Designing Feedback for Collocated Teams using Multimodal Learning Analytics

Vanessa Echeverria Barzola

B.Sc., M.Sc.

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

University of Technology Sydney

Connected Intelligence Centre

2020

Vanessa Echeverria (2020). Designing Feedback for Collocated Teams using Multimodal Learning Analytics.

Supervisors:

Prof. Simon Buckingham Shum, University of Technology Sydney

Dr. Roberto Martinez-Maldonado, Monash University

Certificate of Original Authorship

I, Vanessa Echeverria Barzola, declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Connected Intelligence Centre at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference 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.

This research is supported by the Australian Government Research Training Program.

Production Note:
Signature removed prior to publication.

Signature: ____________________________

Date: 18/03/20

Acknowledgements

This work could not have been completed without the help support and guidance of many people. I have learnt many invaluable lessons during my doctorate and would like to acknowledge those who have assisted me in my doctoral journey.

First and foremost, I would like to extend my deepest gratitude to my supervisors, Prof. Simon Buckingham Shum and Dr. Roberto Martinez-Maldonado. I am grateful to Simon Buckingham Shum for allowing me to join a fantastic research centre and for guiding me to become a responsible researcher. Your willingness to share your knowledge and innovative insights has helped me shaped my research.

My heartfelt thanks go to Roberto for his continued support, guidance and supervision. During this time Roberto was more than just my supervisor, he was my mentor and friend. I enjoyed every conversation we had, research-related or life-related. Roberto pushed me towards new challenges, he is the person and friend who made all this possible. Roberto was always positive, no matter the challenge. His guidance went beyond my research and has impacted many aspects of my life. Thank you for the countless times you cheered me up and for your unconditional support and kindness through this journey. Thanks for teaching me how to be an organised researcher and that every little step count!

I would also like to acknowledge the contribution of the people who voluntarily dedicated their time and expertise. My sincere thanks go to Prof. Cristina Conati for offering valuable advice on the analysis of the eye-tracking data. Thanks to the excellent staff from the Faculty of Health, without their collaboration and support this research would not be possible. In particular to Dr Tamara Power, Dr. Carmen Axisa, and Dr. Carolyn Hayes for always being engaged in this research, for being encouraging, for opening the doors to their classrooms and being very supportive through every stage of this research. Thanks also to Prof. Doug Elliot, Prof. Tracy Levett-Jones, Dr. Natalie Govind, Ms. Jan Forber, and Mrs. Felicity Smith for their valuable time and insight which helped shape the work presented in this thesis. Thanks to the invaluable support of simulation technicians Michael Cabauatan, Paul Benson, Connie Land and TJ Agudera. Thanks also to the students that were willing to participate in the studies reported in this work.

A special thank you goes out to my internal reviewers, Dr. Simon Knight and Prof. Judy Kay, for providing me with constructive feedback on my final presentation and the first thesis draft.

I want to express my gratitude to my colleagues from the Connected Intelligence Centre. Thanks to Dr. Kirsty Kitto for her constructive advice and for fostering my critical thinking during our fortnightly discussions. I'd like to recognise the assistance that I received from Georgia, Gabrielle, Emma and Ratha for helping me solve administrative issues. Thanks also to Jack and Radhika for being friendly and kind to me. Special thanks to my great friends Sophie Abel and Carlos Prieto with which I shared the ups and downs of this PhD adventure. Thanks to the PhD crew, Shibani, Evelyn Palominos and Gloria Fernandez for being part of the PhD life and for sharing fun times with me.

My gratitude is also extended to the funding bodies and scholarships that provided financial support for my doctorate and the presentation of my PhD work at research conferences. Thanks to the Asian-Pacific Society for Computers in Education, the Connected Intelligence Centre, UTS, and the Escuela Superior Politécnica del Litoral for their scholarships.

I would not have been able to undertake this endeavour without the colleagues who supported me from the beginning of my research career. I'm deeply indebted to Prof. Katherine Chiluiza for mentoring me during the initial stages of my research career and teaching me to strive for the best. Also, I'm grateful to Prof. Xavier Ochoa for inspiring me with his innovative ideas and for introducing me to the Learning Analytics community.

I welcome this opportunity to thank my friends and family. To my new friends in Sydney, thanks for making this place feel like home. To my dear friends in Guayaquil, thanks for being there, not physically but emotionally, sharing your joyfulness in our video calls. Thanks to my family, Fabricio, his wife and kids, Bryan, Mariuxi, and all relatives, for showing me that everything is possible when you are surrounded by love.

Special thanks go to my parents; you are the pillar of who I am, and I owe you everything. Thanks to my dad for teaching me that dedication and hard work is the key to making things happen. Thanks to my mom, this is the result of all your selfless support and prayers. Thanks for your unconditional love, support and for keeping me always motivated.

To my beloved husband, Ivan, thank you for standing with me through all the emotions I've been through during this time. You are one in a million. To my dear kids, Ivan and Annie, thanks for brightening my days with your smiles, kisses and unconditional love. Finally, I thank God for the wisdom and strength he bestowed upon me to finish this work.

Preface and Notes

The following peer-reviewed publications produced during the PhD candidature contribute to this thesis.

Journal Papers

- Roberto Martinez-Maldonado, Doug Elliott, Carmen Axisa, Tamara Power, **Vanessa Echeverria** and Simon Buckingham Shum (2020). Designing translucent learning analytics with teachers: an elicitation process. Interactive Learning Environments.

- **Vanessa Echeverria,** Roberto Martinez-Maldonado, Simon Buckingham Shum, Katherine Chiluiza, Roger Granda, Cristina Conati (2018). Exploratory versus Explanatory Visual Learning Analytics: Driving Teachers' Attention through Educational Data Storytelling. Journal of Learning Analytics, 5(3), (pp. 72-97).

Conference Proceedings

- Roberto Martinez-Maldonado, **Vanessa Echeverria**, Gloria Fernandez Nieto, Simon Buckingham Shum (2020). From Data to Insights: A Layered Storytelling Approach for Multimodal Learning Analytics. Accepted for the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20).

- Simon Buckingham Shum, **Vanessa Echeverria**, Roberto Martinez-Maldonado (2019). The Multimodal Matrix as a Quantitative Ethnography Methodology. In International Conference on Quantitative Ethnography (pp. 26-40). Springer, Cham.

- **Vanessa Echeverria**, Roberto Martinez-Maldonado, Simon Buckingham Shum (2019). Towards Collaboration Translucence: Giving Meaning to Multimodal Group Data. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), Paper 39, 16 pages, ACM, New York, NY, USA.

- Roberto Martinez-Maldonado, **Vanessa Echeverria**, Doug Elliott, Carmen Axisa, Tamara Power, Simon Buckingham Shum (2019). Making the Design of CSCL Analytics Interfaces a Co-design Process: the Case of Multimodal Teamwork in Healthcare. In Proceedings of the 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019, Volume 2 (pp. 859-860). Lyon, France: International Society of the Learning Sciences.

- **Vanessa Echeverria**, Roberto Martinez-Maldonado, Tamara Power, Carolyn Hayes, Simon Buckingham Shum (2018). Where is the nurse? towards automatically visualising meaningful team movement in healthcare education. In International conference on

Artificial Intelligence in Education (AIED 2018). Lecture Notes in Computer Science, vol 10948. Springer, Cham.

- Roberto Martinez-Maldonado, **Vanessa Echeverria**, Olga C Santos, Augusto Dias Pereira Dos Santos, Kalina Yacef (2018). Physical learning analytics: A multimodal perspective. In Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK '18). ACM, New York, NY, USA, (pp. 375-379).
- **Vanessa Echeverria**, Roberto Martinez-Maldonado, Roger Granda, Katherine Chiluiza, Cristina Conati, Simon Buckingham Shum (2018). Driving data storytelling from learning design. In Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK '18). ACM, New York, NY, USA, (pp. 375-379).
- **Vanessa Echeverria**, Roberto Martinez-Maldonado, Simon Buckingham Shum (2017). Towards data storytelling to support teaching and learning. In Proceedings of the 29th Australian Conference on Computer-Human Interaction (OZCHI '17). ACM, New York, NY, USA, (pp. 347-351).
- **Vanessa Echeverria**, Roberto Martinez-Maldonado, Katherine Chiluiza, S Buckingham Shum (2017). DBCollab: Automated feedback for face-to-face group database design. In Proceedings of the 25th International Conference on Computers in Education, ICCE 2017-Main Conference Proceedings. (pp. 156 - 165).

Workshop Papers

- Roberto Martinez-Maldonado, **Vanessa Echeverria**, Kalina Yacef, Augusto Dias Pereira Dos Santos, Mykola Pechenizkiy (2017). How to capitalise on mobility, proximity and motion analytics to support formal and informal education?. In MMLA- CrossLAK 2017 (pp. 39-46). (CEUR Workshop Proceedings; Vol. 1828). CEUR-WS.org.

Workshops

- Roberto Martinez-Maldonado, **Vanessa Echeverria**, Luis P Prieto, Maria Jesus Rodriguez-Triana, Daniel Spikol, Mutlu Curukova, Manolis Mavrikis, Xavier Ochoa, Marcelo Worsley (2018). 2nd CrossMMLA: Multimodal learning analytics across physical and digital spaces. In MMLA- CrossLAK 2018. (CEUR Workshop Proceedings; Vol. 2163). CEUR-WS.org.

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 references at the end of the thesis. Figures from external sources where authors granted permission for usage are cited in their captions.

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 email[1].

[1] Email addresses: Vanessa.I.EcheverriaBarzola@student.uts.edu.au, Roberto.MartinezMaldonado@monash.edu, Simon.BuckinghamShum@uts.edu.au,

Table of Contents

List of Figures

List of Tables

List of Abbreviations

LA	Learning Analytics
MMLA	Multimodal Learning Analytics
HCI	Human Computer Interaction
HCD	Human-Centred Design
CSCL	Computer-Supported Collaborative Learning
CSCW	Computer-Supported Collaborative Work
DS	Data Storytelling
InfoVis	Information Visualisation
QE	Quantitative Ethnography
TUI	Tangible User Interfaces
DBR	Design Based Research
IPA	Interaction Process Analysis
ACAD	Activity-centred Analysis and Design
CPS	Collaborative problem solving
EDA	Electrodermal Activity
HCD-MMLA	Human-Centred Design Multimodal Learning Analytics
HoC	Higher order constructs
OV	Original visualisations
VDS	Visualisations with data storytelling elements
EvisLA	Explanatory Visual Learning Analytics

Abstract

The ability to communicate, be an effective team or group member and collaborate face-to-face are critical skills for employability in the 21st century workplace. Previous research suggests that learning to collaborate effectively requires practice, awareness of group dynamics and reflection upon past activities. However, although having a teacher closely supervising and providing detailed feedback to each group would be ideal, it may be unrealistic in practice. A promising way to approach this challenge could be to capture behavioural traces from group interactions in order to generate comprehensible and actionable feedback to support team reflection. In this sense, Multimodal Learning Analytics (MMLA) is a promising field, offering the potential to track learners' activity across digital and collocated contexts, using emerging sensing and pervasive computing technologies. Most of the research in MMLA has been conducted in lab conditions, to help researchers validate learning theories or generate more comprehensive learner models. However, one of the most underexplored aspects of MMLA has been the generation of feedback to support teaching and learning, and moreover, in authentic locations and activities.

