) team members, and iii) other patient beds.
- MMLA interfaces based on Study 5 (Section 6.3) communicating a written report summarising the main outcomes of the team (see Figure 7.9).

Print

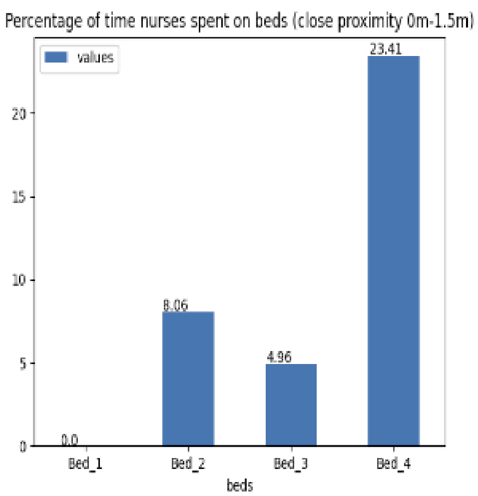
Rule	Team Feedback	Outcome
Timely Naloxone l. Bed 4. Administer Naloxone - 5 - k. Bed 4. Cease PCA	Be careful next time, after the PCA was ceased the patient was overdose and needed timely administration of medication to improve his/her status, Naloxone was the best alternative.	✗
Administer Oxygen f. Bed 4. Administer Oxygen - After - e. Bed 4. Patient Clinical deterioration - respiratory depression	Well done, after patient respiratory depression the patient developed SOB (Shortness of Breath) and the team correctly validated how to provide oxygen supply.	✓
Cease PCA timely k. Bed 4. Cease PCA - 5 - i. Bed 4. Patient Altered conscious state	The patient's conscious state continued to decline, whilst the team did identify that the patient had an ACS, the team did not make the link to the patient's PCA, and therefore this was not ceased. It was expected that the team react in less than 5 minutes Time response: 8.78	✗
Frequent Systematic Assessment g. Bed 4. Systematic Assessment including vital signs - 5 - null	Session 33 - Team A (31st Aug) - Unfortunately, the patient assessment was not completed in a timely manner (i.e every 5 minutes), and may not have been systematic in your approach.	✗
Activate MET call m. Bed 4. Communicate for a MET call - After - e. Bed 4. Patient Clinical deterioration - respiratory depression	Well done. Well done, after any clinical deterioration the team must call for help, activating MET calls.	✓
Prioritisation of beds a. Initial Handover - Priority - za. Bed 4. Patient recovery	Prioritisation was the main purpose of this simulation. Consider this bar chart to REFLECT on how the team approaches the prioritisation of all beds. For the selected period, the team spent **23.41 %** (percent) of their time on **Bed 4** 	

Figure 7.9: Explanatory MMLA interface, team report

7.2.5 Scaling up the MMLA implementation: deployment

The Reference Implementation detailed above provides a proof of concept that the MMLA interfaces (whose prototyping and validation are reported in previous chapters) can indeed be implemented. This section briefly documents ongoing work to validate the usability of the context modeller, and the deployment of the whole system at larger scale. One of the best ways to validate the usefulness of a tool is observing how others adopt it and use it. In the rest of this section I briefly describe how other researchers at another university are making use of the system I developed.

Deploying the context modeller with teachers in a different university.

The paper and digital prototypes described in this thesis were all evaluated with teachers at the University of Technology Sydney who, after several iterations, became familiar with and trusted the novel interfaces. This reference implementation has now been appropriated by other researchers and then deployed in the Nursing program at Monash University, which involved introducing the interfaces to teachers through walkthroughs and interviews.

One week before scaling up the system's trials (described below), seven nursing educators volunteered to participate in walkthroughs of the interfaces and an exploration of the context modeller user interface (male=1, female=6, aged 44.7 years on average, std=7.71). They were asked to translate their pedagogical intentions – what feedback they wanted to give to students, depending on their simulation performance – into a form that could be processed computationally, and instantly rendered for the teacher to verify. Following the methodology for eliciting pedagogical intentions presented in Chapter 4 (in Section 4.2), teachers designed the feedback using the tool to set parameters and conditional rules, which were subsequently rendered as data stories. Teachers followed a think-aloud protocol to share their experiences as they completed the task and checked if the feedback generated was as they intended, iterating the rules as required. Teachers' responses were broadly positive, and their use of the tool showed that the context modeller was usable enough for them to translate their pedagogical intentions into structured rules to operate on multimodal data.

This is a significant step in giving more agency to teachers in MMLA, since up until this point, only a technical researcher (the author) could codify and modify feedback rules. While the context modeller user interface is far from as polished as it could be with further work, nursing academics were able to configure, check, and revise the MMLA feedback visualisations. However, the HCLA approach adopted in this thesis recognises that while an interface may reach a degree of usability, it takes time for teachers to build trust and confidence to be ready to use new tools with students in simulation debriefing sessions. Due to scheduling constraints, it was possible to conduct this evaluation only one week before the next round of MMLA data gathering (detailed next). This was not enough time for teachers to become sufficiently familiar with the interfaces that they were ready to use them "live", which would also require the redesign of their debriefing sessions (i.e., the learning design) to incorporate the automated feedback interfaces.

In sum, this deployment illustrates how the context modeller can not only enable teachers to configure the data capture and data storytelling components of the system but can also contribute to gradually build confidence and trust with teaching teams. It is an important aspect of ethical HCLA practice that teachers' agency is respected, and they are not rushed into using MMLA prototypes that they are not confident to use in front of their students.

Scaling up MMLA infrastructure deployment for massive data collections.

Investigating the MMLA interfaces presented in Chapters 5 and 6 demonstrated that the explanatory MMLA interfaces were used positively to support the reflections of the students on their embodied team activities. Therefore, Simulation 3 (described in Section 3.2.1.3) was selected as the ongoing focus to scale up infrastructure deployment, at a second university (Monash University), also using this simulation exercise, to record 224 students in 38 classes, led by 10 teachers. To the best of my knowledge, this is the largest MMLA dataset gathered to date. The reference implementation was able to generate the feedback interfaces immediately on completion of each simulation exercise, but as explained, the academics were not yet ready to run live debriefs using the interfaces.

Yet, the MMLA infrastructure I created has not only been used for data collection but also to advance teamwork research. For example, the collected multimodal data offers new opportunities to explore different modalities of embodied team activity. As presented in recent work by Zhao et al. (2022), whose authors combine audio and positioning data (collected using the YarnSense reference implementation). In their research, the authors model student communication based on spatial (position) and audio (voice detection) to explore if they can serve as meaningful indicators of team performance according to teacher learning intentions.

7.3 Summary

This chapter addressed **RQ 4:** *To what extent can MMLA interfaces for teachers and students be automatically generated?*

Motivated by the methodology to elicit teachers' pedagogical intentions (Section 5.2) and the positive perceptions of the Explanatory MMLA interfaces presented in Study 1 (Section 5.1), Study 2 (Section 5.2) and Study 5 (Section 6.3), a Functional Architecture and a Reference Implementation have been developed as a proof of concept that MMLA can be implemented. The four components of the Functional Architecture *YarnSense* were inspired by positive perceptions of the modelling techniques (Chapter 4) and the MMLA interfaces (Chapters 5 and 6). Finally, the architecture was implemented and deployed in a significantly scaled up context (Section 7.2).

From a software engineering perspective, which is the focus of this chapter, the architecture: i) is adaptable to different contexts (e.g., different universities and learning contexts); ii) demonstrates its value for massive data collections; iii) implements the end-to-end process, from interface concepts, through to full rendered interfaces; and iv) generates automated Explanatory MMLA interfaces (for 38 classes). Therefore, this chapter has demonstrated the feasibility of designing and implementing a MMLA infrastructure according to sound architectural principles and the automation of the workflow from *feedback design* to *delivery*, through an implementation that scaled well for large nursing classes with new students and teaching teams, at a second university.

Next iterations of this research are planned, as part of a nationally funded project, to extend the implementation and evaluation of this research. The introduction of the tools into debriefings is scheduled for the remainder of 2022. Finally, although *YarnSense* was implemented for nursing teamwork simulations, there is no reason in principle why it could not be adapted to provide teamwork feedback in other domains or to provide an e-research infrastructure for Team Science research.

CHAPTER 8

DISCUSSION, FUTURE WORK, AND CONCLUSION

This chapter synthesises the results of the previous chapters, discussing my results with respect to the research questions first proposed in Chapter 1, and then working to describe the implications of this work. Limitations of my work to date will be considered, along with suggestions for future work that would help to strengthen the field of MMLA. Please consider Figure 8.1 and Table 3.3 as a guide to navigate this discussion.

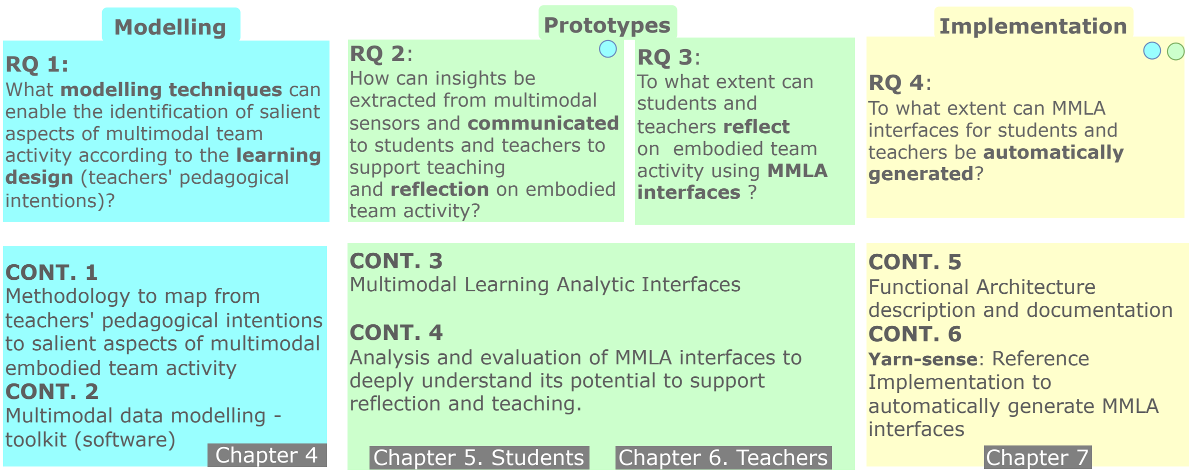


Figure 8.1: Summary of the thesis research questions and contributions (Chapters 4-7)

8.1 Summary of results

Revisiting the research questions (presented in Section 1.2), I can summarise the results of the different studies as follows:

8.1.1 Research Question 1

RQ1: *What modelling techniques can enable the identification of salient aspects of multimodal team activity according to the learning design (teachers' pedagogical intentions)?* (RQ1 in Figure 8.1, blue boxes)

Chapter 4 documented multimodal modelling techniques to gain understanding of how low-level multimodal data can be mapped into higher-order constructs. Using Quantitative Ethnography (QE) (Shaffer, 2017) as an approach to integrate quantitative data with qualitative insights, this thesis implemented different instances of the Multimodal Matrix (MM) (Buckingham Shum et al., 2019a) using logs, physiological, and positioning data. Contributions made by this thesis take the form of adding the following modelling techniques to identify salient aspects of embodied team activity.

- **Activity log modelling: mapping from actions to team errors.** This model detected errors using a rule-based algorithm informed by teachers' pedagogical intentions, with three rules that validate the sequence, timeliness, and frequency of actions (in Section 4.2.1.1).

- **Physiological data modelling: mapping from electrodermal time series data to categorical indicators of arousal.** This model detected students' arousal peaks and classified their intensity into bands from very-low to very-high aroused (in Section 4.2.1.2).

- **Co-presence in interactional spaces.** This model identified the interpersonal interactions of the team members as a proxy of the social bonds between the team members (in Section 4.2.2.2).

- **Socio-spatial formations.** This model identified facing-formations (f-formations) as an indicator of team coordination (in Section 4.2.2.3).

- **Presence in Spaces of Interest.** This model translated positional data into more meaningful indicators of nurse placement in critical care zones (in Section 4.2.2.4).

8.1.2 Research Questions 2 and 3

RQ2: *How can insights be extracted from multimodal sensors and communicated to students and teachers to support teaching and reflection on embodied team activity?* (RQ2 in Figure 8.1, green boxes)

To address RQ2 and RQ3, Chapters 5 and 6 presented the multimodal prototypes that were designed to communicate the results of the multimodal modelling techniques (explored in Chapter 4). They also include results of the evaluation of prototypes with students and teachers.

This thesis investigated both Explanatory and Exploratory prototypes, to communicate information identified in multimodal data that may lead to human insights about embodied team activity. The following list summarises the contribution of this thesis in terms of prototypes.

- **Three explanatory storytelling prototypes** (one per nursing simulation): to communicate team outcomes in terms of i) team errors, ii) nurses' arousal levels, and iii) a prototype explaining the rule-based algorithm used to generate the stories (see Section 5.1), using DS.

- **Four illustrative vignettes of nurses' proxemic behaviours**: to visualise and provide illustrative interpretations of i) nurses' full-proximity networks and role-centred ego networks, ii) nurses' f-formations, and iii) Epistemic Network Analysis representations (see Sections 5.2 and 6.2).

- **Five illustrative vignettes of the Classroom Dandelion**: to visualise the position, trajectory, and body orientation of students and teachers (see Section 6.1).

- **Three Visual-Narrative Interfaces**: to investigate alternative ways to communicate team insights using DS and narrative, with three designs: i) visual data slices, ii) tabular visualisations, and iii) a written report (see Section 6.3).

- **Annotated version of ENA representations**: node position in the ENA representation was confusing to teachers when they referred to the spatial position. For this reason, an annotated version of the floor plan of the ward was designed (see Section 6.4.0.1).

RQ3: *To what extent can students and teachers reflect on embodied team activity using MMLA interfaces?* (RQ3 in Figure 8.1, green boxes)

Five qualitative studies were conducted to evaluate the MMLA interfaces. These prototypes were evaluated by teachers, students, or using a set of illustrative vignettes. This research demonstrated the potential of MMLA interfaces to prompt reflections on their embodied team activity. The following list summarises the type of reflections that became possible using the MMLA interfaces.

- **Explanatory storytelling interfaces**. Students used the data stories to reflect on: i) opportunities to improve, ii) specific skills and knowledge they need to master, iii) strategies to tackle different situations in the future, iv) their accuracy of team / individual judgement on performance, and iv) their collaboration and teamwork strategies (see more details in Section 5.1.5).

- **Nurses' Proxemic Behaviour Vignettes**. Patient-centred ego network helped students and teachers identify the presence of team members in the vicinity of the patient's interactional space. Full social networks can model proximity ties among nurses, indicating whether nurses worked together appropriately during the simulation (e.g., during the delegation phase, nurses are expected to be close to the team leader). ENA representations provided meaning to the coordinates (x and y) of individual team members by modelling interpersonal ties based on proximity, as well as the particular locations adopted by individuals during the simulation (see more details in Section 5.2).

- **Classroom Dandelions**. These were effective in augmenting teachers' sensemaking, compared to heatmaps. Science teachers could reflect on aspects of their practices in the laboratory such as: i) proximity to classroom resources (e.g., boards) and uses of them, ii) proximity to students and interactions with them, and iii) instances of co-teaching and f-formations. Nursing teacher's reflections included: i) their roles during the simulation (e.g., monitoring vs. intervening) and ii) identification of team dynamics. The ethical risks of over-interpretation were discussed (see more details in Section 6.1).