This thesis reports progress in tackling this challenge by designing and validating computer-based feedback, by means of visual representations and narrative, to support effective, guided reflection using multimodal learning analytics evidence. To achieve this, three contributions are presented. The first contribution is a human-centred design method to translate the informal outputs of co-design sessions with teachers and students, into more meaningful group work constructs with clear MMLA design requirements. The second contribution is a modelling approach to add meaning to low-level multimodal group data based on the characteristics of the context (domain expertise, theory, and the learning design). Finally, the third contribution is an approach for augmenting visual representations with data storytelling elements to facilitate the interpretation of group dynamics insights by educators and students. This thesis is developed in the context of two distinct, collocated group work settings, in the domains of collaborative database design and healthcare simulation. Using a Design-Based Research process, a set of *explanatory* interfaces (i.e. interfaces that communicate insights) was designed and validated with teachers and students. The thesis provides timely and necessary groundwork for researchers and practitioners to design visual representations capable of communicating actionable insights, using multimodal data in complex and authentic collaboration scenarios.

Chapter 1: Introduction

The thesis investigates the design and validation of *explanatory* interfaces (those that communicate insights) using multimodal traces to support group activity by providing automated feedback and facilitating guided reflection. This chapter is structured as follows: Section 1.1 presents the context and motivation of this thesis. Section 1.2 introduces the research questions and Section 1.3 outlines the research goals. Next, Section 1.4 describes the contributions of this thesis. Finally, Section 1.5 provides a summary of the thesis structure with an overview of the remaining chapters of the thesis.

1.1 CONTEXT AND MOTIVATION

The ability to communicate, be an effective team/group member and collaborate face-to-face (f2f) are key 21st Century skills for the modern workforce that employers seek when recruiting professionals (Bellanca, 2011; de Lima & de Souza, 2017). Critical thinking, social interaction, negotiation, agreement, and regulation are some of the processes that often occur during the execution of effective group work activity (Dillenbourg, 1999; Salas, Sims, & Burke, 2005). Therefore, group work activities can be considered as a complex human tasks, not only because of the intrinsic complexity of the processes mentioned above, but also because of the many interactions occurring concurrently among team members through multiple channels of communication (Beebe & Masterson, 2003). Due to these challenges, learning to collaborate effectively requires practice, awareness of group dynamics and reflection upon past activities (Liu et al., 2017). Importantly, it often needs close coaching by an expert facilitator to foster beneficial collaborative interaction (Webb, 2009). However, although having a coach or tutor closely supervising each group might be ideal, it may be unrealistic in practice.

The literature suggests that without proper, timely and personalised feedback, groups may find it hard to improve their collaboration skills (Hattie & Timperley, 2007; London & Sessa, 2006). Nicol (2010) reported that students value personalised feedback, seeing it as critical for a successful learning experience, which is why teachers and academic administrators are often developing ways to support the provision of timely and intelligible feedback to students (Boud & Molloy, 2013a; De Luca, 2014). Yet, most of the feedback that students commonly receive is based on the formal assessment of their final products, which provides limited information about different aspects of a group's learning *processes* (Rousseau, Aubé, & Savoie, 2006). Alternatively, teachers may also provide comments during a debrief session based on what they observed but, with limited evidence of the interactions that occurred during the activity, it may be challenging for this reflection to become an effective learning opportunity (Riebe, Girardi, & Whitsed, 2016). This is particularly problematic in collocated settings in which the evidence about what happened during the group activity is ephemeral and hard to log.

During the last decade, pervasive technology has made it possible to capture traces of activity happening in physical spaces, using everyday technology (e.g. laptops and smartphones) and specialised devices (e.g. sensors, wearables and tabletops; Anslow, 2014; Jetter et al., 2012). These technologies are allowing the creation of enriched Computer-Supported Collaborative Learning (CSCL) and Computer-Supported Cooperative Work (CSCW) environments by supporting teachers in the orchestration and awareness of students' interactions (Martinez-Maldonado, Clayphan, et al., 2015; Martinez-Maldonado, Dimitriadis, Kay, Yacef, & Edbauer, 2013); supporting equal participation by mirroring group's interactions (Bachour, Kaplan, & Dillenbourg, 2010); or using these digital traces to inform the relevance of collaboration theories in experimental settings (Pijeira-Diaz, Drachsler, Järvelä, & Kirschner, 2016; Schneider, Sharma, et al., 2018). In short, making group work interactions visible, measurable, and interpretable through data streams captured from collocated environments (such as the classroom) can be useful to raise awareness and visibility of learning processes that previously were invisible or ephemeral (Noroozi et al., 2018).

Multimodal Learning Analytics (MMLA), which is a subfield in Learning Analytics research, offers new insights into learning by exploiting technologically feasible techniques to automatically capture evidence in collocated situations, and further use this evidence to support the provision of computer-based feedback (Blikstein & Worsley, 2016). To accomplish this, MMLA often makes use of emerging sensing and pervasive computing technologies to track students' and teachers' behaviour beyond the capture of clickstreams and keystrokes. A MMLA perspective also includes methods for analysing multiple data streams captured by sensors or logged by interactive devices (Blikstein & Worsley, 2016; Ochoa, 2017). However, current research reveals that most MMLA studies have been conducted in experimental settings with the aim of helping researchers validate learning theories (e.g. Haataja, Malmberg, & Järvelä, 2018; Schneider & Blikstein, 2015); identify critical incidents by visually inspecting multiple streams of data (e.g. Di Mitri, Schneider, Klemke, Specht, & Drachsler, 2019; Noroozi et al., 2018); mine patterns of activity from multiple data streams (Blikstein & Worsley, 2016; Spikol et al., 2017); and generate more complete models of learners (e.g. Schneider & Blikstein, 2015; Worsley & Blikstein, 2014). The challenge with capturing data from multiple sensors and interfaces is that, as the number of data streams grow, it becomes more complex to interpret and make sense of the data (Noroozi et al., 2018; Roll & Winne, 2015). Therefore, providing computer-based feedback based on multimodal group data requires interfaces that scaffold interpretation and sensemaking.

Current research in Learning Analytics (LA) has shown a growing interest in the use of visualisations of students' behaviour and interactions to scaffold the interpretation of data by using a combination of charts, tables and text. “Learning dashboards”[2] typically include such visualisations. Dashboards have gained considerable attention in the field as a way to visualise different traces of learning activities for teacher and students (Bodily et al., 2018; Bodily & Verbert, 2017; Schwendimann et al., 2017). Teacher-facing dashboards are commonly intended to help educators gain a better understanding of their whole course or specific tasks, reflect on their teaching strategies, and identify students who require specific attention (Molenaar & Knoop-van Campen, 2017; Verbert, Duval, Klerkx, Govaerts, & Santos, 2013). Similarly, student-facing dashboards are intended to help students reflect on aspects of their learning behaviour and potentially assist them in their learning, for example, managing time effectively, accessing key learning resources, or gaining a richer picture of their progress (Bodily & Verbert, 2017; Reimers & Neovesky, 2015). Existing learning dashboards communicate multiple insights about students’ experiences, but the visualisations designed are often complex and difficult to interpret (Duval, 2011) especially “at a glance”. Moreover, tutors and students are encouraged to interpret these visualisations in a limited time frame and, even if the data can be interpreted correctly, they may fail to understand the “call to action” for reflection or adapting their behaviour (Drachsler & Greller, 2011; Matcha, Gasevic, & Pardo, 2019). Furthermore, interpreting and making sense of learning dashboards that use multimodal group data can become even more problematic because each data source can reveal a different behaviour in relation to its function (e.g. movement data can be used to signal physical activity, and changes in heart rate to indicate stress).

A potential approach that could scaffold more effectively the interpretation and sensemaking of multimodal group data, is the use of explanatory visual elements to guide people’ attention to the data points that are important during the exploration and interpretation of the underlying data (Knaflic, 2015; Tufte, 2001). Explanatory visual elements are intended to facilitate interpretation and orient the audience by providing context through the emphasis of one or more elements; the addition of annotation/narrative elements; and the removal of unnecessary or irrelevant elements (Kosara & Mackinlay, 2013; Lee, Riche, Isenberg, & Carpendale, 2015). Explanatory visualisations have been widely used by journalists as a means to engage and inform through "interactive graphics and narratives" (Segel & Heer, 2010). Current reported uses of explanatory visualisations have focused on communicating scientific data (Ma, Liao, Frazier, Hauser, & Kostis, 2012); and teaching dynamic networks in an educational context (Bach et al., 2016). However, little is known about how to use explanatory visual elements in learning contexts and whether this approach can scaffold the interpretation and sensemaking of multimodal group data for students and teachers.

[2] In the literature, learning dashboards are also denoted as visual learning analytics, or learning visualisations. A detailed description and its use in learning analytics can be found in Chapter 2.

Moreover, recent studies have suggested that the design and validation of learning dashboards should be formalised and be grounded on human-centred approaches by involving stakeholders during the whole process (Martinez-Maldonado, Kay, Buckingham Shum, & Yacef, 2019; Sedrakyan, Malmberg, Verbert, Järvelä, & Kirschner, 2018). Previous studies have investigated the potential of co-design methods to explore, identify, and test prototypes, before providing a mature solution of learning analytics tools (Holstein, McLaren, & Aleven, 2017; Martinez-Maldonado, Pardo, et al., 2015; Prieto-Alvarez, Martinez-Maldonado, & Anderson, 2017). In MMLA research, little attention has been paid to justifying the benefits of capturing and analysing data streams in authentic scenarios to then provide actionable insights on the learning process back to teachers and students (Ochoa, Weibel, Worsley, & Oviatt, 2016; Oviatt, 2018; Worsley, 2018). In this sense, using co-design techniques may help researchers and designers to formalise the design requirements for the development of a MMLA solution, tailored to teachers and students' needs.

This thesis reports progress in tackling the different challenges outlined above, by investigating **explanatory multimodal learning analytics interfaces that communicate key aspects of the group work activity to guide reflection.**

Figure 1.1 summarises this thesis in relation to its context, goals, contributions and validation. This thesis follows a Human-Centred Design approach to investigate teachers' and students' requirements for the design and validation of explanatory interfaces, that can be generated from multimodal evidence. The next sections introduce the research questions, goals and contributions related to this thesis.