- **ENA representations**. According to the teachers, the ENA representations were useful i) to see the differences among teams and ii) to support instructors to address the simulations (e.g., letting students to figure out situation by their own without much help). (see more details in Section 6.2.5)

- **Visual-Narrative Interfaces**. Each visual narrative interface supported teacher reflections at the individual, team, and class level. The data slices provided team and individual outcomes of specific learning outcomes, results indicated that they are effective for students to navigate details but not for teachers to gain a wide view of the classroom activity. In contrast, the tabular visualisation supported teachers to quickly identify team outcomes but details of individual teams activity get lost. Finally, the written report provided a summary of the class outcome that can inform the learning design, but will require additional time to make sense of the content (see more details in Section 6.3.5).

8.1.3 Research Question 4

RQ4. *To what extent can MMLA interfaces for students and teachers be automatically generated?* (see Figure 8.1, yellow boxes).

Two contributions were presented in Chapter 7 in response to RQ4.

- **Functional Architecture** (YarnSense): to automatically generate data stories driven by teachers' pedagogical intentions. YarnSense materialises all the lessons learnt in the five qualitative studies described above. Four components encompass the requirements and needs of stakeholders to support their reflections on their embodied team activity. The context modeller (component 1) allows teachers to include their pedagogical intentions. Next, the multimodal data (component 2) and teachers' intentions are used to drive the multimodal modelling (component 3) to automatically generate enhancements in the form of data stories (component 4). The four components thus work together to identify salient aspects of embodied team activity and communicate them as data stories aligned to teachers' pedagogical intentions to support students' reflections.

- **Reference Implementation** in Nursing Simulations: to demonstrate the feasibility of implementing automated Multimodal Data Storytelling for embodied team learning in real scenarios.

8.2 Implications

This section considers the implications of my research for supporting: i) embodied team learning, ii) teaching and students' reflection on embodied team activity, iii) teachers' and students' sensemaking, and iv) LA design.

8.2.1 Implications for supporting Embodied Team Research

The CSCW field has contributed rich insights into the value of reflective practices to improve team performance (Diamant et al., 2008), foster purposeful learning (Stahl, 2011), develop professional practices (Leinonen et al., 2005), and the re-design of group work practices (Prilla et al., 2013). In the work presented by Diamant et al. (2008) the author identifies the need for more research to understand critical aspects of successful team performance, such as situational awareness, in knowledge-intensive projects (e.g., software development, customer service, corporate training). In this regard, this thesis contributes modelling techniques to identify salient aspects of embodied team activity, for example, by analysing multimodal data through the lens of "Proxemics" theory.

Recognising the role of reflection in improving teamwork practices, the modelling techniques and MMLA interfaces reported in this thesis can provide inspiration to generate new evidence-based tools to support embodied team learning. One of the consequences of making evidence of teamwork visible is that people can be held responsible for their actions (Erickson and Kellogg, 2000). This was recognised by teachers in studies 1 and 3 (Sections 5.1 and 6.1). Specifically, to develop professional practice, the MMLA interfaces explored in this thesis support teachers and students in prompting reflections about their nursing simulations and co-teaching practices.

From a methodological and practical point of view, this thesis contributes to the call by CSCW researchers (e.g., Diamant et al. (2008); Fitzpatrick and Ellingsen (2013)) for a deeper and granular understanding of collaborative work in physical settings. The methodology presented in this thesis enables the elicitation and formalisation of teachers' pedagogical intentions, contributing to providing contextual tools to reflect. That way, the analysis and visualisations provide relevant insights about aspects of the embodied team activity. For example, the interfaces presented in Study 1 (Section 5.1) helped the students reflect on their roles and the accountability of multidisciplinary teams. Discussing responsibilities is a key approach for teams to conciliate, mediate, and solve problems collaboratively (Robertson et al., 2010). In the same way, the methodology used in Study 2 (Section 5.2) addressed authentic questions about the spatial behaviours of nurses that teachers may have about a particular learning context. In that study, the modelling and the MMLA interfaces evidenced aspects of micro mobility and spatial arrangements of the team, which are of interest to the CSCW community (Fitzpatrick and Ellingsen, 2013).

From a practical point of view, the MMLA interfaces (exploratory and explanatory) support embodied team research by providing evidence-based tools that support reflection on embodied team activity. For example, although social network analysis (SNA) has been used in CSCW research to identify proximity between team members (e.g., using sociometric badges (Kim et al., 2008) or infrared tags (Montanari et al., 2018)), modelling indoor positioning data as proxemic constructs enables the generation of more complete models. In this thesis, I went beyond modelling proximity between two people by also considering: the context in which they encounter one another, where precisely these interactions occur, the locations where individuals perform work, and micro mobility aspects (such as f-formations), which can provide more nuanced indicators of team behaviours.

Although my modelling approach originated from clinical simulation and laboratory science scenarios, based on prior CSCW research, I see significant potential for its use in a wide range of teamwork settings. This would meaningfully facilitate evidence-enhanced professional training, data-enabled workplace optimisation, and fundamental research on professionalisation. Furthermore, the proxemic constructs modelled in this thesis could also be used to meaningfully analyse team performance in other contexts. For example, modelling co-presence and f-formations is now enabling the investigation of how multiple teachers interact with students in the classroom, as recently explored by Martinez-Maldonado et al. (2020b). Similar space-time team dynamics could be modelled for the case of team sport events (e.g., see positioning visualisations of basketball matches in Goldsberry (2012)). Continuous and automatic modelling of f-formations has been explored for the purpose of designing interactive systems (Rädle et al., 2014). The combination of such f-formation data with daya on the presence of team members in meaningful spaces could also be extended to training settings, such as team firefighting, where indoor positioning has also started to be used (Wake et al., 2019); and workspace analysis that seeks to understand how interior elements and displays influence collaborative behaviours of office workers (Radoi et al., 2017).

8.2.2 Implications for supporting teaching and students' reflections on embodied team activity

The contributions explained in Chapter 7 of this thesis present a tool (context modeller) for **teachers to be in control of the analytics** that they want to use to support their practice. That way, teacher's expertise can be combined with the use of technology to render personalised MMLA interfaces. This was documented by Pardo (2018), whose work explained that these tools can be defined as "technological exoskeletons", where

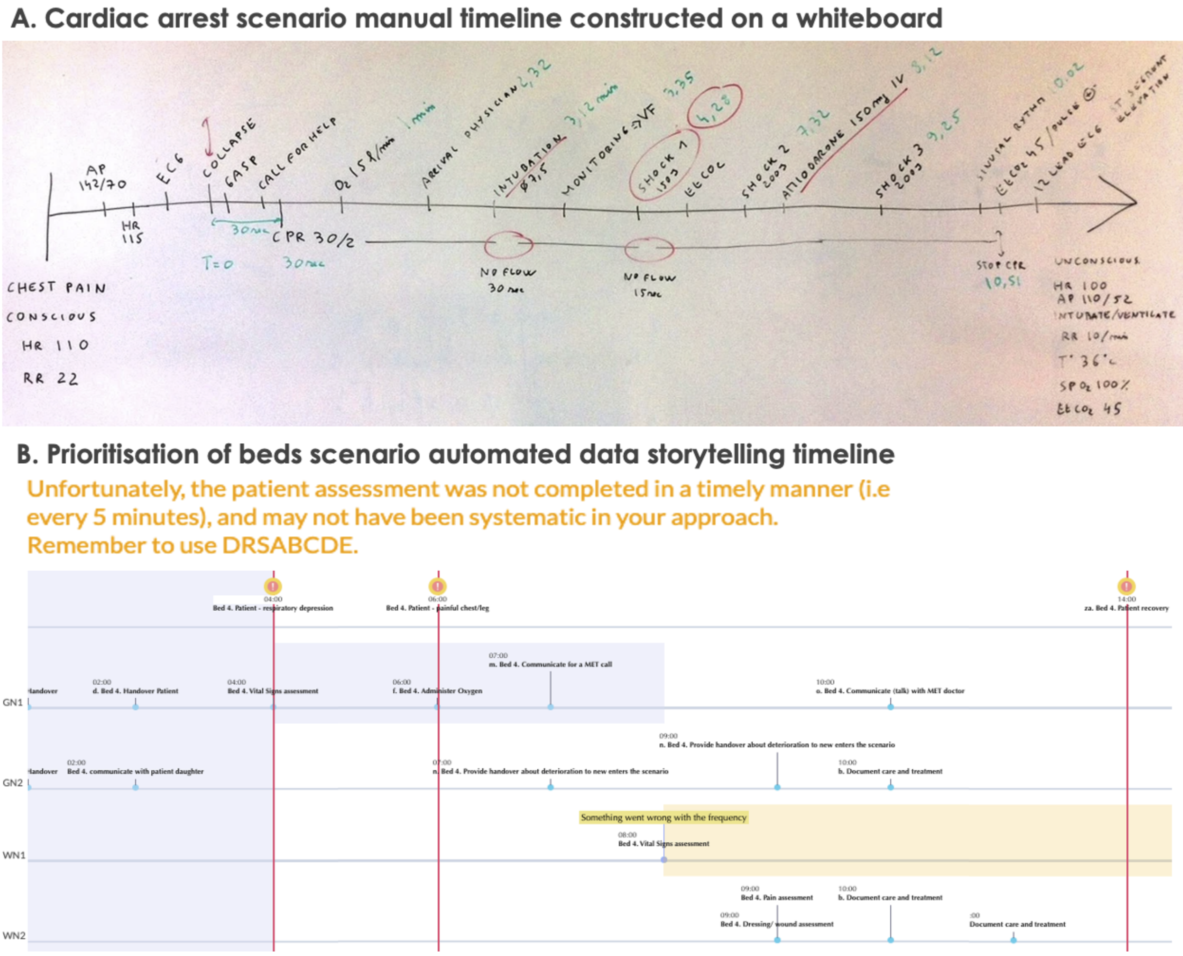


Figure 8.2: Timeline Debriefing Tool. A) Cardiac arrest scenario manual timeline constructed on a whiteboard (Secheresse and Nonglaton, 2019). B) Prioritisation of beds scenario automated data storytelling timeline.

the expertise of the teacher can be enhanced by technology. The findings of this research indicated that the MMLA interfaces can be used to support teachers in the classroom. For example, interfaces help teachers compare different teams that performed at the same time, which is currently not possible without evidence (e.g., watching the video of teams performing) and can be a burden for educators (Tuck, 2012). In the same way, nursing teachers indicated that the MMLA interfaces are potentially useful to guide their reflection sessions at the end of their classes. In fact, the MMLA interfaces presented in Study 1 (5.1) and Study 5 (Section 6.3) use the notion of a timeline representation to show detailed information on the individual and team performance of nurses. As recent research by Secheresse and Nonglaton (2019) indicate, timeline representations are pedagogically useful to guide conversation during debriefing. From this work, Figure

8.2.A presents an example of a timeline built manually by students on a whiteboard. Although I perceive limitations to constructing the timeline from memory and using physical media to represent it (e.g., the floor with a rope and stickies), I also recognise the value of such interactive and participatory activities with physical artifacts. Automated timeline representation (see Figure 8.2.B) can support teachers in planning their debriefing sessions, and a promising avenue for future research is to facilitate participatory engagement around such data-generated visualisations (e.g., following a "participatory quantitative ethnography" approach (Buckingham Shum et al., 2021)).

The studies show consistently positive responses from teachers and students that the way in which data was communicated provoked **deeper reflection** about patient-care, teamwork, and co-teaching. Nursing debriefings are dependent on expert teacher observation (but often stretched over multiple teams) and the partial (sometimes stressed, and always biased) memories of students. As a result, teachers recognised the value of capturing objective evidence of embodied team activity and rendering it visible as MMLA interfaces. For students, the use of tools for reflection, such as layered prototypes (Section. 5.1) or the visual-narrative interfaces (Section 6.3), can serve as tools for **reflection based on evidence** collected through multimodal data (Martinez-Maldonado et al., 2020a). Similarly, evidence of co-teaching is very difficult to capture Veteska et al. (2022), with videos being the regular source of it. However, reflection on co-teaching is challenging due to time constraints Sanchez et al. (2019).

In both contexts, team-based reflection on past team activity is considered by some authors the most crucial element in team training (Fanning and Gaba, 2007; Sawyer et al., 2016). It is through reflection that team members identify their learning gaps and develop strategies for improving them. From a teaching and learning perspective, the use of evidence, conveyed through appropriate visualisations, is critical to providing feedback to students and practitioners, particularly if they come from different backgrounds and perspectives, such as in multidisciplinary medical teams (Salik and Paige, 2020). The ultimate aim of this thesis was to provide evidence-based reflection tools from the analysis of large amounts of multimodal data to support reflection on the dynamics of embodied teamwork.

Through multimodal interfaces, this thesis has illustrated the potential of automatically generating evidence about team correct actions and procedures, positioning strategies, and nurses' stress levels. This evidence supports teachers in addressing the authentic pedagogical intentions they have when monitoring, assessing, and reflecting on embodied team activity. For instance, modelling and visualising team's time responses

(Sections 4.2.1 and 5.1 respectively) to critical situations facilitates students to reflect on **aspects they can improve in future scenarios** and **reduce misconceptions** about their outcomes (e.g., thinking that team responses were quick when they were not). Moreover, the analysis associated with the data vignette related to co-presence in interactional spaces (Section 5.2.1.1) facilitates rapid *comparison of team performance* in terms of **patient-centre care**, and **central role that the team leader** should play at critical moments. Similarly, automated detection of f-formations could be useful in distinguishing **team dynamics** (e.g., work delegation) or **detect clinical errors** that should be minimised by following effective collaborative practices. As illustrated in Section 5.2.2.1, the common error of not having another member of the team to validate the administration of the medication can be addressed in training (e.g., during the debriefing). Medication errors are still prevalent in professional nursing practice, and thousands of people die each year as a result of these errors (Gates et al., 2019), which makes this contribution highly significant in its potential to solve a widely recognised problem. Besides, automatically identifying how **teams use the space** during critical situations, and the transitions between such spaces (as illustrated in Section 5.2.3.1), could contribute to the recent interest in **mapping nurses' workflows** in emergency wardrooms with the purpose of improving clinical practice (Petrosoniak et al., 2019), or improving the architectural design of the wardrooms (Marini, 2019).

MMLA interfaces also support teachers reflections for the co-teaching scenario. The Classroom Dandelion exploration indicated that teachers can use it for **characterising effective ways to approach students** and **use classroom resources**. This is important since teachers have been reported to be often encouraged to co-teach in large open learning spaces without developing the ability to collaborate with peers and use the space effectively (Mackey et al., 2017). This makes it possible to identify the patterns that characterise effective co-teaching practice, as illustrated in the vignettes presented in Sections 6.5 and 6.6, and as suggested in emerging literature Montgomery and Akerson (2019). Thus, teachers appreciated the value of classroom dandelions to provide a quick understanding of their spatial behaviours, such as their interactions with classroom resources and students.

In addition to teachers and students, other educational stakeholders can also benefit from this research. For example, subject coordinators can use the MMLA interfaces to validate if there are possible improvements to incorporate into their planning. As documented by Mangaroska et al. (2020), analytics can be used to inform the learning design. The findings of this research suggested that the written reports (presented in

Section 6.3) can serve to **document performance of a unit** (Wiley et al., 2020) or to **revise the learning design** (e.g., reinforce the importance of doing specific procedures that the majority of groups missed). In addition, teachers considered the timeline as an important outcome for teachers to re-organise aspects of the simulation or the classroom settings.

In sum, team reflection commonly occurs under the guidance of a facilitator or teacher (Viller, 1991), but this can be highly challenging for teachers to provide. This is perhaps why teachers, students, and practitioners are increasingly recognising the added value of capturing objective evidence of embodied team activity and rendering it visible to support reflections after team simulation (Echeverria et al., 2019; Martinez-Maldonado et al., 2020a) or co-teaching sessions.

8.2.3 Visualisation guidance and storytelling for learning analytics

As detailed in Chapters 5, 6, and 7 Explanatory and Exploratory interfaces were generated as part of this thesis. **Explanatory interfaces** were detailed in Studies 1, 5 (Sections 5.1 and 6.3), and the proof of concept in Chapter 7, and these visualisations intentionally seek to reduce complexity by incorporating DS principles such as focussing users' attention on specific target features. I did find that using complex data (e.g., physiological data) introduce additional complexities in communicating student's data, such as complex visualisations (e.g., ENA representations in Section 6.2), which require additional training for users to make sense of them (see Study 4 in Section 6.2).

This research found that the explanatory MMLA interfaces had been positively evaluated by academics, teachers, and students. For example, the results presented in study 1 (Section 5.1.5) show that the enhancements using DS principles helped the students identify misconceptions, think about strategies to address the errors they made (Figures 5.5 and 5.6), and reflect on the levels of arousal (Figures 5.7 and 5.8) they may have experienced during the simulations. Students identified errors that they had missed prior to engaging in the proposed enhanced designs and reported, through interviews and a survey (see results in Figure 5.10). In general, DS enhancements **assisted students in focussing their attention on the expected learning goals** (see simulation 1 in Section 3.2.1.1 and simulation 2 in Section 3.2.1.2). This evidence supports the proposal that **DS principles can be designed to assist in communicating insights** found in the multimodal learning data, helping to address problems with visual LA products documented by other researchers Corrin and De Barba (2014); Matcha et al. (2019); Schwendimann et al. (2016). Teachers perceived positively salient aspects of the visual-narrative interfaces and **messages** included in the visual representations (Figures 6.11, 6.12, 6.13 and 6.14). The evaluation results made evident how a feature such as **colour** was used by teachers to visually group elements and interpret them as part of the same message (e.g., orange in text, points and shaded areas were used to communicate a group error). However, teachers particularly mentioned that certain parts of the message needed to be complemented, particularly time responses. In this sense, Chapter 7 demonstrated how the reference implementation can assist teachers in generating and modifying customised messages and visualisations, according to their own pedagogical intentions. Likewise, I acknowledge that additional categories of rules for interrogating the data could emerge according to the context, but the ones presented in this thesis (see Table 4.1) are generally applicable for this and other learning contexts.

Although the studies presented in this thesis were conducted in the context of complex, multimodal embodied team learning situations, there is no reason why a storytelling approach could not be implemented to aid in the interpretation of more conventional LA visualisations supporting non-collocated teamwork. For example, in his doctoral thesis Al-Doulat (2021) investigates how to support teacher/student sensemaking on student performance by automatically generating narrative, using structured data stored in online systems. Guiding students while interpreting their data is a feature that is missing in most current LA dashboards (according to the most recent review in (Matcha et al., 2019)). This thesis illustrated the specific case of making evidence on *errors*, *arousal*, and *spatial behaviours* available for computational analysis and scaffolding the interpretation of such data in a meaningful way.

Hence, the contribution of the explanatory MMLA interfaces should be seen as a set of instances of the extensive work that needs to be conducted in guiding students and teachers in interpreting multimodal learning data by aligning the visual representations with the learning goals and teacher's pedagogical intentions. The literature suggests that future work could involve extracting insights about other high-order features of learning in addition to errors, arousal, and spatial behaviours. For instance, there is interest in investigating learning strategies (e.g., (Matcha et al., 2019)), progress in achieving certain goals (e.g., (Ifenthaler and Yau, 2020)), or effectiveness of collaboration (e.g., (Er et al., 2021)).

In addition, **Exploratory MMLA interfaces** were generated and evaluated as part of the exploration of alternative ways to communicate aspects of embodied team activity to support reflection. Studies 2 (in Section 5.2), 3 (in Section 6.1), and 4 (in Section 6.2) detailed the different exploratory prototypes. The key role that data may have on teachers' sensemaking is a rapidly growing topic of interest in the LA field (Campos et al., 2021). However, turning human activity and behaviour into data points does not guarantee that educational stakeholders will be able to do anything practical with it (Selwyn and Gašević, 2020). For that reason, exploratory prototypes require an additional expert evaluation to validate how useful they can be in supporting learning or teaching purposes. As experts in their learning context, teachers are generally willing to participate in the complex visualisation exploration process to identify key insights about embodied team activity.

As suggested by Campos et al. (2021), there is a timely need for HCI and learning analytics communities to explore a broader range of instructional contexts and jointly enrich the typology of teacher sense making of various learning data. Motivated by this, studies 3 (in Section 6.1) and 4 (in Section 6.2) explored teachers' sensemaking of spatial data representations using SNA and the f-formation notion. Although spatial data can bring additional challenges for sensemaking because the vast amounts of sensor data can be both overwhelming and stripped of valuable contextual cues, teachers engage with the different visualisation and gain insights from them. This was done in two different contexts, a physics laboratory and a simulated hospital ward: two different educational contexts that both involve dynamic, embodied practice and require teachers' reflection and analysis on spatial behaviours.

However, this research also found that to make the most of the exploration process, teachers will require additional guidance. The additional support can be in the form of: i) explanations on how to read the visual interpretations; ii) provision of relevant/critical moments to explore (e.g., 2 minutes after the patient deterioration); iii) use of colour to differentiate roles (e.g., red and blue for main and secondary teachers); and iv) for the Classroom Dandelion, transitions from coordinates (points) to trajectories and rotations data (dandelion). Taking into account these guidance aspects, **teachers identified the characterisation of spatial (attentional) behaviours of teachers and students**. For instance, they differentiated monitoring or scanning the classroom behaviours versus remaining focused on a classroom resource or small groups of students (thanks to aspects in the design named, the spotlight metaphor and the density stacking, see Section 6.1). Teachers also identified and compared independent teams and teams seeking help. Additionally, teachers envisaged how they could appropriate visualisation into their regular teaching practice, the classroom dandelion mainly as a resource for self-inquiry, to train novice teachers, and the ENA representation for exploration of spatial patterns and to be discussed during the debriefing. This is of particular importance in higher education where lecturers and teaching assistants often lack basic pedagogical training (Grunspan et al., 2018) or feedback about spatial behaviours.

In summary, although teachers gained insights from exploratory prototypes, the results confirmed that teachers benefit from additional orientation cues, **they still needed guidance to disambiguate the interpretation of the spatial behaviours**. Explanatory prototypes are potentially more useful for being used to support teaching and learning, while exploratory prototypes are mainly useful for research purposes. In addition, both types of visualisation should be considered as evolving prototypes, because constant evaluation with educational stakeholders can reveal the needs for improvements in the design.

8.2.4 Implications for supporting LA design

The modelling approaches and the MMLA interfaces in this thesis were motivated by the current challenges faced by LA. The first challenge is the growing need for Human-Centred Design (HCD) methods to involve stakeholders in design processes. This is still relevant, as HCD processes are one of the topics of interest included in the 12th annual Learning Analytics and Knowledge Conference (LAK) call for contributions [1]. In fact, the concept of Human Centred Learning Analytics (HCLA) was recently coined to refer to the sub-community of LA researchers and practitioners interested in using the body of knowledge and practice from design communities, such as participatory design and co-design, in data-intensive educational contexts (Buckingham Shum et al., 2019b; Dimitriadis et al., 2022). From an HCLA perspective, this thesis contributes with: 1) modelling techniques that embed **teachers' assessment criteria** and **pedagogical intentions** in the multimodal model, detailed in Chapter 4; and 2) an elicitation methodology of **students' spatial behaviours expected by teacher**, detailed in Section 4.2.2.1. From an HCLA perspective, the modelling techniques in this thesis consider the voices of the teacher and provide tools that are aligned with the student's needs for reflection. Thus, the MMLA approach that I used is in fact a form of "deep HCLA" by incorporating the teacher's worldview into the feedback tool at multiple levels (guiding analysis and generating feedback).

Although HCD approaches can support the design of LA solutions that are tailored

[1] https://www.solaresearch.org/events/lak/lak22/general-call/

to the specific needs of stakeholders, it has also been argued that HCLA faces the novel situation that (unlike most HCI contexts) end-users are not experts in the task at hand. By definition, students have not mastered the domain of study and invariably know very little about the processes of learning in the formal sense (Buckingham Shum et al., 2019b). Thus, researchers and designers must discern when to accommodate students' views on what will help them learn and when to favour teachers' views. In developing this thesis, I considered both teachers and students perceptions on the prototypes. As experts and final users, teachers' explorations of the prototypes were valuable in the process of co-designing the visualisations (based on their pedagogical intentions) and identifying meaningful patterns in complex visual representations. As final users, the students' perceptions were valuable in validating key improvements to the prototypes and envisaging potential uses of the interfaces.

Another challenge of the LA community is the need to align LA design with the learning design. Very little design work or empirical work has demonstrated the visual means to make this alignment explicit in the design of student–facing interfaces (see the recent review in (Mangaroska and Giannakos, 2018)). This thesis contributes to addressing these gaps, expanding the work initiated by Echeverria et al. (2018b) who validated a DS approach, driven by the learning design, with teachers under controlled conditions. This thesis extends this work by 1) implementing six MMLA interfaces in real *in-the-wild* scenarios (see Chapters 5) and 6), and 2) proposing an architecture and reference implementation to automatically generate multimodal data stories (see Chapter 7). The careful alignment of data stories created from the analysis of student data, with the learning intentions of the activity, offers rich opportunities to support students and identify potential changes needed in the learning designs.

An additional challenge for LA is the need to examine aspects of the adoption of LA methods and approaches (Prieto et al., 2019). A potential damage from the generation of automated tools with lack of context is the reduction in the adoption of LA solutions in the classroom. This can be mitigated by engaging teachers and students in the design process. In this way, they can: i) decide what information is to be included in the learner model and interfaces (e.g., MMLA interfaces); and ii) participate in iterative prototyping, to gauge risks and increase the usefulness of prototypes. Teachers have been the most commonly involved stakeholders in the **co-design** studies, since the first studies in education such as that of Holstein et al. (2017), using **co-design techniques** to identify teachers' data needs using prototypes. This thesis keeps in the design loop both teachers and students, combining MMLA interfaces and **walkthroughs**, to validate its potential to support students' reflections and teaching and made future design decisions. This is similar to the work presented by Wise and Jung (2019), who use detailed interviews with teachers to inform design decisions for the teacher dashboard.

In sum, LA designers should enquire and include educational stakeholders to contribute to LA solutions in order to adapt them to specific needs and contexts. However, HCD approaches are not easy to implement; they require a close relationship with stakeholders and hard work to build and sustain trust in the LA solution. Although these approaches can be time-consuming, they are worth it in the sense that they often end up generating tools that support real learning and teaching needs.

8.3 Limitations and Remaining Challenges

The research in this thesis has provided well-grounded conceptual contributions that should advance research and practice in the MMLA, LA, CSCL, and CSCW communities. Nevertheless, this research has also faced several kinds of challenge, described below.

8.3.1 From a risk of over-interpretation perspective

Explanatory MMLA interfaces provide a "prescriptive" way to communicate insights from data. In doing so, the risk of over-interpretation is evident. As illustrated in the studies, teachers and students generally perceived the exploration of data stories as a positive experience. In study 1 (Section 5.1), students did not explicitly mention potential problems related to bias in the prototypes. However, students could have questioned how their arousal levels were determined or whether the errors made were correctly detected. To mitigate this risk, in the studies, each MMLA interface was carefully aligned with the teacher's learning intentions. However, exploratory MMLA interfaces such as the Classroom Dandelion, revealed several instances of potential over-interpretation of the classroom situations. Although teachers were often reminded that they were only seeing indoor positioning data, they attributed other meanings to these data in terms of task or social aspects (e.g., effective communication). This occurred mainly in nursing simulations (Figures 6.7 and 6.8), where teachers expect students to communicate effectively in teams.

Mostly, teachers' over-interpretations did coincide with what actually happened, based on the video, because students are meant to communicate and collaborate according to the learning task, and it is expected that they do so especially when they stand face-to-face for some time. However, the problem of over-interpretation is not exclusive to the dandelion diagrams, as it occurred also while teachers inspected the x and y coordinates and it has been reported as a widespread concern in the sector of learning analytics (Alhadad, 2016; Selwyn and Gašević, 2020). It follows that if multimodal data stories are to be embedded as either a teaching support tool or for professional development, it is critical (as with all instruments) to clarify their scope: explaining the rules used to detect errors or indicating that positioning data depict only spatial behaviours, and verbal or video data are needed to make stronger inferences about collaboration or effective communication.

Although the instances of over-interpretations were very low across the studies, explanatory MMLA interfaces can certainly introduce bias in the way annotations are written, how rules are created, and how accurately sensors capture data. This thesis concurs with more cautious view, since the validity and reliability required for automated grading (assessment embodied teamwork summatively) is significantly beyond the current maturity of this infrastructure. This thesis lowers the stakes by focusing on formative feedback, to provoke deeper reflection and dialogue, in which the human agents determine the ultimate meaning and consequences.

8.3.2 From the learning situation specificity

While clinical simulations and the science laboratory were authentic and reflect how nurses professionals are commonly trained and co-teaching behaviours, the data collected to generate the multimodal data stories were from a relatively small sample size of teams (e.g., 25 nursing students, 5 teams for study 5 6.3) performing a particular embodied team activity. Furthermore, although teachers were always interested in participating in the exploration of MMLA interfaces, some scenarios involved only a few teachers responsible for different classes (e.g., 4 in study 5 6.3). It is possible that other teachers could have reacted differently or were expected to evaluate specific team behaviours and that the design approach documented here may not transfer to other kinds of simulation exercise which requires the tracking of student activity that is too complex to automate. For instance, the spatial behaviours expected by the teachers (in Section 5.2) are context and task dependent, which means that they will change depending on the simulation. Consequently, some of the modelling techniques presented in Chapter 4 may not be applicable to address questions in new contexts. This limitation was recently documented in the literature review by Praharaj et al. (2021a) on co-located collaboration (CC) measures, as the authors found that CC is scenario dependent and the collaboration indicators can vary depending on the scenario, its goal, and context. This limitation is something the LA community should continue exploring.