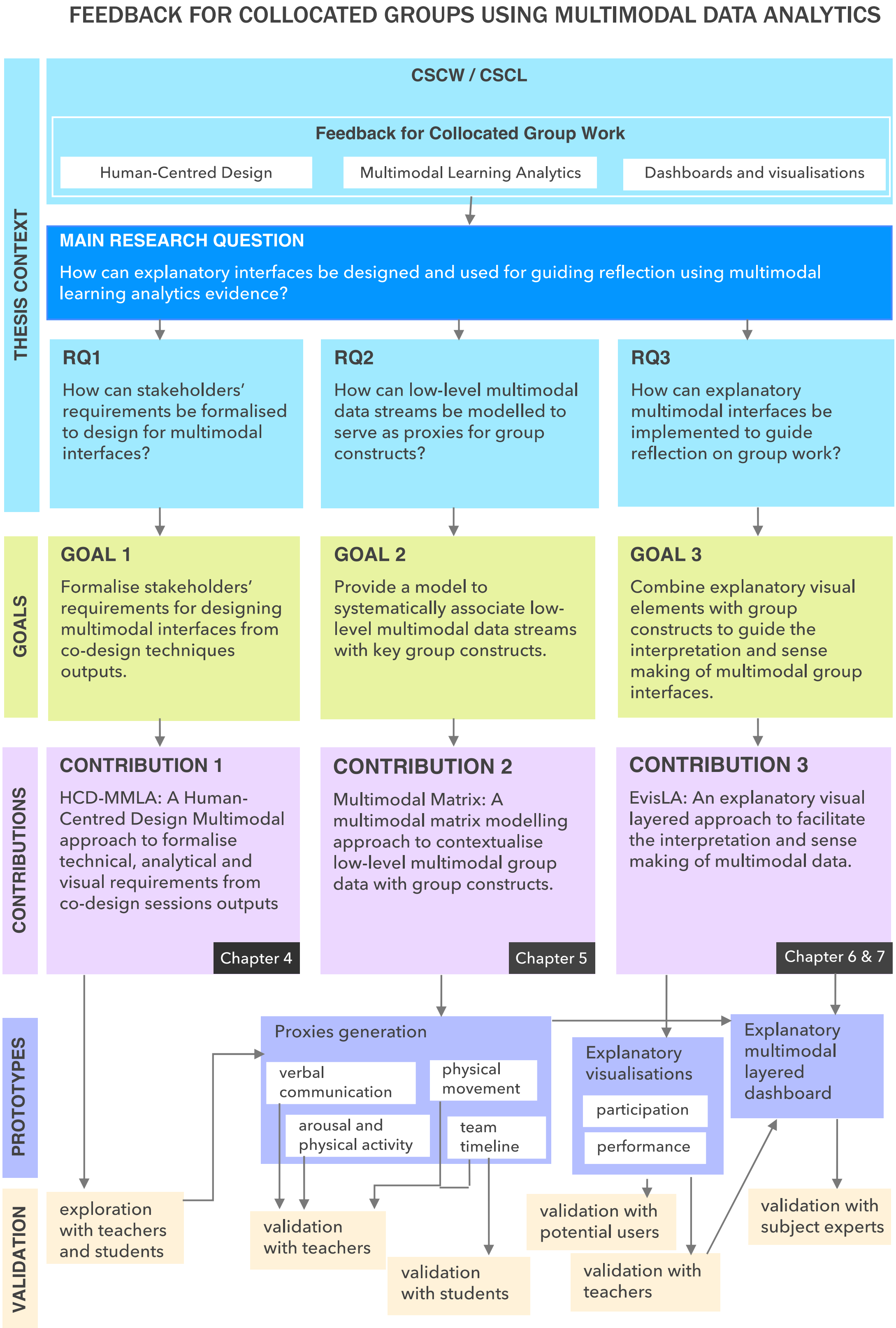


Figure 1.1: Overview of the context, goals, contributions and validation of this thesis.

1.2 RESEARCH QUESTIONS

This thesis raises the following broad question**: How can explanatory interfaces be designed and used for guiding reflection using multimodal learning analytics evidence?** This broad question has been divided in three sub-questions:

1. *How can stakeholders' requirements be formalised to design for multimodal interfaces?* MMLA solutions need to be evaluated to justify the cost of embedding sensing technologies into learning environments in light of the potential benefits for teaching and learning (Ochoa, 2017). This first question addresses the issue of defining the requirements to guide researchers and designers about the technical decisions when developing MMLA. Thus, a first challenge is to formally gather information from stakeholders, such as teachers, students, or administrative staff, to understand and define the learning context in relation to - not directly observable - learning expectations, learning intentions, student's behaviours and attitudes, or more explicitly, *learning constructs*. By bringing Human Centred Design (HCD) methods, this information can be formalised by considering teacher's experience and knowledge, driven by educational theory or from the learning design of the activity; and then be translated into key group constructs (e.g. teamwork, leadership). Then, these group constructs can be mapped to observable and detectable behaviours, the sensing technology to detect those behaviours, the analytical capability to extract those behaviours and the visual mechanisms to represent key constructs.
2. *How can low-level multimodal data streams be modelled to serve as proxies for group constructs?* This question tackles the problem of contextualising multimodal data with the semantics of group constructs in a structured way. One particular challenge is to analyse the activity through a holistic lens that captures relevant aspects of group work. A second challenge is the enrichment of quantitative data streams with qualitative insights, needed to make sense of the vast group interaction data available from the multimodal data capture.
3. *How can explanatory multimodal interfaces be implemented to guide reflection on group work?* This third question seeks to explore how group visualisations can be enhanced to communicate insights, thus becoming into explanatory visualisations. One challenge is to validate which visual elements drive user's attention of group visualisations. A second challenge is to combine explanatory visual elements with the multimodal data to scaffold the interpretation and sensemaking of complex information.

1.3 RESEARCH GOALS

Having described the thesis context, aim and research questions, the main goals of the thesis are formulated below (see Figure 1.1 - Goals):

1. *Formalise stakeholders' requirements from multimodal interface co-design sessions.* While co-design techniques can help to collect critical information about what teachers and students need, these informal ideas must be formalised in order to derive implementable requirements. The purpose of this goal is to formalise: i) group constructs explored from the learning design; teachers' experience or knowledge; or educational theory; and ii) teachers' and students' needs for supporting the provision of feedback.

2. *Provide a model to systematically map low-level multimodal data streams to key group constructs.* Once the requirements for creating multimodal group interfaces have been defined, the next step is to associate the group low-level multimodal data with group constructs to contextualise the underlying data to the learning environment. Therefore, it is necessary to provide a systematic way to model the structure and combine low-level sensor data to higher order constructs (informed by domain knowledge, theory, or other pedagogical requirements from the previous goal) that can help explain such data in a specific context. This goal proposes a model to map the semantics of qualitative interpretation from key group constructs to low-level group data streams by *i)* framing the group interaction analysis in the physical, social, epistemic and affective dimensions of group activity; and, *ii)* providing meaningful information from the processing of low-level data streams.

3. *Combine explanatory visual elements with group constructs to guide the interpretation and sensemaking of multimodal group interfaces.* It is challenging to provide visualisations that effectively communicate insights about the learning activity, and even more challenging if multimodal group data adds complexity to the interpretation and sensemaking of those visualisations. Inspired on Information Visualisation (InfoVis) and Data Storytelling (DS) techniques, this goal aims to implement *explanatory visualisations* to drive users' attention to focus on meaningful group constructs that will facilitate interpretation and sensemaking. Specifically, the purpose of this goal is two-fold: *first*, the contextualisation of visualisations by linking constructs with visual elements; and *second*, the scaffolding of multimodal evidence that invite the interpretation and sensemaking of one insight at a time.

1.4 RESEARCH CONTRIBUTIONS

The contributions of this thesis comprise three approaches to design, model and implement explanatory multimodal interfaces (see Figure 1.1- Contributions):

1. *HCD-MMLA: A Human-Centred Design Multimodal Learning Analytics approach to formalise technical, analytical and visual requirements from co-design sessions.* This contribution assists in the formalisation of outputs from co-design sessions into technical, analytical and visual requirements to design multimodal interfaces. This procedure is formalised by developing a *mapping instrument* to help researchers, designers and practitioners to identify: i) group work

constructs, sub-constructs and behavioural markers; ii) sensors needed to capture those behavioural markers; iii) multimodal metrics and analytics derived from the sensors; and, iv) potential design elements that can be included a dashboard (Chapter 4).

2. *Multimodal Matrix: A multimodal matrix modelling approach to contextualise low-level multimodal group data with group constructs.* The purpose of this approach is to provide a conceptual and computational model that transforms low-level data into meaningful data representations by considering group constructs. Group constructs, who are defined as concepts, skills and attitudes that students externalise during learning episodes, can be identified from knowledge expertise or grounded by literature in educational research. These constructs can be observed as students' behavioural markers, with each behaviour associated with one or many multimodal observations. The multimodal observations (i.e. meaningful representations) are derived from sensor data and stored into a matrix-like representation, where each column is a meaningful observation (e.g. position data can be translated into meaningful zones of movement) and each row corresponds to the time when that observation occurred. The resulting matrix can further be analysed to extract patterns or visualise group behaviours and performance (Chapter 5).

3. *EvisLA: An explanatory visual layered approach to facilitate the interpretation and sensemaking of multimodal data.* The third contribution advances the design and implementation of explanatory multimodal interfaces, by highlighting key aspects from the learning context that are relevant in the multimodal group data. This is accomplished by aligning visualisations with the intended learning context. The learning context can be grounded from theory, the learning design, learning outcomes or teachers' intentions, as demonstrated in the operationalisation of the HCD-MMLA approach. At the same time, the multimodal group data can be transformed into meaningful data and visualisations, by applying the Multimodal Matrix approach. Then, each visualisation is enriched with visual elements to drive attention to important aspects of the activity (Chapter 6). A layered user interface enables the user to filter and combine different visualisations, each of which focuses on a specific 'data story' (Chapters 7).

1.5 THESIS STRUCTURE

The remainder of this thesis is structured as follows:

Chapter 2: Background and Related Work. This chapter provides an overview of the provision of feedback in CSCL and CSCW fields, specifically, the growing interest in designing feedback solutions in authentic learning environments and the exploration of learning visualisations as way to support feedback in collocated groups. Moreover, this chapter identifies the current gap in the literature motivating the need to investigate the design and validation of learning visualisations to support feedback by communicating insights about group work using multimodal learning analytics.

Chapter 3:: Learning Contexts and Research Methodology. This chapter describes two learning contexts: 1) a database group work activity; and 2) teamwork in nursing simulations. The underlying multimodal data gathered from these contexts, and resulting visualisations were further used in exploratory studies described in Chapter 5, 6 and 7 to directly address the research questions. It also explains how a Design-Based Research methodology has been harnessed for conducting this research.

Chapter 4: Mapping from Multimodal Data to Group Constructs and Analytics Requirements. This chapter addresses RQ1 by presenting the approach to formalise MMLA requirements to design multimodal interfaces. The chapter documents the application of the approach, in the context of the nursing teamwork simulations (learning context 2), to translate the outcomes of student/teacher co-design sessions into more formal requirements for multimodal interfaces.

Chapter 5: Giving Meaning to Multimodal Group Data. This chapter addresses RQ2 by introducing the Multimodal Matrix. To illustrate the purpose of the Multimodal Matrix approach, this chapter details its application to the teamwork simulation scenario (learning context 2), which led to the generation of four visualisations. Finally, this chapter provides insights into the views of teachers and students on the usefulness of the generated visualisations and makes further recommendations for improving the visualisations.

Chapter 6: Exploring the Potential of Explanatory Visualisations. This chapter addresses RQ3 by describing a *learning design-driven data storytelling* approach. This chapter illustrates the approach by re-designing a set of visualisations in the context of a database group work activity (learning context 1), resulting in a set of explanatory visualisations. It also reports the results of two studies to validate the visualisations. The first preliminary study gives recommendations for improvements of explanatory visualisations with potential users. The second study reports teachers' reflections on the usefulness of data storytelling elements from the exploration of a set of improved explanatory visualisations.

Chapter 7: A Multimodal Layered Approach for Teamwork Reflection. This chapter addresses RQ3 by describing the EvisLA approach, which aligns the outcomes from the Multimodal Matrix and the learning design-driven data storytelling approaches. This chapter illustrates the applicability of this approach by re-designing and generating an explanatory layered interface in the context of teamwork nursing simulations (learning context 2). It then reports a capstone study of this thesis involving a guided walkthrough of the interface with nursing simulation experts to elicit their assessments of its potential.

Chapter 8: Discussion and Conclusions. This chapter summarises the contributions related to each research question, draws conclusions and indicates some of the implications of the research findings. Limitations and challenges of the contributions and suggestion for further research in this field are considered.

Chapter 2: Background and Related Work

This chapter presents the background and challenges in the research field that motivate the design and validation of feedback for collocated groups to guide reflection. First, this chapter describes the background and foundations of group work, and feedback in collocated settings, along with associated challenges. Then, it reviews current state in analytics for group work to support the understanding of group's interactions and performance, including the use of multimodal data to analyse and visualise collaboration. Next, it highlights the challenges of creating interfaces that guide the interpretation and sense making of multimodal data consumed by non-experts' users, such as teachers and students. Finally, it reviews the current state of learning dashboards and discusses the challenges which this thesis aims to address through the design and validation of explanatory multimodal learning analytics interfaces that communicate key aspects of the group work activity to guide reflection.

2.1 INTRODUCTION

A growing body of research has examined the provision of personalised feedback mechanisms to coach groups, aiming at improving the teaching and learning experience. At the same time, there has been an emergent interest in exploring new ways for capturing and analysing group's interactions and performance in collocated settings through multimodal data, mostly to help researchers understand different aspects of collaboration, support learning theories or distinguish high achieving from low achieving groups. However, there is a relative paucity of empirical research focusing on designing and exploring how multimodal data captured from collocated group work activities can help teachers and students guide the reflection of groups' skills and practices. As sensors and technology improve and become more readily available, researchers should exploit the underlying data effectively by providing guidance during the exploration of multimodal and complex data, which in turn may be difficult for teachers and students draw conclusions about the group activity and performance.

This thesis considers different areas of research in order to design effective feedback for collocated groups. Figure 2.1 shows the key areas of research that are relevant to this thesis and are reviewed in this chapter. This background chapter provides a discussion of existing work in Computer Supported Collaborative Work (CSCW) and Computer Supported Collaborative Learning (CSCL) that tackles the provision of feedback for collocated group work, specifically at the intersection of

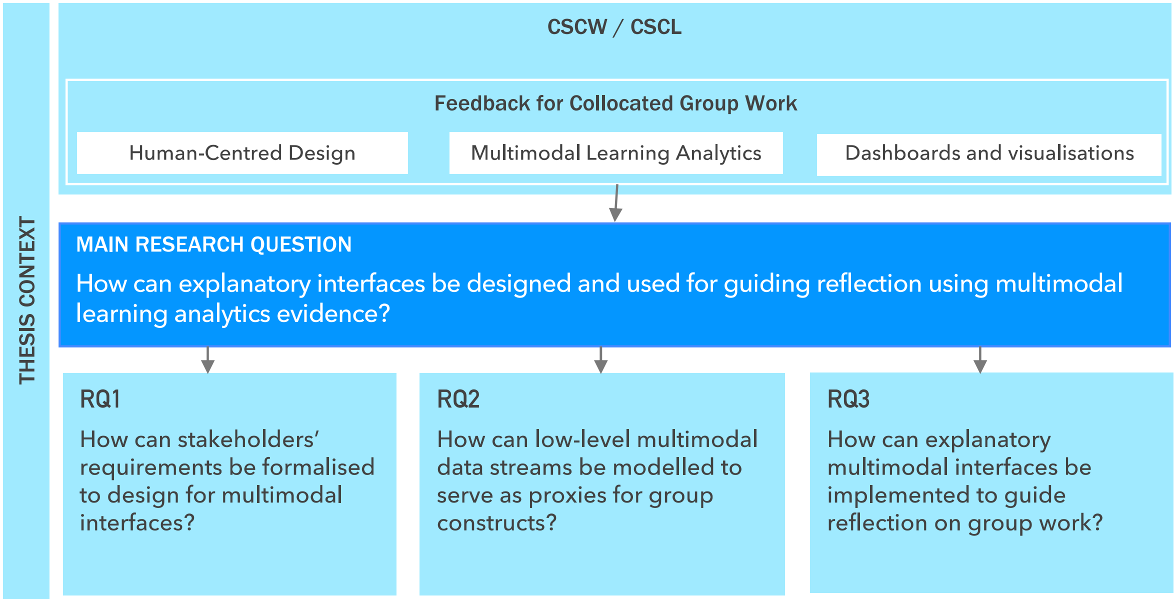


Figure 2.1: Partial diagram that represents the research context and key research questions this thesis addresses.

Human-Centred Design, Multimodal Learning Analytics, and Learning Analytics dashboards and visualisations.

In line with the main research question presented in Section 1.2, this chapter introduces foundational research on group work and feedback in collocated classroom settings and the challenges that are imposed by the learning environment. This chapter also reviews existing literature in analytics for group work to support the understanding of group's interactions and performance, including the use of multimodal data. Furthermore, this chapter presents a review of current learning dashboards to support group work and highlights the challenges and gaps of designing and producing learning dashboards using multimodal data.

This chapter is structured as follows. First, Section 2.2 introduces the background and foundational research in group work and emphasises the challenges that teachers and students face in these group work activities. Then, Section 2.3 describes the background and foundational research in feedback for collocated group work and highlights the challenges in the provision of feedback. Next, Section 2.4 reviews the current state-of-the art in analytics for group work, discusses how multimodal learning analytics may be used to analyse and visualise group work comprehensively and highlights the challenges and gaps of harnessing multimodal group data to generate interfaces to support teaching and learning. Section 2.5 discusses the relevant research in learning dashboards that support collocated group work activities, describes the techniques that may help guide the interpretation and visualisation of complex multimodal group data and lists the challenges and gaps in research about investigating learning dashboards that exploit and guide the interpretation of multimodal group data. Finally, Section 2.6 summarises the challenges and issues that are addresses in this thesis.

2.2 BACKGROUND AND FOUNDATIONS OF GROUP WORK

2.2.1 What is group work?

In both professional and educational contexts, it is usual to work in groups. This is because some tasks are complex, often requiring multiple people to contribute with their expertise and skills to be accomplished successfully. A ***group*** can be defined as *"two or more individuals who are connected by and within social relationships"* (Forsyth, 2009, p. 3). Many varieties of groupwork configurations can be found in the literature. For example, families, sport teams, associations and colleagues at school are considered groups (Beebe & Masterson, 2003; Johnson & Johnson, 1991). Therefore, many other characteristics are also relevant to define a group, such as:

- *Interdependence:* members in a group are influenced by each other in terms of their actions, outcomes, and feelings. Interdependence between individuals exists to some extent or degree, depending on the group activity. In **low-interdependent** groups, members are less influenced by other member's choices or actions (e.g. groups working in brainstorming ideas); while in **high-interdependent** groups, members actions are influenced into a greater extent (e.g. firefighting or soccer teams).

- *Goal-oriented:* all groups have a purpose or goal. This goal may vary, depending on the activity to be performed. Therefore, group members dedicate all their efforts to attain and accomplish the goal.

- *Interpersonal interactions:* when two or more people work together, interpersonal interactions typically occur with the purpose of achieving the task. This interaction may happen in the form of verbal or non-verbal communication. The Interaction Process Analysis (IPA) classifies two types of group interactions: **task** and **relationship** interactions. Task interactions are actions or behaviours that contribute to the group's purpose. Relationship interactions are related to emotional connections among group members, such as supporting, resolving conflicts or undermining criticism.

- *Sense of belonging:* group members should perceive that they belong to a group, and this group acts as a unit. This also refers to the group's cohesiveness, which relates to the strength of links that tie each member to the group.

- *Structure:* groups do not exist merely by joining people, as some type of structured relationship may be present. Interactions among group members bring organisation to the group by implicitly or explicitly defining roles and norms. A **role** prescribes some general behaviours and a set of actions a member could exhibit (e.g. leader, follower). **Norms** can be set in a group as a list of behaviours that members should or should not perform.

- *Size:* The number of group members determines the interdependence and connections in a group. For instance, **small groups** are likely to be more connected, and members are able to keep track of the interactions among each other member. In contrast, **large groups** tend to be less connected, as it may result complex for each member to keep track of the many interactions that could be devised.

While in the literature the terms ***group*** and ***team*** are sometimes used interchangeably (cf. Cohen & Bailey, 1997; Kozlowski & Ilgen, 2006), some works have point to some differences when referring to each other. As Forsyth (2009) argued: "*Not all groups are teams. Teams require more from the members in the way of collaboration and coordination"*. In seminal works on teamwork science, a ***team*** is referred as: *"two or more individuals with specified roles interacting adaptively, interdependently, and dynamically toward a common and value goal, within an specific time span*" (Salas, Burke, & Cannon-Bowers, 2000, p. 341). A common distinction refers to the role definition. In teams, members have more differentiated roles, with their specific goals and set of skills that need to be fulfilled collectively towards the solution of a task. In contrast, groups may have less defined roles, expecting that members will contribute individually to the task according to their prior knowledge. Another difference between groups and teams relies on the members interdependency. A team is highly interdependent, which means that each individual action is influenced by other member's actions in a coordinated manner. On the contrary, groups are less interdependent, meaning that each member is responsible of individual actions and contributions without affecting other members actions in a greater extent. Finally, while groups goals may be broad or discussed in general terms before or during the activity, in teams, goals are clear, specific and are well-defined before the activity starts (Beebe & Masterson, 2003).

A review presented by Zoltan and Vancea (2015) suggested that a team can be considered as a specialised type of group (see for example Cohen & Bailey, 1997; Paulus, 2000). In general terms, a team can be defined as a structured group working with well-defined goals requiring coordinated interactions; and mutual trust and accountability to complete certain tasks (Forsyth, 2009). In this thesis, the term *team* will be used to describe a group where specific roles have been defined, each member has been assigned with a role, members exhibit a high interdependency on their actions and a shared goal has been defined towards the completion of the activity. In addition, this thesis is focused on small group work (e.g. 3-7 people), either working in teams or groups.

2.2.2 What is Collaborative Learning?

Collaborative Learning is an educational approach in which two or more people learn or attempt together (Dillenbourg, 1999; Gokhale, 1995; Johnson & Johnson, 1986). Dillenbourg's seminal work (1999) stated that collaborative learning it is not just a mechanism or a method. Instead, the author argues that the nuances of collaborative learning are defined by a situation and the interactions determined by the group work activity. For instance, a group work activity can promote a more collaborative situation (e.g. face-to-face small groups working in a math problem) or a less collaborative situation (e.g. in an online setting where more than 40 students are exchanging ideas in a forum). Likewise, group members could perform more collaborative interactions (e.g. negotiation, task execution) or less collaboratives ones (e.g. giving instructions, seeking for information).

In a learning context, collaborative learning activities are implemented in classrooms by teachers to get students work together in groups or teams. During a collaborative activity, group members display different behaviours such as sharing perspectives, exchanging arguments, clarifying meanings, negotiating conflicts and formally externalising knowledge at a personal and group level (Panitz, 1999; Stahl, 2006). It has been demonstrated that collaborative learning can foster critical thinking (Gokhale, 1995). It can also promote the development of cognitive, social, and metacognitive skills (Dillenbourg, 1999; Järvelä et al., 2015; Stahl, 2006).

Collaborative learning involves a myriad of approaches, such as cooperative learning (e.g. jigsaw, pairs check), problem-centred instruction (e.g. cases studies, simulations), discussion groups (e.g. open-ended discussions) that can be implemented in formal and informal learning settings (Smith, 1992). It may also contain a well-defined (i.e. specific goals and expected solutions) or ill-defined (i.e. no clear goals or expected solutions) task or problem that students try to solve. Independently of the collaborative learning approach or collaborative task, this thesis takes a broad perspective by investigating different learning scenarios and strategies, focusing on the effective provision of feedback in face-to-face group work activities. For purposes of clarity, the term group work will be used interchangeably to refer to teamwork or collaboration.

2.2.3 What makes a group effective?

Collaborating in collocated (f2f) settings provides unique benefits that are not easy to achieve in digitally mediated forms of group work (Johnson, Sutton, & Poon, 2000; Olson, Teasley, Covi, & Olson, 2002). Literature suggests that the rich, multimodal communication channels in f2f interaction can promote social bonding (Nardi & Whittaker, 2002), trust building (Wilson, Straus, & McEvily, 2006), increase creativity (Gloor et al., 2012; Nardi & Whittaker, 2002) and productivity (Olson et al., 2002). Moreover, the ability to communicate, be an effective team/group member and collaborate are key skills for employability in the 21st century workplace (de Lima & de Souza, 2017) and a competency that needs attention in undergraduate programs (Buckingham Shum & Crick, 2016).

Research in small groups has focused on the analysis of effective groups in function of their cognitive, affective and behavioural processes that enable members to combine their resources to solve a task (Cohen & Bailey, 1997; Kozlowski & Ilgen, 2006). Group effectiveness can be measured by group's performance, attitudinal or behavioural outcomes (Cohen & Bailey, 1997).