8.3.3 From the adoption and integration of the MMLA interfaces as part of the learning design

The evidence from evaluations of the prototype interfacess (studies 1 5.1.5 and 5 6.3) is that nursing simulation debriefing sessions can be guided by using MMLA interfaces. Although the MMLA interfaces described in this thesis are already fully automated, and this thesis introduces a reference implementation for Healthcare Simulations as discussed in Chapter 7 (Sec. 7.2.5), they have not yet been used in debriefing sessions immediately after simulations, which will take place shortly, in the next iteration over the remainder of 2022. The architecture and reference implementations (in Sections 7.1 and 7.2) are major advances enabling such automation, and hence integration, of these MMLA interfaces into immediate debriefing sessions. I further envisage this automation enabling replay tools to support students in subsequent reflective assignments.

8.3.4 From the ethics and the risk of harmful surveillance perspective

O'Neil (2016) has documented the damage to teacher employment conditions from automated assessments based on poor data and black-box algorithms. The use of data in education should begin with the assumption that such data are commonly incomplete (Kitto et al., 2018). In their work, Martinez-Maldonado et al. (2020b) documented that teachers were against the use of positioning data to measure their performance, because data can easily be misinterpreted by others who may not be aware of the context and other sources of evidence needed to gain a rich understanding of the complexity of classroom practices. In line with this, this thesis encourages other researchers interested in using the MMLA interfaces (presented in Chapters 5 and 6) to use them only to support teachers and students for professional development purposes. Similarly, surveillance of activities and collection of classroom data should not harm student progress (Slade and Prinsloo, 2013). Therefore, this thesis cautions in the strongest terms against using visualisations of the sort presented here, with the purpose of summatively assessing the performance of teachers or students.

Prescriptive data stories can raise privacy and accountability concerns. Teachers explicitly stated that positioning data about themselves should be shared with peers for the purpose of helping others to learn from mistakes or "good" practice, in line with Martinez-Maldonado et al. (2020b)'s envisaged uses of classroom proxemics to support co-teaching. Students were not concerned about others looking at their data, which was the opposite of what teachers expected (as reported in study 1 5.1). Although privacy issues were raised for teachers and students, they also suggested ways to address some of these through pedagogical strategies and interface design features. For instance, teacher suggested the **de-identification** of the students, so that they should conduct the whole class reflections without naming individual students.

Ethic considerations are a high-priority topic in the LA community. The recent literature review by Yan et al. (2022) highlighted ethical challenges for MMLA interfaces to be scalable and sustainable. For example, equality concerns, as the technology used may have language limitations (e.g., only software in English) that can exclude specific regions or cultures.

8.3.5 From a practicality perspective

To scale up this thesis proposal, there are some considerations in terms of the technology and complexity of the installation. The software developed may not have identified specific salient aspects of teamwork (e.g., f-formations) due to abnormal values in the rotation data captured. These abnormalities were caused by some nurses adjusting the devices (e.g., waist bags containing their positioning trackers) and thus altering the multimodal data captured (e.g., rotation values). For some of the studies, these data were manually corrected after inspecting video footage of the simulations, but future work should consider this as a potentially disruptive factor for automatic analysis, particularly for in-the-wild settings.

Preparing the classroom for data collection still requires the physical installation of software, as reported by Yan et al. (2022), it remains a critical challenge to scale up MMLA innovations. An alternative to mitigate this challenge is to explore technology, which is easy to install and maintain (e.g., using video-based systems or beacons); for example, indoor positioning tracking systems are rapidly evolving and dropping in price (Yan et al., 2022). Additional research is needed to think about alternative technologies to make the installation more agile, and also strategies for data capture to increase accuracy and avoid abnormal values in in-the-wild settings.

8.4 Future Work

This section envisages several avenues for future work to build on the progress made in this thesis.

High-performance teamwork research. The reference implementation in Chap. 7 demonstrated the feasibility of implementing multimodal data storytelling. The implementation performed well for different simulations conducted at a different university, collecting data from 254 students. This is promising evidence that the implementation can scale up for other scenarios and larger datasets. I envisage that this advance opens the possibility of a **"science of teamwork" research infrastructure** for high-performance, embodied teamwork research. For the first time, it is possible to formalise hypotheses about multiple dimensions of embodied teamwork (using the Context Modeller), and evaluate these empirically through the automated generation of visual analytics.

Exploration of other modelling techniques and multimodal data. I would like to emphasise that teachers' expectations are context-dependent and task-dependent, which means that they will change depending on the simulation. Consequently, the designs presented in Section (see modelling techniques in Chapter 4) may not be applicable to easily identify errors, interpret arousal levels, or validate spatial behaviours in new contexts.

This thesis explored different social aspects of embodied team activity, including physical social presence (developed in Chapter 4.2.2.2), social arrangements (developed in Chapter 4.2.2.3), and scripted and emerging roles (developed in Chapter 4.2.1). Although the audio data was collected in the different studies, none of the modelling techniques or interfaces uses those data. However, I can see great potential in using audio and video to gain insight into effective communication. Exploring this and other modalities can complement stories to support students. Although in most teaching, it remains impractical to replay the entire video recording of classroom activity, these clearly contain rich information not captured in the proposed prototypes (e.g., dandelion diagrams), including the content of verbal interactions, body language, gaze, and the actions students and teachers perform. These enable deeper reflection on higher-order competencies that go beyond the analysis of the behaviours explored in this thesis (e.g., effective teaching, teamwork, and leadership). In addition to exploring other modalities, it would be of interest to explore other features of the modalities explored in this thesis. For example, although we used body positioning and orientation to generate positioning prototypes;

these are just two features of many other ways people interact in physical learning spaces. It also involves investigating the impact of modelling arousal data using alternative heuristics to convert peaks into categories or presenting the insights using different combinations of text narrative and visual enhancements. Or, combine different modalities to generate new and more informative learner data stories.

Exploration and evaluation of data visualisation design. This work should be seen as the first steps toward the further research needed to create interfaces that support the interpretation of the analytics outputs (Echeverria et al., 2018a; Wiley et al., 2020). Additionally, more experimental studies could be conducted to identify which specific elements of the LA interface drive students' visual attention (e.g., such as the study using eye-trackers conducted by Echeverria et al. (2018b)). Some immediate design improvements from the results are summarised in the next point.

Improvement of the design and evaluation of MMLA interfaces. Particularly in study 3 (see dandelion study 6.1), the dandelion diagrams were always explored by teachers after exploring the x-y coordinate data plots. The rationale was that, given the qualitative nature of the study, I wanted to identify variations in the sensemaking process by disclosing body orientation data after the coordinates. However, I recognise that the process of comparing dandelion diagrams with a visualisation of data points and asking open questions about them could have reduced opportunities for design innovation along new trajectories, as argued by Greenberg and Buxton (2008). Moreover, with a larger sample and a different study design, it would be possible to explore variations in the features of the dandelion diagram or other visualisation techniques representing the body orientation data for usability evaluation. This further exploration goes beyond the scope of this paper but is an avenue for future work to refine the design concepts introduced here.

Additional research on Explainable AI. The notion of stories to *explain* the analytical outcomes was introduced in Study 1 (see Section 5.1). The most questioned feature in the designs was the exposure of the semiformal "pseudocode" *if–then* rules. Nursing students, who rarely have programming skills, found these difficult to interpret. Consistent with the growing interest in explainable AI solutions (Wang et al., 2019), this points to the need to find ways to communicate to non-technical people how a particular insight was obtained from the data to encourage transparency and trust in LA systems. Based on the assumption that greater transparency encourages adoption, I argue that explainability can be considered at two levels: (1) explainability in the way feedback is communicated, for example, explainable visual analytics providing a clear narrative that explains what went well or wrong during a simulation experience, and why; and (2) explainability at the level of feedback generation, this is, how the machine picked a story. Current work in explainable AI tends to focus predominantly on the latter, as a way to provide transparency in automated decisions made by the machine (Hagras, 2018), however, in education, this is an area that still needs work, since the interactions between students and automated processes can be operationalised into learning opportunities (e.g., for students and teachers to understand the reasons behind the feedback that has been generated). For example, Echeverria et al. (2018b) presents an approach for explaining student data to drive teachers' interpretations of visual LA. This thesis introduces the context modeller for teachers to capture their pedagogical intentions into the system, but other ways to capture them can be explored in future work.

Exploration of other contexts. The mixed-method approximation presented in this thesis focused on small samples of authentic classes (that is, four educational contexts that feature two architectural designs in the classroom, simulations of health care, and science labs). Future work can explore the capabilities of the Multimodal Data Stories to support reflection in other educational disciplines (beyond Health and Science education) and educational levels (e.g., primary or secondary schools), and for other classroom architectures. I envisage the use of the Functional Architecture (Section 7.1) to expand the exploration and generation of other modalities and multimodal data stories.

Explanatory MMLA interfaces as part of the learning design. The application of DS is in its infancy and more exploratory work is needed before conducting longitudinal studies. In addition to this, the DS approach fits well in the learning contexts presented in this thesis, where simulations and science laboratories occur in a short period of time and in complex and special learning spaces. I chose not to artificially change the learning design of the activity for research purposes. Instead, I valued in-the-wild exploration of DS (in an authentic classroom setting). Therefore, this study focuses on this kind of authentic experience. This leaves open for future work the possibility of longitudinally focused research to understand the sustained use of DS and narrative tools in LA dashboards and reports.

8.5 Conclusion

Given the limitations of current visual analytics used in educational contexts, this thesis anticipates that approaches such as Data Storytelling will grow in importance to help students make the most of the new forms of feedback that are becoming possible. DS is a promising approach to facilitate such assistance and to augment teacher–led reflection. This thesis has presented the modelling techniques, design rationale, and expert critiques of multimodal learner data stories, in the context of three authentic nursing simulation classrooms and one science lab. The evidence from the results indicates that the affordances of each design assisted different teaching and learning tasks. The students envisaged a range of potential uses for the stories to support their reflection on teamwork practices to (i) recall the activity they performed, (ii) think about what can be improved, and (iii) gain new information or knowledge as a result of the reflection process.

From the teaching perspective, this thesis also reports that teachers envisaged multiple potential uses of such data (i) for them to reflect on their own teaching; (ii) to support the development of teaching skills of novice teachers and teaching assistants without strong pedagogical training; (iii) the provision of feedback to students engaged in developing spatial capabilities in healthcare; and (iv) to support further research assessing the design of learning spaces and in identifying effective workflows in both teacher and student teamwork.

The exciting potential of *teacher-driven* data stories is that once LA researchers/designers understand the affordances of different representations, they are not restricted to choosing just one to try and meet the diverse needs of different stakeholders and tasks, but can offer multiple views, as automatically generated, linked visualisations, placing the agency in the hands of teachers and students to switch between them as suits the context. Finally, this thesis has advanced the state of the art, moving beyond validating user interface prototypes, to a functional architecture and reference implementation of automated MMLA interfaces enhanced with DS principles for Healthcare Simulations. I hope that this work provides the conceptual and technical foundations for future work.

APPENDIX A

APPENDIX

Protocol interview used to elicit students' perceptions of prototypes presented in Section 5.1.

This interview will be run (13-15 of August) using the focus group technique.

Goal: Validate the potential value offered by a second version of the layered interface that Vanessa tested with the teacher.
NOTE: The protocol will be similar but the layers will now be slightly different (focused on one layer per "**expected critical actions**"), this is the validation with a different simulation, Adverse Reaction to antibiotics.

The Critical actions (sent by email and discussed with Carmen) are the following:

Expected Critical actions for this simulation

1. Perform an initial set of vital signs (after the lab tutor reads the handover)
2. Administer the IV antibiotic to the patient
3. Perform another set of vital signs after the patient complains of chest tightness
4. Stop the IV antibiotic after the patient complains of chest tightness and erythematous torso rash
5. Perform an ECG after the patient complains of chest tightness
6. Call the doctor to review the patient after stopping the IV antibiotics

TASK	ACTIVITIES	TIME
Introduction	Introduce yourselves. Explain that "*the goal of the session is to validate the potential value offered by a prototype in helping you to reflect on what went well and what didn't in your simulation last week.*"	2 minutes
Students preconception about the simulation (INDIVIDUAL)	For this, we want to start with a short activity for you to remember what you did. Pre-test Instructions: ❏ Show a template ❏ Students use **black/blue pen** ❏ Students write down critical actions they remember ❏ Turn the piece of paper down (avoid distractions)	1 minute Instructions 4 minutes Writing down

Figure A.1: Interview protocol to elicit students' perceptions of prototypes presented in Study 1

Exploring the **baseline timeline**(timeline without any multimodal layer) Instructions: Participants will explore the **baseline** timeline to understand the activity, the actions performed and the performance of the team using a think aloud protocol	❏ Record screen ❏ Think aloud protocol ❏ Show the timeline Can you identify your role in this timeline? **Key question (don't forget to ask):** ❏ **Do you think there were mistakes or actions that were correctly performed during the simulation? Which ones? Why?** Triggering questions if they don't say anything: ❏ Do you think the team worked together? ❏ Do you think the team reacted quickly? Why?	5 mins
Explanation of Critical Actions (Interviewer)	Explain to students about the **critical actions slide** (things that are important to be done as pointed by Carmen) Explain to participants how to move into different layers How to select different layers of information	2 mins

Figure A.2: Interview protocol to elicit students' perceptions of prototypes presented in Study 1

Exploring the interface Participants will **explore the interface** using a think aloud protocol. We are going to give them the total freedom of exploring the interface to observe their reactions	❑ Record screen ❑ Think aloud Questions to explore after they click on the layers: ❑ Do you think that these **layers add value** to the timeline I presented first for you to identify (vanilla o baseline timeline): 1. What went well and what didn't? 2. How do you react during the simulation, in terms of time and teamwork?	10 mins
Comparison of Layers	Question about comparing with their initial pre-conception ❑ How do you think you **performed according to the timeline** in the screen? Good? Poorly? Well? Triggering question: ❑ From a scale **from 1 to 10**, can assess your performance? AROUSAL	
Usefulness of Layers	❑ Which one do you prefer the most? WHY? How you read the first and the second one? Ask to turn the initial page around. Students are asked to add/correct the timeline if needed with a **red pen** ❑ Now that you know the critical actions, would you **change the draft** in your timeline or would you leave it the same? You can **add/cross out actions** that you missed in your draft. *How do you think you performed comparing the timeline in the screen with the one that you drafted in paper?* Have you been **comfortable wearing all this equipment** on you while in the simulation? ❑ Empatica? ❑ Belt bags? ❑ Video recording? ❑ How do you think this interface can be **used**? ❑ Help tutors during debrief? (yes/no) How?	

Figure A.3: Interview protocol to elicit students' perceptions of prototypes presented in Study 3

	❑ Help your team during debrief? (yes/no) How? ❑ As an individual tool? (yes/no) How? ❑ As an assessment tool? (yes/no) Why? ❑ Would you **share your data** with ...? Prompting hints ❑ your teacher? (for what purpose?) ❑ other students? (for what purpose?) RULES ❑ Present the rules in the presentation ❑ Do you **understand** the rules? ❑ Do you **agree** with the **rules** that were used to highlight team's performance (for example….[pick one or two rules based on what you highlighted as a mistake or the group reacting slow])? If **not, why**? And what would you **change**? (Question 5) ❑ Would you like to see the rules in the interface? **(Question 5a)** **Questionnaire** DEMOGRAPHIC QUESTIONNAIRE	

Figure A.4: Interview protocol to elicit students' perceptions of prototypes presented in Study 4

BIBLIOGRAPHY

Ahuja, K., Kim, D., Xhakaj, F., Varga, V., Xie, A., Zhang, S., Townsend, J. E., Harrison, C., Ogan, A., and Agarwal, Y. (2019). Edusense: Practical classroom sensing at scale. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 3(3). doi: https://doi.org/10.1145/3351229.