Communication, collaboration and coordination strategies have been identified as key behavioural processes that can lead to effective group work (Dillenbourg, 1999; Salas et al., 2000). When group members work together, they must clearly communicate thoughts, support each other's ideas, assume leadership (implicitly or explicitly) participate equally or according to their roles. However, not all groups are effective despite the benefits that working together may bring (Barron, 2003). Communication issues, social loafing, unpreparedness, unequal participation and free riding are some of the problems that most groups commonly face when the activity unfolds (Riebe et al., 2016).

Previous works have identified that effective group's processes (e.g. planning, communication, monitoring, etc.) and behaviours (seeking for information, providing feedback to peers, etc.) makes the difference between successful and unsuccessful groups (e.g. Rousseau et al., 2006; Salas et al., 2005). Therefore, it is important for tutors and groups to understand the underlying processes and outcomes that lead to an effective and high performing group. This could help tutors to facilitate group effectiveness, or to support groups in the selection of strategies for accomplishing current or future goals. Salas' seminal work (2005) suggested a list of five core mechanisms that promotes group work effectiveness namely leadership, mutual performance monitoring, backup behaviour, adaptability and orientation, which are operationalised across domains. In addition, Salas and colleagues (2015) further suggested that the context where the activity unfolds also influences effective group processes and performance. Therefore, it remains a challenge to elicit context specific group performance metrics because it requires a full understanding of the task for each group activity (Gilbert et al., 2018).

2.2.4 Challenges in collocated group work

Even though literature suggest many benefits for collocated group work activities, there are still some challenges that hinders the adoption of these activities by negatively influencing teachers and students' satisfaction and preventing effective collaboration learning strategies to students:

First, collocated group work has remained an ephemeral activity due to the challenging task of capturing and analysing the many interactions occurring physically. Due to large classes, time allocation, and content coverage (Riebe et al., 2016), it is challenging for teachers and facilitators to provide close coaching and foster beneficial collaborative interactions (Gillies, Ashman, & Terwel, 2007).

Second, significant effort has been invested in exploring digital traces of online experiences, where logs can be easily captured, to make group interactions visible, infer patterns of behaviour and broaden the understanding of group activity in various contexts (Martinez-Maldonado, Kay, et al., 2019). By contrast, much more needs to be done to provide effective group work support using objective evidence to promote the understanding of groups' processes and facilitate group coaching.

Third, personalised support on group practices is essential to help groups attain their goals (Gilbert et al., 2018). Due to the many group's interactions happening as the activity unfolds, teachers may not be able to notice all behaviours or misconceptions of individuals. In addition, formal practices to provide support on group's processes or performance tend to occur occasionally, or very late in the learning process (e.g. final subjects in a professional program) (Kozlowski & Ilgen, 2006). Therefore, it is important to detect where a group, or a member is struggling in order to provide effective, actionable and timely support.

To sum up, this section brought the foundational concepts of group work, collaborative learning and what makes a group effective. Several challenges have been identified in the literature that hinders a proper nurturing of effective group work skills. This can be summarised as the use of evidence towards guiding groups to become effective. The next section describes how group's evidence can be formalised into feedback mechanisms for supporting the improvement of group's performance.

2.3 FEEDBACK FOR COLLOCATED GROUP WORK

2.3.1 What is feedback?

Feedback can be broadly defined as any type of information provided by teachers, peers or external agents intended to improve learners' performance (Hattie & Timperley, 2007). In particular, high quality feedback should include information about the intended goals, information about the current student's performance, and information that can help students to develop strategies to close the gap between their current performance and a desired goal. Educational designers should consider the provision of effective feedback before, during and after a task to ensure that students' learning outcomes are met. Thus, students have the opportunity to demonstrate their learning progress. The literature suggests that the provision of feedback has positive effects in learning when it is given effectively (i.e. avoiding criticism or negative and personal comments) and timely (immediate or delayed) to students (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006). Teachers commonly perceive the provision of feedback as an opportunity to communicate the actual state of students and compare it with a desired state. From a students' perspective, it is an opportunity for reflecting upon the activity and adjust for subsequent performance (Boud & Molloy, 2013b).

Effective feedback encourages students to posit their learning along these questions: 1) *"where am I going?* 2) *how am I going?* and, 3) *where to next?"* (Hattie & Timperley, 2007). Moreover, as suggested by Nicol (2010), to improve the quality of feedback, it should be:

- **understandable**: learners can understand easily the information being provided;

- **specific and selective**: the information provided should describe only two or three key points that could lead to the improve of learner's performance;
- **timely**: delivered on time (immediate or delayed) to give learners enough opportunities to reflect on the information provided;
- **transferable**: information provided should not be focused only on the task but also on the process and;
- **balanced**: the information should point at both positive and negatives areas of improvement.

In sum, the provision of effective feedback should specify: *i) When the feedback is given*: before, during or after the learning task or problem. For example, a pre-task with its corresponding feedback could make students aware of their current knowledge and what is needed to improve the performance of the actual task. Feedback given during the activity (i.e. real-time) allow students to regulate their current knowledge and change it according to the expected learning goals or outcomes. Finally, feedback provided after (e.g. post-hoc) the activity evaluates students' final outcomes with the expected outcomes. *ii) What information is given*: task-oriented (e.g. correct or incorrect response) or process-oriented (e.g. learning processes required to complete a task). *iii) How the feedback is given:* using different media to communicate current students' learning states depending on the learning environment. For example, students may receive verbal feedback during a class activity, a written feedback in a homework activity (e.g. an essay), and an audio or video recorded in an online course. Finally, iv) *To whom the feedback is targeted:* at an individual level (e.g. written feedback to each student) or at a group level (e.g. verbal feedback to a whole classroom). Furthermore, when giving feedback at a group level, it should be considered that the feedback can be directed to individual members, the whole group or a combination of both (Gabelica, Bossche, Segers, & Gijselaers, 2012; London & Sessa, 2006). The next section reviews literature specifically about the provision of feedback to collocated groups.

2.3.2 Providing feedback to collocated groups

This section revises three important aspects aiming at providing feedback to collocated groups: a) how feedback is given in classrooms; b) what type of feedback is commonly given to groups and; c) to whom the feedback is targeted.

2.3.2.1 *How feedback is given in classrooms?*

Strategies to provide feedback to groups in everyday classroom practice includes teachers' observations and peer- and self-feedback (Marzano, Norford, Paynter, Pickering, & Gaddy, 2001; Strijbos, 2010; Strijbos & Sluijsmans, 2010). During a collaborative activity, teachers are encouraged to *observe* and monitor current groups' interactions and performance in a glimpse and provide immediate feedback to groups regarding their current learning state. However, this can result in a partial view of the group performance, causing students' dissatisfaction and impeding the correct fostering of group skills (Riebe et al., 2016). To address this issue, digital tools have been developed to support teachers' observations of groups' interactions and skills. An example of this is Group Spinner (Kharrufa, Rix, Osadchiy, Preston, & Olivier, 2017), a tool that enables teachers to record, visualise and compare in-class observations of different aspects of the learning activity and group's interactions at different times. Nonetheless, contextual information about the learning activity is not recorded by this tool, impeding teachers to provide feedback about task performance.

On the other hand, a good teaching and learning practice is to engage students into the feedback loop. Peer- and self-feedback is an effective approach to help students reflect on their progress (Strijbos & Sluijsmans, 2010). Mostly, teachers provide to students a rubric to evaluate their group processes or performance, during or at the end of the activity (Ohland et al., 2012). However, teacher should also consider objective evidence to provide a comprehensive feedback that not only rely on subjective evidence.

Once that teachers have gathered information about group's performance, the next step is to inform groups about any detected gap between the actual and the expected performance to improve collaboration efficiently. This is done by comparing the information gathered from observations or self-report measures and the learning goals and outcomes established a priori (Taras, 2005). Also, the feedback can be addressed at an individual level (i.e. to each group member) or at a group level (i.e. all members without distinction).

Within an educational context, feedback is distinguished either as formative or summative (Strijbos, 2010; Taras, 2005). Formative feedback, referred as assessment for learning, aims to provide a comprehensive view of the group's learning process, involving cognitive, social and motivational aspects of collaboration. Formative feedback interventions can occur several times during the collaborative activity. In contrast, summative feedback, denoted as assessment of learning, intends to evaluate how well the group or members performed by assigning a single score. This feedback is considered isolated from the learning process, focusing on cognitive aspects of collaboration and is commonly given at the end of the activity by the teacher.

Depending on the learning design and outcomes of the activity, feedback can be targeted at an individual-level, group-level or classroom level. Different strategies have been suggested for this purpose (Marzano et al., 2001). For example, teachers may provide individual feedback during the enactment of the activity by asking a task-related questions to a random group member. Also, teachers may opt to give group feedback by assigning a single score to the group's solution once the activity finished (i.e. summative assessment; Strijbos, 2010). Further, teachers may deliver a rough classroom feedback based on the observations and inquiries from groups during the collaborative activity.

2.3.3 Challenges in the provision of feedback to collocated groups

Previous research has demonstrated the potential of group feedback to enhance group effectiveness, if it is clear, timely, directly to groups, and all groups receive the same attention (Gabelica et al., 2012). However, this practice sometimes is not realistic. In authentic learning contexts, teachers found it a demanding and a time-consuming task. This is given due to several factors:

First, large classes may hinder the provision of effective feedback, as teachers cannot be present in all interaction's instances of all groups. Thus, feedback provided during teachers' observations does not always reflect the group's performance in its totality, but only an excerpt of what happened (Riebe et al., 2016).

Second, peer- and self- feedback may not be the best option to provide formative and summative feedback. The feedback given by peers may be biased, as peer's and self- perceptions could overestimate the real group performance (Riebe et al., 2016) or peers may not be experienced to provide formative feedback.

Third, it is challenging to provide formative feedback to groups, as it can be hard for teachers determine cognitive, social and motivational aspects of collaboration through sporadic observations. Literature points out that it is important to provide feedback not only on cognitive aspects of collaboration, but also on social and motivational aspects, as these are critical for the future workforce {Salas, 2015 #699. For summative feedback, most of the teachers assess groups only on their task-related aspects of collaboration by assigning a single score on the final solution. This may cause discrepancy among group members if not all of them equally contributed to solving the task.

In sum, given the limitations of these conventional group feedback approaches, collocated collaboration analytics hold promise as a way to gather and analyse objective data from classrooms, by using pervasive technology and up to date devices (Martinez-Maldonado, Kay, et al., 2019). In this way, this thesis aims to provide feedback tools based on evidence gathered from collocated group work activities. The next section revises literature and current work on group work analytics, and how it can be used to provide effective feedback to groups.

2.4 STATE-OF-THE-ART IN ANALYTICS FOR GROUP WORK[3]

2.4.1 How collocated group work is framed?

In the last two decades, Computer-Supported Collaborative Learning (CSCL) and Computer-Supported Cooperative Work (CSCW) research have widely investigated the adoption of technology to understand, analyse and support collaboration in online and face-to-face contexts (Dillenbourg, 1999; Jeong, Hmelo-Silver, & Yu, 2014; Schmidt & Bannon, 2013; Stahl, 2006). Empirical research in

[3] Parts of this section have been published in CHI'19 (Echeverria, Martinez-Maldonado, & Shum, 2019).

CSCL and CSCW has adopted different perspectives, theoretically grounded, to analyse the complexity of group work activities by decomposing it into multiple dimensions. For example, the Interaction Process Analysis (Bales, 1950) was proposed as an approach to analyse group processes by categorising an individual or group behaviour into four major areas: *positive reactions* (socio-emotional), *attempted answers* (task), *questions* (task) and *negative reactions* (socio-emotional). The Activity Theory Framework (Lewis, 1997) describes a group activity through relationships (links) of six elements (nodes): *tools*, *subject*, *object*, *rules*, *community* and *division of labour*. The activity-based computing (ABC) framework (Bardram, 2005) decomposes group activity into *tasks*, *materials*, *time*, and *users*. The Blended Interaction framework (Jetter, Reiterer, & Geyer, 2014) structures the CSCW design space into four spaces: *individual* and *social* interaction, the *task*, and the physical *space*. More recently, an approach based on the Activity-Centred Analysis and Design (ACAD) framework (Martinez-Maldonado, Goodyear, Kay, Thompson, & Carvalho, 2016) provided a three-dimensional view of group activity: (1) *physical*, which includes the physical and digital space and objects; input devices, screens, software, material tools, furniture; (2) *epistemic*, which includes both implicit and explicit knowledge oriented elements that shape the participants' tasks and working methods; and (3) *social*, which includes the variety of ways in which people might be grouped together (e.g. dyads, trios); scripted or emerging roles, and divisions of labour. In addition to these three dimensions, (4) *affective* aspects have been identified in foundational theoretical work (Vygotski, 1987) as critical to understand CSCW work and teaching and learning (Miller, 2005), even though it can often remain *invisible* (Roth, Tenenberg, & Socha, 2016). Figure 2.2 depicts a graphical representation of the

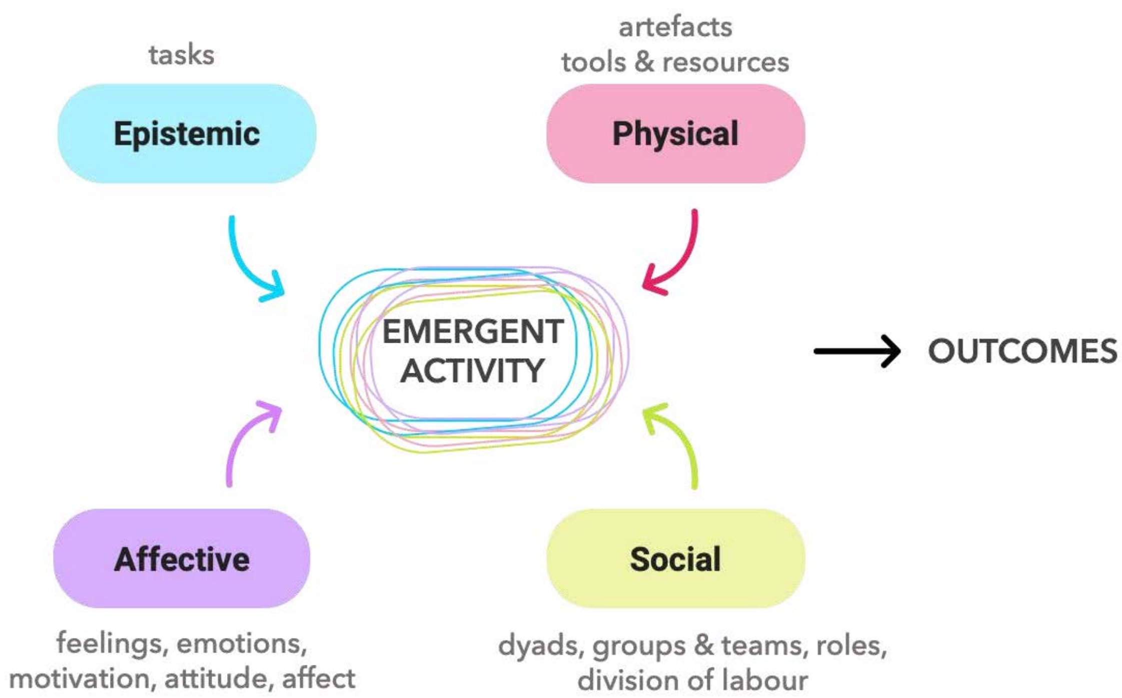


Figure 2.2: Graphical representation of adapted version of the ACAD framework (Goodyear & Carvalho, 2014) to analyse collaboration in terms of *social*, *epistemic*, *physical* and *affective* dimensions.

adapted ACAD framework to analyse collaboration in terms of *social*, *epistemic*, *physical* and *affective* dimensions. The emergent activity unfolds the many relationships and interactions among these four aspects resulting in group outcomes of the activity, which can be related to the initial expectations, learning goals or any other anticipated intention chosen by the teacher or the learning designer.

It is worth noting that this section does not intent to argue about the advantages and disadvantages of each of these frameworks. Instead, this section provided a succinct overview of the frameworks that have been used in CSCL and CSCW research to analyse collaboration through different lenses. Further, this thesis used an adapted version of the ACAD framework to analyse collaboration in terms of *social*, *epistemic*, *physical* and *affective* dimensions.

2.4.2 How collocated group work is commonly analysed?

There has been a growing interest in exploring the value of group data to generate understanding of both observable and hidden patterns in f2f settings by capturing behavioural traces through sensors and devices. For example, substantial effort has been posed in extracting quantitative *non-verbal speech* features from dialogue in contexts such as brainstorming (Kim, Chang, Holland, & Pentland, 2008), problem-solving (Mohan, Sun, Lederman, Full, & Pentland, 2018; Oviatt & Cohen, 2013; Viswanathan & VanLehn, 2017b), and group meetings (Bachour et al., 2010; Müller, Huang, & Bulling, 2018; Nakano, Nihonyanagi, Takase, Hayashi, & Okada, 2015; Nihei, Nakano, & Takase, 2017; Sturm, Herwijnen, Eyck, & Terken, 2007).

Other substantive works have considered a wider range of aspects of collocated group work. Some studies have analysed the relationship between traces of *physical activity* such as proximity (Martinez-Maldonado, Echeverria, Yacef, Dos Santos, & Pechenizkiy, 2017), motion (Cukurova et al., 2018) and posture (Kim et al., 2008) and students' competencies during group tasks. Traces of *head motion* have been used to investigate how group members pay attention to objects and people in close proximity; effective social dynamics (Sturm et al., 2007); the author of meaningful utterances (Nihei et al., 2017), or low rapport (Müller et al., 2018). *Face* and *hand tracking* features, extracted from videos, have been used to analyse affective states (Neubauer, Woolley, Khooshabeh, & Scherer, 2016), physical formations (Spikol et al., 2017) and synchronicity (Müller et al., 2018). Other 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, Reilly, & Schneider, 2018; Haataja et al., 2018). *Gaze* features have been used to identify participation styles (Nakano et al., 2015) and assess the quality of group work (Schneider & Pea, 2017). Multitouch screens and input devices can also serve to log group activity (Martinez-Maldonado & Goodyear, 2016). For example, *touch events* have been be used to predict group performance (Martinez-Maldonado, Yacef, et al., 2015; Viswanathan & VanLehn, 2017b) and *stroke* features from digital pens have been used to identify expertise (Oviatt & Cohen, 2013) and to assess participation (Nakano et al., 2015).

Overall, the studies presented above show how group data may serve to complement the analysis of collocated group activities. Yet, most prior work has decomposed the complexity of group work, focusing on *particular aspects of collaboration* (i.e. physical, epistemic, social and affective), with a few exceptions described below. The next section describes how multimodal learning analytics can be applied as an approach for comprehensively analysing group work data.

2.4.3 Multimodal Learning Analytics for supporting collocated group work

Learning Analytics (LA) is a research area focused on conducting data-driven analysis from an educational context aimed at improving teaching and learning practices and understanding learning processes (Siemens & Baker, 2012). Since its adoption in the last decade, several methods, tools and analysis have been proposed for many formal and informal learning settings in which students, teachers, researchers or administrators are the main actors and beneficiaries of these analytical solutions (Ferguson, 2012).

In the broad spectrum of LA, Multimodal Learning Analytics (MMLA) has emerging as a promising subfield, focusing on capturing and analysing learning traces coming from multiple sources in authentic and complex scenarios, which are not necessarily supported by computers (Blikstein, 2013; Ochoa, 2017). *Multimodality* refers to a theory to understand the different channels of communication or '*modalities*' (e.g. voice, gesture, writing) people use to interchange information (Kress & Van Leeuwen, 2001). In a learning context, multimodality aims to investigate how multiple modalities influence teaching and learning practices (Jewitt, 2006). Also, multimodal analysis has been of particular relevance in education as an approach to analyse comprehensively behaviours and interactions at different levels (e.g. teacher, students, objects) and correlate these computationally extracted behaviours with learning theories or as a triangulation approach to reduce the workload of researchers by automatically coding teachers and students' behaviours for posterior analysis (Blikstein, 2013). Current multimodal approaches in classroom-based research include: i) supporting teaching and learning through multimodal systems (Jewitt, 2008; Lotherington & Jenson, 2011) and; ii) understanding learners' interactions during different learning activities, which are complex by nature (Blikstein & Worsley, 2016).

There is a growing body of research that has started to apply MMLA approaches into experimental and authentic collocated scenarios (Chua, Dauwels, & Tan, 2019; Di Mitri, Schneider, Specht, & Drachsler, 2018; Kim, Sottilare, Brawner, & Flowers, 2018; Ochoa, 2017). MMLA approaches in collocated scenarios have opened up opportunities to collect multimodal group traces to automatically generate group metrics (e.g. Martinez-Maldonado, Kay, & Yacef, 2013), contribute to the elucidation of collaboration theories (e.g. Malmberg et al., 2018), visualise the learning phenomena (Noroozi et al., 2018) and provide effective feedback to enhance awareness and reflection (e.g. Kim et al., 2008). The following body of literature reviews research where multimodal data has been used for 1) analysing collaboration and 2) generate multimodal interfaces, either in a broader context or in teaching and learning environments.

2.4.3.1 *Multimodal data for analysing collaboration*

Previous research has examined collaboration traces for understanding group's processes and performance in controlled groups discussions, brain storming tasks and meetings scenarios, beyond a learning context (see a review in Gatica-Perez, 2009). This large body of research has investigated how features extracted from multimodal data could help to understand social dynamics (e.g. equity) and model collaboration behaviours (e.g. collaborative vs. non-collaborative episodes, participation styles, rapport) For example, Sturm et al. (2007) investigated how visual multimodal data influenced the social dynamics during meetings. Speaker activity and gaze direction were automatically calculated and displayed during the group meeting as circles in a shared space. Neubauer et al. (2016) correlated speech utterances; facial and emotional expressions; and heart variability with team's cohesion to understand resilience to stress in a computer-based high workload task. In another study, Nakano et al. (2015) assessed participation styles in problem-solving discussions using manually annotated gaze targets, speech utterances, annotated speech interactions and writing actions. Ward, Pirkl, Hevesi, and Lukowicz (2017) detected instances of ad-hoc physical collaborations between group members using speech and physical movement measures (e.g. energy, zero-crossing count, mean, stdev) during a construction scenario. In a recent study, Müller et al. (2018) identified low rapport during natural interactions within small groups discussions from speech activity, facial expressions, hand motion, gaze and combined speech and gaze features.

In *collaborative learning contexts*, previous studies have examined collaboration in relation to group's behaviours (e.g. coordination, regulation) and learning, automatic identification of roles (e.g. experts), and assessment of the quality of collaboration. For instance Oviatt and Cohen (2013) correctly predicted all dominant experts students while solving math problems collaboratively from the Math Data Corpus, which included digital pen traces, speech and video features and annotated data. Using the same corpus, Ochoa and colleagues (2013) determined that simple features such as writing speed and key math terms are good discriminators of student's expertise (i.e. experts and non-experts).