Aktinson, P. and Hammersley, M. (1998). Ethnography and participant observation. *Strategies of Qualitative Inquiry. Thousand Oaks: Sage*, pages 248–261.

Al-Doulat, A. (2021). *FIRST: Finding Interesting StoRies about STudents - An Interactive Narrative Approach to Explainable Learning Analytics*. PhD thesis, The University of North Carolina at Charlotte, North Carolina. url: https://www-proquest-com.ezproxy.lib.uts.edu.au/docview/2572575458/BC7ACEA917A14A76PQ/1?accountid=17095.

Alhadad, S. S. (2016). Attentional and cognitive processing of analytics visualisations: Can design features affect interpretations and decisions about learning and teaching. *Show Me The Learning. Proceedings ASCILITE*, pages 20–32. doi: https://doi.org/10.18608/jla.2018.52.5.

Alhadad, S. S. (2018). Visualizing data to support judgement, inference, and decision making in learning analytics: Insights from cognitive psychology and visualization science. *Journal of Learning Analytics*, 5(2):60–85. doi: https://doi.org/10.18608/jla.2018.52.5.

Ali, N. and Peebles, D. (2013). The effect of gestalt laws of perceptual organization on the comprehension of three-variable bar and line graphs. *Human Factors*, 55(1):183–203. doi: https://doi.org/10.1177/0018720812452592, PMID: 23516801.

Alzoubi, D., Kelley, J., Baran, E., Gilbert, S. B., Karabulut Ilgu, A., and Jiang, S. (2021). TeachActive Feedback Dashboard: Using Automated Classroom Analytics to Visualize

Pedagogical Strategies at a Glance. *Conference on Human Factors in Computing Systems - Proceedings*. doi: https://doi.org/10.1145/3411763.3451709.

Amulya, J. (2004). What is reflective practice. *Centre for Reflective Community Practice at the Massachusetts Insti-tute of Technology*. url: http://www.itslifejimbutnotasweknowit.org.uk/files/whatisreflectivepractice.pdf.

An, P., Bakker, S., Ordanovski, S., Paffen, C. L., Taconis, R., and Eggen, B. (2020). Dandelion diagram: Aggregating positioning and orientation data in the visualization of classroom proxemics. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI EA '20, page 1–8, New York, NY, USA. Association for Computing Machinery. doi: http://doi.org/10.1145/3334480.3382795.

An, P., Bakker, S., Ordanovski, S., Taconis, R., Paffen, C. L., and Eggen, B. (2019). Unobtrusively enhancing reflection-in-action of teachers through spatially distributed ambient information. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–14, New York, NY, USA. Association for Computing Machinery. doi: https://doi.org/10.1145/3290605.3300321.

Australian Nursing & Midwifery Federation (2020). Double-checking of medication prior to administration.

Bachour, K., Kaplan, F., and Dillenbourg, P. (2010). An interactive table for supporting participation balance in face-to-face collaborative learning. *IEEE Transactions on Learning Technologies*, 3(3):203–213. doi: http://dx.doi.org/10.1109/TLT.2010.18.

Barrera-Leon, L., Corno, F., and De Russis, L. (2020). Systematic variation of preattentive attributes to highlight relevant data in information visualization. In *2020 24th International Conference Information Visualisation (IV)*, pages 74–79. IEEE. doi: https://doi.org/10.1109/IV51561.2020.00022.

Bassiou, N., Tsiartas, A., Smith, J., Bratt, H., Richey, C., Shriberg, E., D'angelo, C., and Alozie, N. (2016). Privacy-Preserving Speech Analytics for Automatic Assessment of Student Collaboration. *INTERSPEECH 2016*, pages 888–892.

Basu, S., Jana, N., Bag, A., Mahadevappa, M., Mukherjee, J., Kumar, S., and Guha, R. (2015). Emotion recognition based on physiological signals using valence-arousal model. In *2015 Third International Conference on Image Information Processing (ICIIP)*, pages 50–55. IEEE. doi: https://doi.org/10.1109/ICIIP.2015.7414739.

Berragan, L. (2011). Simulation: An effective pedagogical approach for nursing? *Nurse Education Today*, 31(7):660–663. doi: `https://doi.org/10.1016/j.nedt.2011.01.019`.

Blikstein, P. (2013). Multimodal learning analytics. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, LAK '13, page 102–106, New York, NY, USA. Association for Computing Machinery. doi: `http://dx.doi.org/10.1145/2460296.2460316`.

Blikstein, P. and Worsley, M. (2018). Multimodal learning analytics and assessment of open–ended artifacts. In *Learning Analytics in Education*, chapter 4, page 89–112. Information Age Publishing, Charlotte, NC, USA, d. neimi, r. d. pea, b. saxberg, and r. e. clark, eds. edition.

Bodily, R. and Verbert, K. (2017a). Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies*, PP:1–1. doi: `https://doi.org/10.1109/TLT.2017.2740172`.

Bodily, R. and Verbert, K. (2017b). Trends and issues in student-facing learning analytics reporting systems research. In *7th LAK Conference*, pages 309–318, New York, USA. ACM Press. doi: `https://doi.org/10.1145/3027385.3027403`.

Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Res. Physiol.*, 3(2):77–101. doi: `https://doi.org/10.1191/1478088706qp063oa`.

Braun, V. and Clarke, V. (2012). Thematic analysis. In Cooper, H., Camic, P., Long, D., Panter, A., Rindskopf, D., and Sher, K., editors, *APA handbook of research methods in psychology, Vol. 2.*, page 57–71. American Psychological Association, Washington, USA. doi: `https://psycnet.apa.org/doi/10.1037/13620-004`.

Britton, L. and Anderson, K. (2010). Peer coaching and pre-service teachers: Examining an underutilised concept. *Teaching and Teacher Education - TEACH TEACH EDUC*, 26:306–314. doi: `http://dx.doi.org/10.1016/j.tate.2009.03.008`.

Brooks, C., Carroll, A., Gillies, R. M., and Hattie, J. (2019). A matrix of feedback for learning. *Australian Journal of Teacher Education (Online)*, 44(4):14–32. doi: `https://search.informit.org/doi/10.3316/ielapa.329362104776530`.

Brudy, F., Suwanwatcharachat, S., Zhang, W., Houben, S., and Marquardt, N. (2018). Eagleview: A video analysis tool for visualising and querying spatial interactions of people and devices. In *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*, ISS '18, page 61–72, New York, NY, USA. Association for Computing Machinery. doi: http://dx.doi.org/10.1145/3279778.3279795.

Buckingham Shum, S., Arastoopour Irgens, G., Moots, H., Phillips, M., Shah, M., Vega, H., and Wooldridge, A. (2021). Participatory Quantitative Ethnography. In *Symposium, Proc. 3rd International Conference on Quantitative Ethnography.*, Online. url: https://idealab.sites.clemson.edu/papers/ICQE21_pqe.pdf.

Buckingham Shum, S., Echeverria, V., and Martinez-Maldonado, R. (2019a). The multimodal matrix as a quantitative ethnography methodology. In Eagan, B., Misfeldt, M., and Siebert-Evenstone, A., editors, *Advances in Quantitative Ethnography*, pages 26–40, Cham. Springer International Publishing. doi: https://doi.org/10.1007/978-3-030-33232-7_3.

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

Campos, F. C., Ahn, J., DiGiacomo, D. K., Nguyen, H., and Hays, M. (2021). Making sense of sensemaking: Understanding how k–12 teachers and coaches react to visual analytics. *Journal of Learning Analytics*, pages 1–21. doi: https://doi.org/10.18608/jla.2021.7113.

Cannon-Bowers, J., Tannenbaum, S., Salas, E., and Volpe, C. (1995). Defining competencies and establishing team training requirements. *Team Effectiveness and Decision Making in Organizations*, pages 333–380. url: https://www.libs.uga.edu/reserves/docs/scans/defining%20competencies.pdf.

Card, S. (2004). Keynote address: From information visualization to sensemaking: Connecting the mind's eye to the mind's muscle. In *IEEE Symposium on Information Visualization*, pages xii–xii.

Carvalho, P. V., Vidal, M. C., and de Carvalho, E. F. (2007). Nuclear power plant communications in normative and actual practice: A field study of control room operators' communications. *Human Factors and Ergonomics in Manufacturing*, 17(1):43–78.

Ceneda, D., Gschwandtner, T., May, T., Miksch, S., Schulz, H. J., Streit, M., and Tominski, C. (2017). Characterizing guidance in Visual Analytics. *IEEE Trans. Visualization & Comput. Graph.*, 23(1):111–120. doi: 10.1109/TVCG.2016.2598468.

Cha, J. S., Athanasiadis, D., Anton, N. E., Stefanidis, D., and Yu, D. (2021). Measurement of Nontechnical Skills During Robotic-Assisted Surgery Using Sensor-Based Communication and Proximity Metrics. *JAMA Network Open*, 4(11):e2132209–e2132209.

Chen, Q., Li, Z., Pong, T.-C., and Qu, H. (2019). Designing narrative slideshows for learning analytics. In *2019 IEEE Pacific Visualization Symposium (PacificVis)*, pages 237–246.

Chin, H., Cheong, C., and Taib, F. (2017). Instructional proxemics and its impact on classroom teaching and learning. *International Journal of Modern Languages And Applied Linguistics*, 1:69.

Chng, E., Seyam, M. R., Yao, W., and Schneider, B. (2022). Toward capturing divergent collaboration in makerspaces using motion sensors. *Information and Learning Sciences*.

Ciolek, T. M. (1983). The proxemics lexicon: A first approximation. *Journal of Nonverbal Behavior*, 8(1):55–79. doi: `https://doi.org/10.1007/BF00986330`.

Cohen, S. G. and Bailey, D. E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management*, 23(3):239–290.

Conole, G. (2012). *Designing for learning in an open world*, volume 4. Springer Science & Business Media.

Cooper, A. B. and Tisdell, E. J. (2020). Embodied aspects of learning to be a surgeon. *Medical teacher*, 42(5):515–522.

Corrin, L. and De Barba, P. (Nov. 23–36, 2014). Exploring students' interpretation of feedback delivered through learning analytics dashboards. In *Proc. 31st Annu. Conf. Australasian Society for Computers in Learning in Tertiary Education (ASCILITE'14)*, page 629–633, Dunedin, New Zealand.

Creswell, J. W. (2014). *A concise introduction to mixed methods research*. SAGE publications.

Cristani, M., Paggetti, G., Vinciarelli, A., Bazzani, L., Menegaz, G., and Murino, V. (2011). Towards Computational Proxemics: Inferring Social Relations from Interpersonal Distances. In *2011 IEEE Third Int'l Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third Int'l Conference on Social Computing*, pages 290–297. IEEE.

Cross, N. and Clayburn Cross, A. (1995). Observations of teamwork and social processes in design. *Design Studies*, 16(2):143–170.

Cuff, R. N. (1980). On casual users. *Int. J. Man-Mach. Stud.*, 12(2):163–187. doi: https://doi.org/10.1016/S0020-7373(80)80016-2.

Dael, N., Bianchi-Berthouze, N., Kleinsmith, A., and Mohr, C. (2015). Measuring body movement: Current and future directions in proxemics and kinesics. In *APA handbook of nonverbal communication.*, pages 551–587. American Psychological Association.

Dafoulas, G., Cardoso Maia, C., Ali, A., and Augusto, J. (2018). Collecting sensor-generated data for assessing teamwork and individual contributions in computing student teams. *EDULEARN18 Proceedings*, 1:11156–11162. doi: http://doi.org/10.21125/edulearn.2018.2759.

Davidsen, J. and Ryberg, T. (2017). "This is the size of one meter": Children's bodily-material collaboration. *International Journal of Computer-Supported Collaborative Learning*, 12(1):65–90.

de Lima, Y. O. and de Souza, J. M. (2017). The future of work: Insights for cscw. *2017 IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pages 42–47.

Desjarlais, M. and Smith, P. (2011). A comparative analysis of reflection and self-assessment. *International Journal of Process Education*, 3(1):3–18.

Desmond, P. A. and Hoyes, T. W. (1996). Workload variation, intrinsic risk and utility in a simulated air traffic control task: Evidence for compensatory effects. *Safety Science*, 22:87–101.

Dewey, J. (1986). Experience and education. In *The educational forum*, volume 50, pages 241–252. Taylor & Francis.

Dewey, J. (1997). *How we think*. Dover Publications.

Di Mitri, D., Schneider, J., and Drachsler, H. (2021). Keep Me in the Loop: Real-Time Feedback with Multimodal Data. *International Journal of Artificial Intelligence in Education*.

Di Mitri, D., Schneider, J., Klemke, R., Specht, M., and Drachsler, H. (2019). Read between the lines: An annotation tool for multimodal data for learning. In *Proceedings* *of the 9th International Conference on Learning Analytics & Knowledge*, LAK19, page 51–60, New York, NY, USA. Association for Computing Machinery. doi: `https://doi.org/10.1145/3303772.3303776`.

Di Mitri, D., Schneider, J., Specht, M., and Drachsler, H. (2018). From signals to knowledge: A conceptual model for multimodal learning analytics. *Journal of Computer Assisted Learning*, 34(4):338–349.

Diamant, E. I., Fussell, S. R., and Lo, F.-l. (2008). Where did we turn wrong? unpacking the effect of culture and technology on attributions of team performance. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, CSCW '08, page 383–392, New York, NY, USA. Association for Computing Machinery.

Dich, Y., Reilly, J., and Schneider, B. (2018). Using physiological synchrony as an indicator of collaboration quality, task performance and learning. In Penstein Rosé, C., Martínez-Maldonado, R., Hoppe, H. U., Luckin, R., Mavrikis, M., Porayska-Pomsta, K., McLaren, B., and du Boulay, B., editors, *Artificial Intelligence in Education*, pages 98–110, Cham. Springer International Publishing.

Dickinson, T. L. and McIntyre, R. M. (1997). A conceptual framework for teamwork measurement. In *Team performance assessment and measurement: Theory, methods, and applications.*, Series in applied psychology., pages 19–43. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US.

Dimitriadis, Y., Martinez-Maldonado, R., Sarmiento, J. P., Campos, F., Santos, P., and Barreiros, C. (2022). Third international workshop on human centred learning analytics. *Companion Proceedings of the 12th*, page 189.

D'Mello, S. K., Bosch, N., and Chen, H. (2018). Multimodal-multisensor affect detection. In *The Handbook of Multimodal-Multisensor Interfaces: Signal Processing, Architectures, and Detection of Emotion and Cognition-Volume 2*, pages 167–202. Association for Computing Machinery and Morgan & Claypool. doi: `https://doi.org/10.1145/3107990.3107998`.

Domínguez, F., Ochoa, X., Zambrano, D., Camacho, K., and Castells, J. (2021). Scaling and Adopting a Multimodal Learning Analytics Application in an Institution-Wide Setting. *IEEE Transactions on Learning Technologies*, 14(3):400–414.

Dumais, S., Jeffries, R., Russell, D. M., Tang, D., and Teevan, J. (2014). Understanding user behavior through log data and analysis. *Ways of Knowing in HCI*, pages 349–372. doi: https://doi.org/10.1007/978-1-4939-0378-8_14.