Using interactive tabletops as a computer-supported technology for collaboration, Martinez-Maldonado (2014) aimed to analyse collaboration occurrences using speech features and tabletop's touch logs. Verbal and physical participation; and indicators of symmetry of talk and physical actions were extracted from students while working in concept mapping activities to model collaborative and non-collaborative situations and mine patterns to find collaboration strategies.

Using a similar approach, Viswanathan and VanLehn (2017a) extracted speech features and tablet gestures from dyads working in a math problem and model episodes of collaboration using machine learning techniques. Using Tangible User Interfaces (TUIs) and a depth camera, (Schneider & Blikstein, 2015) extracted body movement features, touch-based logs and TUI actions from dyads working in a science problem and predicted student learning gains. Schneider and colleagues (2016; 2018) further investigated dyads' visual coordination strategies in a problem-solving tasks from TUI logged actions, speech and eye-tracking features (e.g. gaze, pupils' dilatation, fixation, saccades). In an experimental scenario, Grover et al. (2016)'s preliminary study attempted to assess collaboration of dyads working in a pair-programming task. From body posture, eyes gaze, and clickstreams, Grover and colleagues proposed to model the quality of collaboration (i.e. high, medium or low).

Another strand of collocated collaboration studies have looked at collaboration quality from small groups interventions in open-ended tasks, such as collaborative problem solving (CPS) activities. Spikol et al. (2017) analysed face and hand tracking features; and physical interaction with physical objects from triads working in a CPS guided activity, aimed to identify which features are better to predict collaboration success using machine learning techniques. Using human coded data, Spikol and colleagues reported that distance between hands and between learners are crucial to predict student's performance. In a recent follow up study, Spikol et al. (2018) utilised the same dataset to present a comprehensive analysis with advanced machine learning techniques. Similar to the results presented before, authors found that hands and face distance between students are key to predict the quality of the final group solution.

Affective aspects of collaboration have been examined in a lesser extent. In a recent study, Malmberg et al. (2018) identified regulation learning episodes during a problem-solving collaborative task using physiological arousal and facial emotion expressions from multimodal data collected from EDA sensors and high-quality videos. Malmberg and colleagues demonstrated a high correlation between high-level interaction (i.e. confusion moments) and regulation learning episodes. In another study, Worsley and Blikstein (2018) analysed gesture, audio and bio-physiological data and manual-coded data gathered from dyads working in a hands-on activity to find correlations between multimodal features and the design quality and learning outcomes.

Overall, the studies presented above show how multimodal learning analytics may serve to complement the analysis of f2f activity. Table 2.1 presents an overview of these studies in relation to the group work activity, multimodal data captured, the type of analysis performed and its collaboration aspects. Most prior work has de-composed the complexity of group work, focusing on *particular aspects of collaboration*. Yet, some works have explored the potential of analysing multiple sources of data associated with more than one dimension of collaboration, such as *social and epistemic* (Oviatt & Cohen, 2013; Schneider et al., 2016), *social and affective* (Malmberg et al., 2018), *social and physical* aspects (Cukurova et al., 2018; Spikol et al., 2017), or social, physical and epistemic aspects (Spikol et al., 2018; Spikol et al., 2017). Most previous works have been conducted

Table 2.1: Overview of studies analysing collocated groups using multimodal data in learning contexts

Group work activity	Data	Analysis	Reference
triads working on math problems	digital pen traces, voice, video, human annotations	predicting experts and non-experts	(Ochoa, Chiluiza, Méndez, et al., 2013; Oviatt & Cohen, 2013)
small groups working in concept mapping activities	voice, tabletop's touch logs	modelling collaborative and non-collaborative situations mining collaboration strategies	(Martinez-Maldonado, 2014)
dyads working in a math problem	voice, tablet logs, human annotations	predicting collaborative and non-collaborative situations	(Viswanathan & VanLehn, 2017a)
dyads working in a science problem	body movement, tabletop's touch logs, TUI logs, human annotations	predicting student learning gains	(Schneider & Blikstein, 2015)
dyads working in problem-solving tasks	TUI logs, voice, eye-tracking, human annotations	correlating visual coordination strategies	(Schneider et al., 2016; Schneider, Sharma, et al., 2018)
dyads working in a pair-programming task.	body posture, eye-tracking, app logs, human annotations	predicting collaborative and non-collaborative situations	(Grover et al., 2016)
triads working in a CPS activity	Facial expressions, hand movements, physical actions, human annotations	predicting collaboration success and quality of the final group solution	(Spikol, Ruffaldi, Dabisias, & Cukurova, 2018; Spikol, Ruffaldi, Landolfi, & Cukurova, 2017)
small groups working in a CPS activity	Hand movements, head direction	Students' CPS competence levels	(Cukurova, Luckin, Millán, & Mavrikis, 2018)
triads working in problem-solving task	physiological arousal, facial emotion expressions, human annotations	correlating group regulated learning situations	(Malmberg et al., 2018; Noroozi et al., 2018)
dyads working in a hands-on activity	gesture, voice, physiological data, human annotations	correlating design quality correlating learning outcomes	(Worsley & Blikstein, 2018)

under controlled conditions (Cukurova et al., 2018; Dich et al., 2018; Malmberg et al., 2018; Spikol et al., 2017) where it is easier to isolate aspects of group work. However, there is a growing body of research attempting to bring multimodal innovations into authentic scenarios, such as in classrooms (Evans, Wobbrock, & Davis, 2016; Martinez-Maldonado, 2014; Mohan et al., 2018; Oviatt & Cohen, 2013) and the workplace (Müller et al., 2018; Nihei et al., 2017).

2.4.3.2 *Multimodal data for generating multimodal interfaces*

In terms of multimodal interfaces to support the provision of feedback in a broader context, prior research has mostly been limited to mirroring collaboration indicators (e.g. speech time, participation) during group activities to raise awareness about group's behaviour (Soller, Martínez, Jermann, & Muehlenbrock, 2005). Most of these systems have offered simple representations of non-verbal speech features, to enhance awareness and accountability of verbal participation (Bachour et al., 2010); support self-reflection (Tausch, Ta, & Hussmann, 2016); or promote social regulation (Kim et al., 2008). Not many mirroring systems have included other modalities of data. For instance, a relevant work on implementing and generating real-time feedback to participants in small group meetings using multimodal information was presented by Sturm et al. (2007). They proved that speaking interaction and gaze behaviour indicators, can be displayed as concentric circles in a shared area to influence participation. In another study, Kim et al. (2008) investigated the effect on social dynamics while visualising members speech participation and physical interactivity in a personal display. From speech features and body movement features, authors implemented a network-like graph interface, where the nodes represented each group member and links denoted interactivity estimated from physical proximity and speaking time.

Moving towards *multimodal interfaces to support teaching and learning*, few studies have reported the development of tools to externalise group's processes, quality of collaboration or trends in group's activity (Chua et al., 2019). For example, in small groups working on a concept mapping task, Martinez-Maldonado, Yacef, et al. (2015) estimated several collaboration indicators from speech utterances and touch logged actions, from which they generated four different visualisations: *i)* a gauge chart to inform detected overall collaboration among members; *ii)* a radar chart showing the symmetry of touch and speech activity overlapped for each participant; *iii)* interactions with objects (i.e. links, notes) among participants differentiated by a colour (e.g. red, green, yellow as for participant 1, 2 and 3, respectively); and *iv)* tabletop actions, speech utterances and partial group solutions displayed in many timelines (one timeline per participant and per behaviour and one timeline for partial group solutions). These visualisations were generated in a post-hoc exploration aimed to support *researchers and developers* to get insights about high-performing and low-performing groups' collaboration strategies.

In a healthcare simulation task, Martinez-Maldonado, Power, et al. (2017) reported two different visualisations for exploring team's formations and team's learning strategies. The first visualisation depicted team's audio and mobility using heat maps, while the second visualisation illustrated key team's actions performed during the enactment of the activity in a timeline. Authors suggested that heat maps could provoke reflection on the strategies followed by the team and support *teachers* on coaching teams by providing evidence on team's behaviours, which remained invisible before. Also, timelines could offer evidence to *teaching staff* for improving teaching practices by comparing the expected learning design and the executed learning activity.

In a recent preliminary work, Ochoa, Chiluiza, et al. (2018) presented a multimodal transcript visualisation in the form of several vertical timelines of actions to support the assessment of group collaboration. These actions, which can be considered high-level data representations, were extracted from tabletop touch interactions, gaze and speech data. A preliminary evaluation with teachers reported that the multimodal visualisation was able to communicate group's rapport consistently and that teachers were able to evaluate groups based on the information presented in the visualisation.

Finally, a recent work reported a toolkit to support the analysis and visualisation of multimodal data (Noroozi et al., 2018). Specifically, the toolkit includes a dashboard with several timeseries visualisations plotting low-level data recorded from physiological sensors and video and annotated data from human observations. While this toolkit has been planned to be used by researchers to facilitate the visualisation of bio-signals and the analysis of regulation strategies in collaborative settings, authors mentioned that teachers and students may use this tool in a near-future.

Table 2.2 provides an overview of the studies described above in relation to the different types of visual representations and target audience focusing on particular aspects of collaboration. Overall,

Table 2.2: Overview of studies on generating visual multimodal interfaces for collocated groups in learning contexts

Visual representations	Target users	Reference
Detected collaboration in a gauge chart Symmetry of touch and speech activity in a radar chart interactions with other's objects in a network graph Timeline with tabletop actions, utterances and group solution	Researchers, developers	(Martinez-Maldonado, Yacef, & Kay, 2015)
Mobility and audio heat map Key actions in a timeline	Teachers, teaching staff	(Martinez-Maldonado, Power, et al., 2017)
Vertical timeline with tabletop actions, gaze, verbal interventions and emotions	Teachers	(Ochoa, Chiluiza, et al., 2018)
Multiple timelines showing temperature, EDA, BVP and Heart rate per subject Timeline with combined data Timeline with annotated actions Video	Researchers	(Noroozi et al., 2018)

these studies illustrate how multimodal data can be used to generate multimodal interfaces to support the provision of timely post-hoc feedback. However, more empirical evidence is needed to investigate if the visual design choices and information is relevant for teachers, individuals and groups. Furthermore, it is challenging for teachers and students to engage with complex multimodal low-level data visualisations, as these may fail to communicate or guide the interpretation of meaningful learning insights. The next section addresses relevant literature pointing to the need of giving meaning to multimodal low-level data.

2.4.4 Giving meaning to multimodal data

A significant body of HCI work makes use of behavioural traces of human activity captured through sensors or input devices (Dumais, Jeffries, Russell, Tang, & Teevan, 2014; Lazar, Feng, & Hochheiser, 2017). These traces can range from low level logs, such as clickstreams, to non-mediated human action, such as eye movement or gestures. Logs have the benefit of being easy to capture, at scale, without observers influencing the activity and the capture process (Dumais et al., 2014). Logs can be mined, for instance, to cluster user behaviours (Liu et al., 2017), identify archetypal users (Wang, Zhang, Tang, Zheng, & Zhao, 2016), and visualise common paths (Liu et al., 2017). However, while logs can illuminate *what* users do, they often say much less about *why* (Dumais et al., 2014). This is a critical methodological challenge if HCI is to develop principled ways to make sense of the vast quantities of interaction data now available.

A small but growing body of literature in the field of *Learning Analytics* focuses on the question of how one maps "from clicks to constructs" — how low-level system logs can serve as proxies for the higher order constructs that educators and students can understand (Shaffer, 2017; Shute & Ventura, 2013; Wise & Shaffer, 2015). For example, far from the 'big data revolution' signalling the 'death of theory' (Mazzocchi, 2015), Wise and Shaffer (2015) argue that when datasets are so large that spurious statistically significant patterns can be obtained easily, theory is even more important to guide interpretation. Approaches to bridging the traditional divide between quantitative and qualitative methodologies are now being developed and validated, such as *quantitative ethnography* (Shaffer, 2017).