Dykes, B. (2015). Data storytelling: What it is and how it can be used to effectively communicate analysis results. *Applied Marketing Analytics*.

Echeverria, V., Martinez-Maldonado, R., and Buckingham Shum, S. (2019). Towards Collaboration Translucence: Giving Meaning to Multimodal Group Data. *Proceedings of the CHI*, pages 39:1—-39:16.

Echeverria, V., Martinez-Maldonado, R., Granda, R., Chiluiza, K., Conati, C., and Buckingham Shum, S. (2018a). Driving data storytelling from learning design. In *ACM*, pages 131–140. Association for Computing Machinery.

Echeverria, V., Martinez-Maldonaldo, R., Buckingham Shum, S., Chiluiza, K., Granda, R., and Conati, C. (2018b). Exploratory versus explanatory visual learning analytics: Driving teachers' attention through educational data storytelling. *J. Learning Analytics*, 5(3):72–97. doi: 10.18608/jla.2018.53.6.

Er, E., Dimitriadis, Y., and Gašević, D. (2021). Collaborative peer feedback and learning analytics: Theory-oriented design for supporting class-wide interventions. *Assessment & Eval. Higher Educ.*, 46(2):169–190. doi: 10.1080/02602938.2020.1764490.

Erickson, T. and Kellogg, W. A. (2000). Social translucence: An approach to designing systems that support social processes. *ACM Trans. Comput.-Hum. Interact.*, 7(1):59–83.

Eyal, L. and Gil, E. (2022). *Hybrid Learning Spaces — A Three-Fold Evolving Perspective*, pages 11–23. Springer International Publishing, Cham. doi: https://doi.org/10.1007/978-3-030-88520-5_2.

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

Fernandez-Nieto, G., An, P., Zhao, J., Buckingham Shum, S., and Martinez-Maldonado, R. (2022). Classroom dandelions: Visualising participant position, trajectory and body orientation augments teachers' sensemaking. In *CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA. Association for Computing Machinery.

Fernandez-Nieto, G., Kitto, K., and Martinez-Maldonado, R. (2019). Four challenges in crafting multimodal collaboration analytics for non-data experts. In *In CSCL 2019, Collaboration Analytics*, page 4. note: https://collaborationanalytics.files.wordpress.com/2019/06/submission-7-fernandez.pdf.

Fernandez-Nieto, G., Martinez-Maldonado, R., Echeverria, V., Kitto, K., An, P., and Buckingham Shum, S. (2021a). What Can Analytics for Teamwork Proxemics Reveal About Positioning Dynamics In Clinical Simulations? *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1):1–24.

Fernandez-Nieto, G. M., Echeverria, V., Buckingham Shum, S., Mangaroska, K., Kitto, K., Palominos, E., Axisa, C., and Martinez-Maldonado, R. (2021b). Storytelling With Learner Data: Guiding Student Reflection on Multimodal Team Data. *IEEE Transactions on Learning Technologies*, pages 1–14.

Fernandez Nieto, G. M., Kitto, K., Buckingham Shum, S., and Martinez-Maldonado, R. (2022). Beyond the learning analytics dashboard: Alternative ways to communicate student data insights combining visualisation, narrative and storytelling. In *LAK22: 12th International Learning Analytics and Knowledge Conference*, LAK22, page 219–229, New York, NY, USA. Association for Computing Machinery.

Fernandez-Nieto, G. M., Martinez-Maldonado, R., Kitto, K., and Buckingham Shum, S. (2021c). Modelling spatial behaviours in clinical team simulations using epistemic network analysis: Methodology and teacher evaluation. In *LAK21: 11th International Learning Analytics and Knowledge Conference*, LAK21, page 386–396, New York, NY, USA. Association for Computing Machinery.

Figueiras, A. (2014). Narrative visualization: A case study of how to incorporate narrative elements in existing visualizations. In *2014 18th International Conference on Information Visualisation*, pages 46–52.

Fiore, S. M., Wiltshire, T. J., Lobato, E. J. C., Jentsch, F. G., Huang, W. H., and Axelrod, B. (2013). Toward understanding social cues and signals in human–robot interaction: effects of robot gaze and proxemic behavior. *Frontiers in Psychology*, 4:859.

Fitzpatrick, G. and Ellingsen, G. (2013). A review of 25 years of cscw research in healthcare: contributions, challenges and future agendas. *Computer Supported Cooperative Work (CSCW)*, 22(4-6):609–665.

Fletcher, T. D. and Major, D. A. (2017). The Effects of Communication Modality on Performance and Self-Ratings of Teamwork Components. *Journal of Computer-Mediated Communication*, 11(2):557–576.

Forsyth, D. R. (2018). *Group dynamics*. Cengage Learning.

Forsyth, D. R. and Diederich, L. T. (2014). Group dynamics and development. *Handbook of group counseling and psychotherapy*, pages 34–45.

Fransen, J., Weinberger, A., and Kirschner, P. (2013). Team effectiveness and team development in cscl. *Educational Psychologist*, 48:9–24.

Fuglsang, S., Bloch, C. W., and Selberg, H. (2022). Simulation training and professional self-confidence: A large-scale study of third year nursing students. *Nurse Education Today*, 108:105175.

Gates, P. J., Baysari, M. T., Mumford, V., Raban, M. Z., and Westbrook, J. I. (2019). Standardising the classification of harm associated with medication errors: the harm associated with medication error classification (hamec). *Drug safety*, 42(8):931–939.

Gersick, C. J. (1988). Time and transition in work teams: Toward a new model of group development. *Academy of Management journal*, 31(1):9–41.

Gloor, P., Grippa, F., Putzke, J., Lassenius, C., Fuehres, H., Fischbach, K., and Schoder, D. (2012). Measuring social capital in creative teams through sociometric sensors. *Int. J. of Organisational Design and Engineering*, 2:380 – 401.

Goldsberry, K. (2012). Courtvision: New visual and spatial analytics for the nba. In *2012 MIT Sloan sports analytics conference*, volume 9, pages 12–15.

Goodyear, P. (1991). *Teaching knowledge and intelligent tutoring*. Intellect Books.

Goodyear, P. and Carvalho, L. (2014). The architecture of productive learning networks. *Abingdon: Routledge*.

Goodyear, P. and Carvalho, L. (2016). Activity centred analysis and design in the evolution of learning networks. In *Proceedings of the 10th International Conference on Networked Learning*, pages 218–225.

Goodyear, P., Carvalho, L., and Yeoman, P. (2021). Activity-Centred Analysis and Design (ACAD): Core purposes, distinctive qualities and current developments. *Educational*

Technology Research and Development, 69(2):445–464. doi: https://doi.org/10.1007/S11423-020-09926-7.

Greenberg, S. and Buxton, B. (2008). Usability evaluation considered harmful (some of the time). In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 111–120.

Grunspan, D. Z., Kline, M. A., and Brownell, S. E. (2018). The lecture machine: A cultural evolutionary model of pedagogy in higher education. *CBE—Life Sciences Education*, 17(3):es6.

Guo, L., Yin, X., Zeng, X., Tang, L., and Zhang, D. (2022). Nursing practice combined with scene simulation teaching mode in standardized training and teaching for newly recruited nurses in post anesthesia care unit. *Iranian Journal of Public Health*, 51(4):831–838.

Haataja, E., Malmberg, J., and Järvelä, S. (2018). Monitoring in collaborative learning: Co-occurrence of observed behavior and physiological synchrony explored. *Computers in Human Behavior*, 87:337–347.

Hagman, J., Deshler, J., and Speer, N. (2016). Supporting institutional change: A two-pronged approach related to graduate teaching assistant professional development. In *Proceedings of the 19th Annual Conference on Research in Undergraduate Mathematics Education*.

Hagras, H. (2018). Toward human-understandable, explainable AI. *Comput. J.*, 51(9):28–36. doi: 10.1109/MC.2018.3620965.

Hall, E. T. (1963). A system for the notation of proxemic behavior. *American Anthropologist*, 65(5):1003–1026.

Hall, E. T. (1966a). *The hidden dimension*, volume 609. Garden City, NY: Doubleday.

Hall, E. T. (1966b). Proxemic Theory. *CSISS Classics*.

Hall, E. T., Birdwhistell, R. L., Bock, B., Bohannan, P., Diebold, A. R., Durbin, M., Edmonson, M. S., Fischer, J. L., Hymes, D., Kimball, S. T., La Barre, W., McClellan, J. E., Marshall, D. S., Milner, G. B., Sarles, H. B., Trager, G. L., and Vayda, A. P. (1968). Proxemics [and Comments and Replies]. *Current Anthropology*, 9(2/3):83–108.

Hall, K. L., Vogel, A. L., Huang, G. C., Serrano, K. J., Rice, E. L., Tsakraklides, S. P., and Fiore, S. M. (2018). The science of team science: A review of the empirical evidence and research gaps on collaboration in science. *American Psychologist*, 73(4):532.

Hansen, C. and Wasson, B. (2016). Teacher inquiry into student learning: - the tisl heart model and method for use in teachers' professional development. *Nordic Journal of Digital Literacy*, 11:24–49.

Harvey, D. (2005). Social Processes And Spatial Form: An Analysis Of The Conceptual Problems Of Urban Planning. *Papers in Regional Science*, 25(1):47–69.

Hassenzahl, M. and Ullrich, D. (2007). To do or not to do: Differences in user experience and retrospective judgments depending on the presence or absence of instrumental goals. *Interacting with Computers*, 19:429–437.

Hattie, J. A. and Yates, G. C. (2014). Using feedback to promote learning. *Applying science of learning in education: Infusing psychological science into the curriculum*, pages 45–58.

Haythornthwaite, C. (1996). Social network analysis: An approach and technique for the study of information exchange. *Library and Information Science Research*, 18(4):323–342.

Helyer, R. (2015). Learning through reflection: the critical role of reflection in work-based learning (wbl). *Journal of Work-Applied Management*, 7:15–27.

Holstein, K., McLaren, B. M., and Aleven, V. (2017). Spacle: Investigating learning across virtual and physical spaces using spatial replays. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, LAK '17, page 358–367, New York, NY, USA. Association for Computing Machinery.

Holtzblatt, K. (2007). Contextual design. In *The human-computer interaction handbook: Contextual Design*, pages 975–990. CRC press.

Ifenthaler, D. and Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educ. Technol. Res. Develop.*, 68(4):1961–1990. url: `http://link.springer.com/10.1007/s11423-020-09788-z`.

Ingvarson, L. and Rowe, K. (2008). Conceptualising and evaluating teacher quality: Substantive and methodological issues. *Australian Journal of Education*, 52(1):5–35.

Ivankova, N. and Stick, S. (2007). Students' persistence in a distributed doctoral program in educational leadership in higher education: A mixed methods study. *Research in Higher Education*, 48:93–135.

Jetter, H.-C., Reiterer, H., and Geyer, F. (2014). Blended interaction: Understanding natural human-computer interaction in post-wimp interactive spaces. *Personal Ubiquitous Comput.*, 18(5):1139–1158.

Johnson, D., Sutton, P., and Poon, J. (2000). Face-to-Face vs. CMC: Student communication in a technologically rich learning environment. In *17th Annual Conference of the Australian Society for Computers in Learning in Tertiary Education: ASCILITE 2000*, pages 1–15, Coffs Harbour.

Johnson, R. B., Onwuegbuzie, A. J., and Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2):112–133.

Jones, F. H., Jones, P., and Jones, J. L. T. (2007). *Tools for teaching: Discipline, instruction, motivation*. Fredjones. com.

Kaye, D. (2003). *Loosely Coupled: The Missing Pieces of Web Services*. RDS Press.

Kelly, M., Ellaway, R., Scherpbier, A., King, N., and Dornan, T. (2019). Body pedagogics: embodied learning for the health professions. *Medical education*, 53(10):967–977.

Kendon, A. (1976). The f-formation system: The spatial organization of social encounters. *Man-Environment Systems*, 6(01):1976.

Khosravi, H., Kay, J., Sadiq, S., Buckingham Shum, S., Martinez-Maldonado, R., Knight, S., Chen, G., Tsai, Y.-S., Conati, C., and Gasevic, D. (2022). Explainable artificial intelligence in education. *Computers & Education: Artificial Intelligence*, 3(February 2022):1–31.

Kim, T., Chang, A., Holland, L., and Pentland, A. S. (2008). Meeting mediator: enhancing group collaborationusing sociometric feedback. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pages 457–466.

Kitto, K., Shum, S. B., and Gibson, A. (2018). Embracing imperfection in learning analytics. In *Proceedings of the 8th international conference on learning analytics and knowledge*, pages 451–460.

Knaflic, C. N. (2017). *Storytelling With Data: A Data Visualization Guide for Business Professionals*. John Wiley & Sons, New Jersey, 12 edition.

Knight, S., Gibson, A., and Shibani, A. (2020). Implementing learning analytics for learning impact: Taking tools to task. *The Internet and Higher Education*, 45:100729.

Kurihara, K., Goto, M., Ogata, J., Matsusaka, Y., and Igarashi, T. (2007). Presentation sensei: A presentation training system using speech and image processing. In *Proceedings of the 9th International Conference on Multimodal Interfaces*, ICMI '07, page 358–365, New York, NY, USA. Association for Computing Machinery.

Landherr, A., Friedl, B., and Heidemann, J. (2010). A Critical Review of Centrality Measures in Social Networks. *Business & Information Systems Engineering*, 2(6):371–385.

Lee, B., Lee, M., Zhang, P., Tessier, A., Saakes, D., and Khan, A. (2021). Socio-spatial comfort: Using vision-based analysis to inform user-centred human-building interactions. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW3):1–33.

Lei, Y.-Y., Zhu, L., Sa, Y. T. R., and Cui, X.-S. (2022). Effects of high-fidelity simulation teaching on nursing students' knowledge, professional skills and clinical ability: A meta-analysis and systematic review. *Nurse Education in Practice*, 60:103306.

Leinonen, P., Järvelä, S., and Häkkinen, P. (2005). Conceptualizing the awareness of collaboration: A qualitative study of a global virtual team. *Computer Supported Cooperative Work (CSCW)*, 14(4):301–322.

Lichtenauer, J., Shen, J., Valstar, M., and Pantic, M. (2011). Cost-effective solution to synchronised audio-visual data capture using multiple sensors. *Image and Vision Computing*, 29(10):666–680.

Lubold, N. and Pon-Barry, H. (2014). Acoustic-prosodic entrainment and rapport in collaborative learning dialogues. *MLA 2014 - Proceedings of the 2014 ACM Multimodal Learning Analytics Workshop and Grand Challenge, Co-located with ICMI 2014*, pages 5–12. doi: `http://dx.doi.org/10.1145/2666633.2666635`.

Mackey, J., O'Reilly, N., Fletcher, J., and Jansen, C. (2017). What do teachers and leaders have to say about co-teaching in flexible learning spaces? *Journal of Educational Leadership, Policy and Practice*.

Maltese, A. V., Harsh, J. A., and Svetina, D. (2015). Data visualization literacy: Investigating data interpretation along the novice—expert continuum. *Journal of College Science Teaching*, 45(1):84–90. doi: https://doi.org/10.2307/43631889.

Mangaroska, K. and Giannakos, M. N. (2018). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Trans. Learn. Technol.*, 12(4):516–534. doi:10.1109/TLT.2018.2868673.

Mangaroska, K., Sharma, K., Gasevic, D., and Giannakos, M. (2020). Multimodal learning analytics to inform learning design: Lessons learned from computing education. *Journal of Learning Analytics*, 7(3):79–97.

Mariani, B. and Doolen, J. (2016). Nursing Simulation Research: What Are the Perceived Gaps? *Clinical Simulation in Nursing*, 12(1):30–36.

Marini, G. (2019). Towards indoor localisation analytics for modelling flows of movements. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, pages 377–382.

Marquardt, N. and Greenberg, S. (2015). Proxemic interactions: From theory to practice. *Synthesis Lectures on Human-Centered Informatics*, 8(1):1–199.

Marshall, P., Rogers, Y., and Pantidi, N. (2011). Using f-formations to analyse spatial patterns of interaction in physical environments. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*, CSCW '11, page 445–454, New York, NY, USA. Association for Computing Machinery.

Martinec, R. (2001). Interpersonal resources in action. *Semiotica*, 2001(135):117–145.

Martinez-Maldonado, R., Echeverria, V., Elliott, D., Axisa, C., Power, T., and Buckingham Shum, S. (in press). Designing translucent learning analytics with teachers: An elicitation process. *Interactive Learn. Environ.* doi: `https://doi.org/10.1080/10494820.2019.1710541`.

Martinez-Maldonado, R., Echeverria, V., Fernandez-Nieto, G., and Shum, S. B. (2020a). From Data to Insights: A Layered Storytelling Approach for Multimodal Learning Analytics. In *CHI '20*, page 15.

Martinez-Maldonado, R., Echeverria, V., Mangaroska, K., Shibani, A., Fernandez-Nieto, G., Schulte, J., and Buckingham Shum, S. (2021). Moodoo the tracker: Spatial classroom analytics for characterising teachers' pedagogical approaches. *International Journal of Artificial Intelligence in Education*, pages 1–27.

Martinez-Maldonado, R., Kay, J., and Yacef, K. (2013). An Automatic Approach for Mining Patterns of Collaboration around an Interactive Tabletop. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7926 LNAI:101–110.

Martinez-Maldonado, R., Mangaroska, K., Schulte, J., Elliott, D., Axisa, C., and Buckingham Shum, S. (2020b). Teacher tracking with integrity: What indoor positioning can reveal about instructional proxemics. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 4(1). doi: 10.1145/3381017.

Matcha, W., Gašević, D., Uzir, N. A., Jovanović, J., and Pardo, A. (Mar. 4-8, 2019). Analytics of learning strategies: Associations with academic performance and feedback. In *Proc. 9th Learning Analytics & Knowledge Conf. (LAK'19)*, pages 461–470, Tempe, Arizona, USA. ACM. doi: 10.1145/3303772.3303787.

Matzen, L. E., Haass, M. J., Divis, K. M., and Stites, M. C. (2017). Patterns of attention: How data visualizations are read. In Schmorrow, D. D. and Fidopiastis, C. M., editors, *Augmented Cognition. Neurocognition and Machine Learning*, pages 176–191, Cham. Springer International Publishing.

McDonald, N., Schoenebeck, S., and 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):1–23. doi:10.1145/3359174.

Melo, G. M. d., Rebouças, C. B. d. A., Cardoso, M. V. L. M. L., and Farias, L. M. (2013). Nursing team communication with regard pain in newborns: a descriptive study. *Brazilian Journal of Nursing*.

Mentis, H. M., O'Hara, K., Sellen, A., and Trivedi, R. (2012). Interaction proxemics and image use in neurosurgery. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 927–936.

Meyer, M. and Dykes, J. (2019). Criteria for rigor in visualization design study. *IEEE transactions on visualization and computer graphics*, 26(1):87–97.

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

Mondada, L. (2013). Interactional space and the study of embodied talk-in-interaction. *Space in language and linguistics: Geographical, interactional and cognitive perspectives*, pages 247–275.

Montanari, A., Tian, Z., Francu, E., Lucas, B., Jones, B., Zhou, X., and Mascolo, C. (2018). Measuring Interaction Proxemics with Wearable Light Tags. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1):1–30.

Montgomery, M. and Akerson, A. (2019). Facilitating collaboration through a co-teaching field experience. *Networks: An Online Journal for Teacher Research*, 21.

Moore, J., Goffin, P., Meyer, M., Lundrigan, P., Patwari, N., Sward, K., and Wiese, J. (2018). Managing in-home environments through sensing, annotating, and visualizing air quality data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2(3).

Mor, Y., Ferguson, R., and Wasson, B. (2015). Editorial: Learning design, teacher inquiry into student learning and learning analytics: A call for action: Learning design, tisl and learning analytics. *British Journal of Educational Technology*, 46:221–229.

Moreira, A. G. M., De Azevedo, A. L., De Figueiredo, N. M. A., De Oliveira, L. F. D., and De Araújo, S. T. C. (2017). Proxemic behavior of nursing in the hemodialysis setting. *ACTA Paulista de Enfermagem*, 30(4):343–349.

Munzner, T. (2014). *Visualization Analysis and Design*, volume 1. CRC Press, 1 edition.

Nardi, B. and Whittaker, S. (2001). The place of face-to-face communication in distributed work. *Distributed Work*.

Neubauer, C., Woolley, J., Khooshabeh, P., and Scherer, S. (2016). Getting to know you: A multimodal investigation of team behavior and resilience to stress. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, ICMI '16, page 193–200, New York, NY, USA. Association for Computing Machinery.

Nielsen, J. and Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '93*, pages 206–213, New York, New York, USA. ACM Press.

Nilsen, T. R. (2021). Pedagogical intentions or practical considerations when facilitating children's play? Teachers' beliefs about the availability of play materials in the indoor ECEC environment. *International Journal of Child Care and Education Policy*, 15(1).

Noffke, S. E. and Brennan, M. (2005). The dimensions of reflection: A conceptual and contextual analysis. *International Journal of Progressive Education*, 1(3):58–78.

Noroozi, O., Alikhani, I., Järvelä, S., Kirschner, P. A., Juuso, I., and Seppänen, T. (2019). Multimodal data to design visual learning analytics for understanding regulation of learning. *Computers in Human Behavior*, 100:298–304.

Ochoa, X., Chiluiza, K., Granda, R., Falcones, G., Castells, J., and Guamán, B. (Mar. 7-9, 2018). Multimodal transcript of face-to-face group-work activity around interactive tabletops. In *Proc. 8th Learning Analytics & Knowledge Conf. (LAK' 18)*, page 1–6, Sydney, Australia.

Ochoa, X., Chiluiza, K., Méndez, G., Luzardo, G., Guamán, B., and Castells, J. (2013). Expertise estimation based on simple multimodal features. pages 583–590.

OECD (2018). *Embodied learning*.

Ojo, A. and Heravi, B. (2018). Patterns in award winning data storytelling. *Digital Journalism*, 6(6):693–718.

Okita, S. Y. (2012). Social Interactions and Learning. *Encyclopedia of the Sciences of Learning*, pages 3104–3107.

Olguín, D. O., Gloor, P. A., and Pentland, A. (2009). Wearable sensors for pervasive healthcare management. In *2009 3rd International Conference on Pervasive Computing Technologies for Healthcare - Pervasive Health 2009, PCTHealth 2009*. doi: https://doi.org/10.4108/ICST.PERVASIVEHEALTH2009.6033.

Oliver, R., Harper, B., Wills, S., Agostinho, S., and Hedberg, J. (2007). Describing ict-based learning designs that promote quality learning outcomes. *ERA 2010*.

Olson, J. (2002). Research about office workplace activities important to US businesses - And how to support them. *Journal of Facilities Management*, 1(1):31–47.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Onwuegbuzie, A. J. and Leech, N. L. (2005). On becoming a pragmatic researcher: The importance of combining quantitative and qualitative research methodologies. *International Journal of Social Research Methodology*, 8(5):375–387.

Paay, J., Kjeldskov, J., Skov, M. B., and O'Hara, K. (2013). F-formations in cooking together: A digital ethnography using youtube. In Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., and Winckler, M., editors, *Human-Computer Interaction – INTERACT 2013*, pages 37–54, Berlin, Heidelberg. Springer Berlin Heidelberg.

Palominos, E., Levett-Jones, T., Power, T., and Martinez-Maldonado, R. (2019). Healthcare students' perceptions and experiences of making errors in simulation: An integrative review. *Nurse Education Today*, 77:32–39.

Pardo, A. (2018). Feedback is good, but scaling it.... Technical report, BERA, London.

Pascoe, M. C., Hetrick, S. E., and Parker, A. G. (2020). The impact of stress on students in secondary school and higher education. *International Journal of Adolescence and Youth*, 25(1):104–112.

Pautasso, C. and Wilde, E. (2009). Why is the web loosely coupled? a multi-faceted metric for service design. In *Proceedings of the 18th International Conference on World Wide Web*, WWW '09, page 911–920, New York, NY, USA. Association for Computing Machinery.

Pearsall, M., Ellis, A., and Stein, J. (2009). Coping with challenge and hindrance stressors in teams: Behavioral, cognitive, and affective outcomes. organizational behavior and human decision processes, 109(1), 18-28. *Organizational Behavior and Human Decision Processes*, 109:18–28.

Petrosoniak, A., Almeida, R., Pozzobon, L. D., Hicks, C., Fan, M., White, K., McGowan, M., and Trbovich, P. (2019). Tracking workflow during high-stakes resuscitation: the application of a novel clinician movement tracing tool during in situ trauma simulation. *BMJ Simulation and Technology Enhanced Learning*, 5(2):78–84.

Pirolli, P. and Card, S. (2005). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, volume 5, pages 2–4. McLean, VA, USA.

Plano Clark, V. L. and Ivankova, N. V. (2016). *Mixed Methods Research: A Guide to the Field*. SAGE Publications, Inc., Thousand Oaks, California.

Pointon, P. and Kershner, R. (2000). Making decisions about organising the primary classroom environment as a context for learning: the views of three experienced teachers and their pupils. *Teaching and Teacher Education*, 16(1):117–127.

Pokorny, H. and Pickford, P. (2010). Complexity, cues and relationships: Student perceptions of feedback. *Active Learning in Higher Education*, 11:21–30.

Praharaj, S., Scheffel, M., Drachsler, H., and Specht, M. (2021a). Literature review on co-located collaboration modeling using multimodal learning analytics—can we go the whole nine yards? *IEEE Transactions on Learning Technologies*, 14(3):367–385.

Praharaj, S., Scheffel, M., Schmitz, M., Specht, M., and Drachsler, H. (2021b). Towards automatic collaboration analytics for group speech data using learning analytics. *Sensors*, 21(9).

Prieto, L., Sharma, K., Dillenbourg, P., and Rodríguez-Triana, M. (2016). Teaching analytics: Towards automatic extraction of orchestration graphs using wearable sensors.

Prieto, L. P., Martinez-Maldonado, R., Spikol, D., Hernandez-Leo, D., Rodriguez-Triana, M. J., and Ochoa, X. (2017). Multimodal learning analytics and learning analytics across physical and digital spaces 2017. volume 1828, page 4.

Prieto, L. P., Rodríguez-Triana, M. J., Martínez-Maldonado, R., Dimitriadis, Y., and Gašević, D. (2019). Orchestrating learning analytics (orla): Supporting inter-stakeholder communication about adoption of learning analytics at the classroom level. *Australasian Journal of Educational Technology*, 35(4).

Prilla, M., Pammer, V., and Krogstie, B. (2013). Fostering collaborative redesign of work practice: Challenges for tools supporting reflection at work. In Bertelsen, O. W., Ciolfi, L., Grasso, M. A., and Papadopoulos, G. A., editors, *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work, 21-25 September 2013, Paphos, Cyprus*, pages 249–268, London. Springer London.

Raca, M., Kidzinski, L., and Dillenbourg, P. (2015). Translating Head Motion into Attention - Towards Processing of Student's Body-Language. *EDM*.

Rädle, R., Jetter, H.-C., Marquardt, N., Reiterer, H., and Rogers, Y. (2014). Huddlelamp: Spatially-aware mobile displays for ad-hoc around-the-table collaboration. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, pages 45–54.

Radoi, I., Gutu, G., Rebedea, T., Neagu, C., and Popa, M. (2017). Indoor positioning inside an office building using ble. In *2017 21st International Conference on Control Systems and Computer Science (CSCS)*, pages 159–164. IEEE.

Raelin, J. (2002). “i don’t have time to think!” versus the art of reflective practice. *Reflections: The Sol Journal*, 4:66–79.

Raffaghelli, J. E. and Stewart, B. (2020). Centering complexity in ‘educators’ data literacy’ to support future practices in faculty development: a systematic review of the literature. *Teaching in Higher Education*, 25(4):435–455. doi: `https://doi.org/10.1080/13562517.2019.1696301`.

Raj, S., Lee, J. M., Garrity, A., and Newman, M. W. (2019). Clinical data in context: Towards sensemaking tools for interpreting personal health data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 3(1). doi: `https://doi.org/10.1145/3314409`.

Riel, J., Lawless, K. A., and Brown, S. W. (2018). Timing matters: Approaches for measuring and visualizing behaviours of timing and spacing of work in self-paced online teacher professional development courses. *Journal of Learning Analytics*, 5(1):25–40.

Roberts, A. P., Webster, L. V., Salmon, P. M., Flin, R., Salas, E., Cooke, N. J., Read, G. J., and Stanton, N. A. (2022). State of science: models and methods for understanding and enhancing teams and teamwork in complex sociotechnical systems. *Ergonomics*, 65(2):161–187.

Robertson, T., Li, J., O’Hara, K., and Hansen, S. (2010). Collaboration Within Different Settings: A Study of Co-located and Distributed Multidisciplinary Medical Team Meetings. *Computer Supported Cooperative Work (CSCW)*, 19(5):483–513.

Rogers, Y. (2012). *HCI Theory: Classical, Modern, and Contemporary*. Morgan & Claypool Publishers, 1 edition.

Ronda-Carracao, M. A., Santos, O. C., Fernandez-Nieto, G., and Martinez-Maldonado, R. (2021). Towards exploring stress reactions in teamwork using multimodal physiological data.

Rosen, M. A., Dietz, A. S., Yang, T., Priebe, C. E., and Pronovost, P. J. (2014). An integrative framework for sensor-based measurement of teamwork in healthcare. *Journal of the American Medical Informatics Association*, 22(1):11–18.

Roth, W.-M., Tenenberg, J., and Socha, D. (2016). Discourse/s in/of cscw. *Comput. Supported Coop. Work*, 25(4–5):385–407.

Rousseau, V. and Aubé, C. (2010). Social support at work and affective commitment to the organization: The moderating effect of job resource adequacy and ambient conditions. *The Journal of Social Psychology*, 150(4):321–340. PMID: 20718219.

Rousseau, V., Aubé, C., and Savoie, A. (2006). Teamwork behaviors: A review and an integration of frameworks. *Small Group Research*, 37(5):540–570.

Russell, T. (2003). Critical attributes of a reflective teacher: Is agreement possible? In *Conceptualising reflection in teacher development*, pages 150–159. Routledge.

Ryan, L. (2016). *The Visual Imperative: Creating a Visual Culture of Data Discovery*. Morgan Kaufmann, Cambridge, MA, USA.

Salas, E., Burke, C. S., and Cannon-Bowers, J. A. (2000). Teamwork: emerging principles. *International Journal of Management Reviews*, 2(4):339–356.

Salas, E., Goodwin, G. F., and Burke, C. S. (2009). *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches*.

Saldaña, J. (2015). *The coding manual for qualitative researchers*. Sage.

Salik, I. and Paige, J. T. (2020). Debriefing the interprofessional team in medical simulation. In *StatPearls [Internet]*. StatPearls Publishing.

Sanchez, J. E., Humphreys, K., and Carroll, K. (2019). Exploring the co-teaching experience in a graduate-level, principal preparation course. *InSight: A Journal of Scholarly Teaching*, 14:99–112.

Saquib, N., Bose, A., George, D., and Kamvar, S. (2018). Sensei: sensing educational interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(4):1–27.

Sawyer, T., Eppich, W., Brett-Fleegler, M., Grant, V., and Cheng, A. (2016). More than one way to debrief: a critical review of healthcare simulation debriefing methods. *Simulation in Healthcare*, 11(3):209–217.

Scherr, R. E. and Hammer, D. (2009). Student behavior and epistemological framing: Examples from collaborative active-learning activities in physics. *Cognition and Instruction*, 27(2):147–174.

Schield, M. (2004). Information literacy, statistical literacy and data literacy. *IASSIST Quarterly*, 28(2):7–14.

Schneider, B., Sung, G., Chng, E., and Yang, S. (2021). How can high-frequency sensors capture collaboration? a review of the empirical links between multimodal metrics and collaborative constructs. *Sensors*, 21(24):8185.

Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.* Jossey-Bass.

Schraagen, J. M., Schouten, T., Smit, M., Haas, F., Van Der Beek, D., Van De Ven, J., and Barach, P. (2010). Assessing and improving teamwork in cardiac surgery. *Quality and Safety in Health Care*, 19(6):e29–e29.

Schulz, H.-J., Streit, M., May, T., and Tominski, C. (Oct. 13-18, 2013). Towards a characterization of guidance in visualization. In *IEEE Inf. Visualization Conf.)*, page 1, Baltimore, MD, USA.

Schwabe, L., Joëls, M., Roozendaal, B., Wolf, O. T., and Oitzl, M. S. (2012). Stress effects on memory: an update and integration. *Neuroscience & Biobehavioral Reviews*, 36(7):1740–1749.

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Gillet, D., and Dillenbourg, P. (2016). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Trans Learn. Technol.*, 10(1):30–41. doi: 10.1109/TLT.2016.2599522.

Secheresse, T. and Nonglaton, S. (2019). The "Timeline Debriefing Tool": a tool for structuring the debriefing description phase. *Advances in Simulation*, 4(1):29.

Sedlmair, M., Meyer, M., and Munzner, T. (2012). Design study methodology: Reflections from the trenches and the stacks. *IEEE transactions on visualization and computer graphics*, 18(12):2431–2440.

Segel, E. and Heer, J. (2010). Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148.

Sellberg, C., Lindwall, O., and Rystedt, H. (2021). The demonstration of reflection-in-action in maritime training. *Reflective Practice*, 22(3):319–330.

Selwyn, N. and Gašević, D. (2020). The datafication of higher education: discussing the promises and problems. *Teaching in Higher Education*, 25(4):527–540.

Setti, F., Russell, C., Bassetti, C., and Cristani, M. (2015). F-Formation Detection: Individuating Free-Standing Conversational Groups in Images. *PLOS ONE*, 10(5):e0123783.

Shaffer, D. W. (2006). Epistemic frames for epistemic games. *Computers & Education*, 46(3):223–234.

Shaffer, D. W. (2017). *Quantitative ethnography*. Cathcart Press.

Shaffer, D. W., Collier, W., and Ruis, A. R. (2016). A Tutorial on Epistemic Network Analysis: Analyzing the Structure of Connections in Cognitive, Social, and Interaction Data. Technical Report 3.

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Frank, K., Rupp, A. A., and Mislevy, R. (2009). Epistemic Network Analysis: A Prototype for 21st-Century Assessment of Learning. *International Journal of Learning and Media*, 1(2):33–53.

Shaffer, D. W. and Ruis, A. R. (2017). Epistemic Network Analysis: A Worked Example of Theory-Based Learning Analytics. In *Handbook of Learning Analytics*, chapter 15, pages 175–187. Society for Learning Analytics Research (SoLAR).

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

Shorten, A. and Smith, J. (2017). Mixed methods research: expanding the evidence base. *Evidence-Based Nursing*, 20(3):74–75.

Shute, V. and Ventura, M. (2013). *Stealth assessment: Measuring and supporting learning in video games*. MIT press.

Skitka, L., Mosier, K., Burdick, M., and Rosenblatt, B. (2000). Automation bias and errors: Are crews better than individuals? *The International journal of aviation psychology*, 10:85–97.

Skulmowski, A. and Rey, G. D. (2018). Embodied learning: introducing a taxonomy based on bodily engagement and task integration. *Cognitive research: principles and implications*, 3(1):1–10.

Slade, S. and Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, 57(10):1510–1529.

Sorokowska, A., Sorokowski, P., Hilpert, P., Cantarero, K., Frackowiak, T., Ahmadi, K., Alghraibeh, A. M., Aryeetey, R., Bertoni, A., Bettache, K., et al. (2017). Preferred interpersonal distances: a global comparison. *Journal of Cross-Cultural Psychology*, 48(4):577–592.

Sottilare, R. A., Shawn Burke, C., Salas, E., Sinatra, A. M., Johnston, J. H., and Gilbert, S. B. (Jun. 08, 2017). Designing adaptive instruction for teams: A meta-analysis. *Int. J. Artificial Intell. Educ.*, 28(2):225–264. doi: 10.1007/s40593-017-0146-z.

Stahl, G. (2011). Theories of cognition in cscw. In *ECSCW 2011: Proceedings of the 12th European Conference on Computer Supported Cooperative Work, 24-28 September 2011, Aarhus Denmark*, pages 193–212. Springer.

Stolz, S. A. (2015). Embodied learning. *Educational Philosophy and Theory*, 47(5):474–487.

Tan, J. P.-L. and Koh, E. (2017). Situating learning analytics pedagogically: Towards an ecological lens. *Learn.: Res. & Pract.*, 3(1):1–11. doi: 10.1080/23735082.2017.1305661.

Tang, L. M. and Kay, J. (2017). Harnessing long term physical activity data—how long-term trackers use data and how an adherence-based interface supports new insights. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(2).

Taylor, S., Jaques, N., Chen, W., Fedor, S., Sano, A., and Picard, R. (2015). Automatic Identification of Artifacts in Electrodermal Activity Data. *Engineering in Medicine and Biology Conference*, page 10.

Treffinger, D., Young, G., Selby, E., and Shepardson, C. (2002). Assessing creativity: A guide for educators. *National Research Center on the Gifted and Talented*.

Tschimmel, K. (2012). Design thinking as an effective toolkit for innovation. In *ISPIM Conference Proceedings*, page 1. The International Society for Professional Innovation Management (ISPIM).

Tuck, J. (2012). Feedback-giving as social practice: teachers' perspectives on feedback as institutional requirement, work and dialogue. *Teaching in Higher Education*, 17(2):209–221.

Vanneste, P., Raes, A., Morton, J., Bombeke, K., Van Acker, B. B., Larmuseau, C., Depaepe, F., and Van den Noortgate, W. (2021). Towards measuring cognitive load through multimodal physiological data. *Cognition, Technology & Work*, 23(3):567–585.

Verbert, K., Govaerts, S., duval, e., Santos, J., Van Assche, F., Parra, G., and Klerkx, J. (2013). Learning dashboards: An overview and future research opportunities. *Personal and Ubiquitous Computing*, 18.

Veteska, J., Kursch, M., Svobodova, Z., Tureckiova, M., and Paulovcakova, L. (2022). *Longitudinal Co-teaching Projects: Scoping Review*, pages 35–53. Springer International Publishing, Cham.

Viktorelius, M. and Sellberg, C. (2022). The lived body and embodied instructional practices in maritime basic safety training. *Vocations and Learning*, 15(1):87–109. doi: `https://link.springer.com/article/10.1007/s12186-021-09279-z`.

Villanueva, I., Campbell, B. D., Raikes, A. C., Jones, S. H., and Putney, L. G. (2018). A multimodal exploration of engineering students' emotions and electrodermal activity in design activities. *Journal of Engineering Education*, 107(3):414–441. doi: `https://doi.org/10.1002/jee.20225`.

Viller, S. (1991). The group facilitator: a cscw perspective. In *Proceedings of the Second European Conference on Computer-Supported Cooperative Work ECSCW'91*, pages 81–95. Springer. doi: `https://dl.acm.org/doi/10.5555/1241910.1241916`.

Vujović, M. and Hernández Leo, D. (Jun. 17-21, 2019). Shall we learn together in loud spaces? Towards understanding the effects of sound in collaborative learning environments. In *Proc. 13th Int. Conf. on Computer Supported Collaborative Learning (CSCL' 19)*, page 2, Lyon, France. url: `https://hdl.handle.net/10230/37185`.

Vygotsky, L. S. (1997). *The Collected Works of L. S. Vygotsky*, volume 1. Springer US, New York, springer edition.

Wake, J. D., Wasson, B., Bjørgen, E. P., and Heimsæter, F. V. (2019). Supporting firefighter training through visualising indoor positioning, motion and time use data. In *Proceedings of the 16th International Conference on Information Systems for Crisis Response and Management*, pages 1–10. url: `https://norceresearch.brage.unit.no/norceresearch-xmlui/bitstream/handle/11250/2650379/ISCRAM-4.pd`.

Wang, D., Amin, M. T., Li, S., Abdelzaher, T., Kaplan, L., Gu, S., Pan, C., Liu, H., Aggarwal, C. C., Ganti, R., et al. (2014). Using humans as sensors: an estimation-theoretic perspective. In *Proc. IEEE Simposium on information processing in sensor networks*, pages 35–46. IEEE. doi: 10.1109/IPSN.2014.6846739.

Wang, D., Yang, Q., Abdul, A., and Lim, B. Y. (May 4-9, 2019). Designing theory-driven user-centric explainable AI. In *Conf. Human Factors in Computing Systems (CHI'19)*, page 1–15, Glasgow Scotland, U.K. ACM Press. doi: https://doi.org/10.1145/3290605.3300831.

Watanabe, E., Ozeki, T., and Kohama, T. (2018). Analysis of interactions between lecturers and students. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, pages 370–374, New York, NY, USA. ACM. doi: https://doi.org/10.1145/3170358.3170360.

Webster, C. S., Jowsey, T., Lu, L. M., Henning, M. A., Verstappen, A., Wearn, A., Reid, P. M., Merry, A. F., and Weller, J. M. (2019). Capturing the experience of the hospital-stay journey from admission to discharge using diaries completed by patients in their own words: a qualitative study. *BMJ open*, 9(3):e027258. doi: https://doi.org/10.1136/bmjopen-2018-027258.

Wiggermann, N., Smith, K., and Kumpar, D. (2017). What bed size does a patient need? the relationship between body mass index and space required to turn in bed. *Nursing Research*, 66(6):483. doi: https://doi.org/10.1097/nnr.0000000000000242.

Wiley, K. J., Dimitriadis, Y., Bradford, A., and Linn, M. C. (2020). From theory to action: Developing and evaluating learning analytics for learning design. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, LAK '20, page 569–578, New York, NY, USA. Association for Computing Machinery. doi: https://doi.org/10.1145/3375462.3375540.

Wilson, J., Straus, S., and McEvily, B. (2006). All in due time: The development of trust in computer-mediated and face-to-face teams. *Organizational Behavior and Human Decision Processes*, 99:16–33. doi: https://doi.org/10.1016/j.obhdp.2005.08.001.

Winder, I., Delaporte, D., Wanaka, S., and Hiekata, K. (2020). Sensing teamwork during multi-objective optimization. In *2020 IEEE 6th World Forum on Internet of Things (WF-IoT)*, pages 1–6. doi: https://doi.org/10.1109/WF-IoT48130.2020.9221086.

Wise, A. and Jung, Y. (2019). Teaching with analytics: Towards a situated model of instructional decision-making. *Journal of Learning Analytics*, 6:53–69. doi: https://doi.org/10.18608/jla.2019.62.4.

Wise, A. F., Knight, S., and Buckingham Shum, S. (2021). *Collaborative Learning Analytics*, pages 425–443. Springer, Cham. doi :"10.1007/978-3-030-65291-3$_2$3".

Wooldridge, A. R., Carayon, P., Shaffer, D. W., and Eagan, B. (2018). Quantifying the qualitative with epistemic network analysis: A human factors case study of task-allocation communication in a primary care team. *IISE Transactions on Healthcare Systems Engineering*, 8(1):72–82. url: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6201247/.

Worsley, M. (2014). Multimodal learning analytics as a tool for bridging learning theory and complex learning behaviors. In *Proceedings of the 2014 ACM Workshop on Multimodal Learning Analytics Workshop and Grand Challenge*, MLA '14, page 1–4, New York, NY, USA. Association for Computing Machinery. doi: https://doi.org/10.1145/2666633.2666634.

Worsley, M. (2018). Multimodal Learning Analytics ' Past , Present , and , Potential Futures. In *CrossMMLA@LAK*. url: http://ceur-ws.org/Vol-2163/paper5.pdf.

Worsley, M., Anderson, K., Melo, N., and Jang, J. (2021a). Designing Analytics for Collaboration Literacy and Student Empowerment. *Journal of Learning Analytics*, 8(1):30–48. doi: http://doi.org/10.18608/jla.2021.7242.

Worsley, M., Martinez-Maldonado, R., and D'Angelo, C. (2021b). A new era in multimodal learning analytics: Twelve core commitments to ground and grow mmla. *Journal of Learning Analytics*, 8(3):10–27. doi: https://doi.org/10.18608/jla.2021.7361.

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

Yang, D. F. and Goodyear, P. (2014). *Design Narrative*, pages 41–47. SensePublishers, Rotterdam. doi: https://doi.org/10.1007/978-94-6209-530-4_7.

Yeoman, P. and Wilson, S. (2019). Designing for situated learning: Understanding the relations between material properties, designed form and emergent learning activity. *British Journal of Educational Technology*, 50(5):2090–2108. doi: `https://doi.org/10.1111/bjet.12856`.

Zhang, Z. and Sarcevic, A. (2015). Constructing Awareness Through Speech, Gesture, Gaze and Movement During a Time-Critical Medical Task. In *ECSCW 2015: Proceedings of the 14th European Conference on Computer Supported Cooperative Work, 19-23 September 2015, Oslo, Norway*, pages 163–182. Springer International Publishing, Cham. doi: `http://link.springer.com/10.1007/978-3-319-20499-4_9`.

Zhao, L., Yan, L., Gasevic, D., Dix, S., Jaggard, H., Wotherspoon, R., Alfredo, R., Li, X., and Martinez-Maldonado, R. (2022). Modelling co-located team communication from voice detection and positioning data in healthcare simulation. In *LAK22: 12th International Learning Analytics and Knowledge Conference*, pages 370–380. doi: `https://dl.acm.org/doi/10.1145/3506860.3506935`.

Zoltan, R. and Vancea, R. (2015). Organizational work groups and work teams–approaches and differences. *Ecoforum Journal*, 4(1):13.