Theoretically-motivated analytics can be designed in a principled manner for higher order constructs such as students' "conscientiousness" (Shute & Ventura, 2013), or "crowd-sourced learning" capacity to learn in a MOOC (Milligan & Griffin, 2016). These and most other examples are from contexts where a single modality of clickstream is associated with higher level constructs.

In collocated scenarios, where multiple streams of data in different modalities can be captured from each person, giving meaning to the data logs is an even more challenging task. Current approaches to give meaning to multimodal data include: *i)* comparing multimodal traces of activity with human-annotated video data (e.g. see review in Di Mitri et al., 2019); *ii)* automatically coding multimodal data according to epistemic frames (e.g. Worsley & Blikstein, 2018); and *iii)* examining co-occurrence of events or associating certain indicators of activity with collaboration, task performance and learning outcomes (e.g. Dich et al., 2018; Haataja et al., 2018; Spikol et al., 2017). However, as argued by Shaffer (2017): *"if we want to integrate different sources of data, [their representations] have to span a similar amount of the semantic space, of the meaning, that we are attributing to the data"* (p. 148). While previous approaches attempted to give meaning to multimodal data at some extent, a more holistic and systematic approach that connects low-level multimodal group data with meaningful data representations and considers the complexity of group work is needed in order to make visible group work interactions and performance.

2.4.5 Challenges and gaps

Overall, these studies provided clear evidence for the potential of MMLA approaches to analyse collocated group work activities comprehensively and generate timely multimodal interfaces. Such studies helped identify the following gaps in this research field:

First, there is a need for **empirical studies that apply and evaluate MMLA approaches to support teaching and learning strategies**. Current state-of-the-art in multimodal analysis have demonstrated the feasibility of gathering and analysing multimodal data coming from several sensors in classroom-based scenarios. However, most of these investigations have been carried out for research purposes to extend the understanding of collaboration theories or different aspects of collaboration connected to learning outcomes and performance. To advance in the learning analytics field, MMLA studies should be linked with learning theories; and teaching and learning practices to potentially identify pitfalls and personalise learning.

Second, there is a need for the **development of effective MMLA feedback interfaces to coach collocated groups**. As mentioned above, effective feedback has the potential to impact learning positively. Current works have been limited to the development of *mirroring tools* to support students' awareness, yet these only supports partly group's behaviour and performance. By contrast, research should move towards the development of *guiding systems*, which are a way of coaching effectively groups, by proposing helpful information regarding current and desired learning outcomes.

Third, there is a need for **empirical studies and examples to support the provision of MMLA feedback in collocated groups**. As indicated above, there is a lack of empirical studies and tools that support the provision of feedback in collocated groups using multimodal data. This could be partly because most of the efforts have been devoted to data gathering, integration and modelling challenges in authentic scenarios. However, much work is needed to develop and validate novel MMLA feedback interfaces tailored to teachers and student's needs.

Fourth, there is a need for **approaches to give meaning to multimodal data and generate effective MMLA feedback.** The design of multimodal interfaces plays an important role on generating effective MMLA feedback, as multimodal data from collocated groups can result in complex low-level visual interfaces, hindering the communication of relevant information. Moving beyond solving the technical data fusion problem; it is necessary to provide mechanisms to represent multimodal data in a compelling way and provide tailored feedback to teachers and students.

The next section covers current research and challenges in learning dashboards to support group work, and background literature to design effective learning dashboards and offer a guided interpretation to teachers and students.

2.5 GUIDING THE INTERPRETATION OF LEARNING DASHBOARDS[4]

2.5.1 Learning dashboards

In recent years, educational dashboards and visualisations (which are also referred to as "visual learning analytics"; Duval, 2011) have gained considerable attention in the field of learning analytics (see reviews in Bodily & Verbert, 2017; Schwendimann et al., 2017). Indeed, for some people, learning analytics are synonymous with the dashboards that teachers (Molenaar & Knoop-van Campen, 2017) and students (Schumacher & Ifenthaler, 2017) are increasingly encountering in a wide range of educational products. Borrowing from big data and web analytics in other domains, a learning analytics dashboard aggregates indicators, from one or more sources, about students' activity and/or learning, using one or more visualisations (Schwendimann et al., 2017; Verbert et al., 2013). *Teacher-facing dashboards* are intended to help educators gain a better understanding of their whole course, reflect on their teaching strategies and identify students who require specific attention (Molenaar & Knoop-van Campen, 2017; Verbert et al., 2013). Similarly, *student-facing dashboards* are intended to help students reflect on aspects of their learning behaviour and potentially assist them in, for example, managing time effectively, accessing key learning resources, or gaining a richer picture of their progress (Bodily & Verbert, 2017; Reimers & Neovesky, 2015).

Although *learning dashboards* and other visual learning analytics have been getting significant traction in recent years (Bodily & Verbert, 2017; Schwendimann et al., 2017), there have also been numerous reports pointing to the limitations and possible pitfalls of rolling out these products without further research and development work (e.g. Aguilar, 2016; Corrin & de Barba, 2015; Jivet, Scheffel, Drachsler, & Specht, 2017; Reimers & Neovesky, 2015; Teasley, 2017). The design of effective dashboards in educational contexts is complex. First, a recent review pointed to the absence of design choice justifications reported in several research outputs presenting educational dashboards usage (Bodily & Verbert, 2017). A current review finds poor evidence of grounding in the learning sciences (Jivet, Scheffel, Specht, & Drachsler, 2018). Another survey reports that both commercial and scientific dashboards commonly feature poor interface design and lack of usability testing, and the choice of what data is to be visualised does not commonly correspond with what students and educators really want or need (Reimers & Neovesky, 2015), in part, because they are not regularly consulted as part of the design process (Holstein et al., 2017).

[4] Parts of this section have been published in LAK'18 (Echeverria, Martinez-Maldonado, Granda, et al., 2018) and in the Journal of Learning Analytics (Echeverria, Martinez-Maldonado, Chiluiza, Granda, & Conati, 2018).

In parallel to the visual and design factors noted above, researchers and designers can easily overlook the learning context and the audience for whom these visualisations have been created (Schwendimann et al., 2017). This may cause a disparity between users' and designers' perspectives. Sometimes, designers and researchers want to communicate multiple insights or dimensions of data about students' experience. This can lead to the design of complex visualisations that are often hard to interpret especially "at a glance" (Duval, 2011). Other times, the effectiveness of several learning dashboards relies on teachers' or students' ability to understand graphs and make sense of (often complex) information in order to take meaningful actions or change behaviour (Greller & Drachsler, 2012). Moreover, teachers and students are encouraged to interpret these visualisations in a limited time and, even if the data can be interpreted correctly, they may fail to understand the "call to action" for reflection, or adapting their behaviour.

2.5.2 Learning dashboards to support group work

There has been considerable work exploring the value of learning dashboards in collaboration environments as tools aimed at promoting participation (e.g. DiMicco & Bender, 2007; Janssen, Erkens, Kanselaar, & Jaspers, 2007), group awareness (e.g. Gergle, Kraut, & Fussell, 2013; Janssen, Erkens, & Kirschner, 2011; Martinez-Maldonado, Clayphan, Yacef, & Kay, 2014; Phielix, Prins, Kirschner, Erkens, & Jaspers, 2011) and shared regulation (e.g. Järvelä et al., 2015). Data collected from CSCL environments can be at an individual and group level and dashboards can be used by teachers, individuals or groups. In CSCL, *student-facing dashboards* prompt individuals and groups to actively regulate their learning, raise awareness to promote behavioural changes and fine-tune their learning strategies for future activities. *Teacher-facing dashboards* offer orchestration opportunities by monitoring individual and group activities to take informed decisions about individual and group learning outcome (van Leeuwen, Rummel, & van Gog, 2019).

Several tools have been developed over the past years with the purpose of supporting social visibility, accountability and awareness in collaborative environments (Erickson & Kellogg, 2000; Erickson et al., 1999; Liu & Nesbit, 2020). These tools can be distinguished according to its purpose in the collaboration process: 1) mirroring the current state of interaction; 2) comparing the current state of interaction to the desired state; and 3) offering advice or guidance, to teachers, individuals or the whole group (Soller et al., 2005).

The first type of tools are *mirroring tools* to increase group members awareness of their actions during a collaborative activity (Soller et al., 2005). It is expected that teachers and students reflect upon this information to take informed decisions and make changes on their actions accordingly. Usually, aggregated data is displayed in real-time as the activity unfolds, through charts or metaphors in a shared display or personal displays. These tools have proven effective in visualising and keeping track of group and task performance, social presence and individual and group's actions to facilitate coordination, increase their productivity and reduce errors (Salas, Cooke, & Rosen, 2008). One of the most prominent works of mirroring tools in collocated environments is *Second Messenger* (DiMicco et al., 2007), a group mirroring tool (Figure 2.3) which captured group members' speech and displayed a histogram in a shared display revealing member's participation aimed at *influencing group's behaviour.*

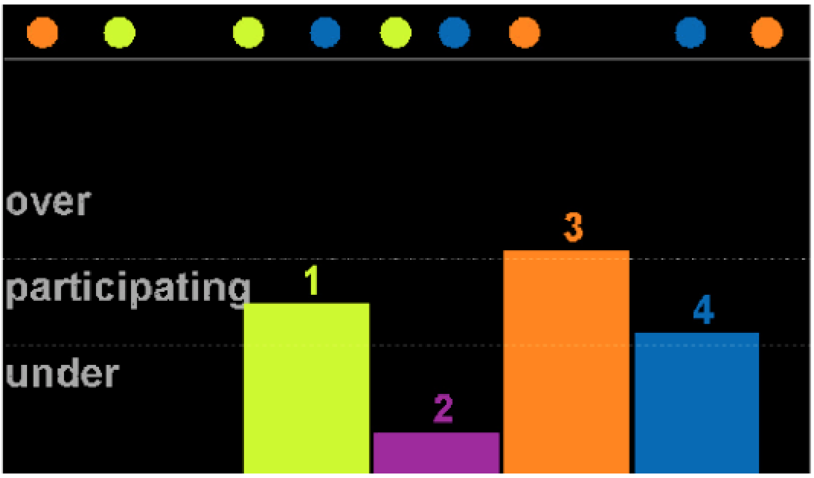


Figure 2.3: ***Second Messenger*****, a mirroring visualisation tool displaying group members' speech during a meeting session (DiMicco, Hollenbach, Pandolfo, & Bender, 2007)**

The second type of tools are *alerting tools* which are intended to compare group's current behaviour with a desired condition. Teachers may find these tools very useful for orchestrating collaborative activities as alerting tools can inform critical situations where students need more attention. In collocated environments, for example, Martinez-Maldonado, Clayphan, et al. (2015) designed and deployed the *MTFeedback* tool to enhance teacher's awareness of collaborative activities in authentic classroom environments (Figure 2.4). MTFeedback generated automatic visualisations and notifications to *direct teachers' attention more effectively* and give informed interventions and tailored feedback about epistemic aspects of the task. MTFeedback showed one visualisation per each group's solution and its matching proportion with an expert solution. It also displayed a notification square around each group visualisation indicating the group with more errors (e.g. red square) or the group with half of the solution similar to the expert and no errors (e.g. green square).

Finally, the third type of tools includes *advising or guiding tools* aimed at providing additional information that could guide the interpretation of the displayed information. These tools can support and facilitate a teacher-led reflection to guide students towards effective collaboration. Soller et al. (2005) reviewed advising tools deployed in distributed and online learning environments. These tools prompted information to facilitators or students to improve groups' social or task-related aspects of the activity. However, there is a lack of research in the design and validation of guiding tools for collocated collaboration, and broadly, in the learning analytics field.

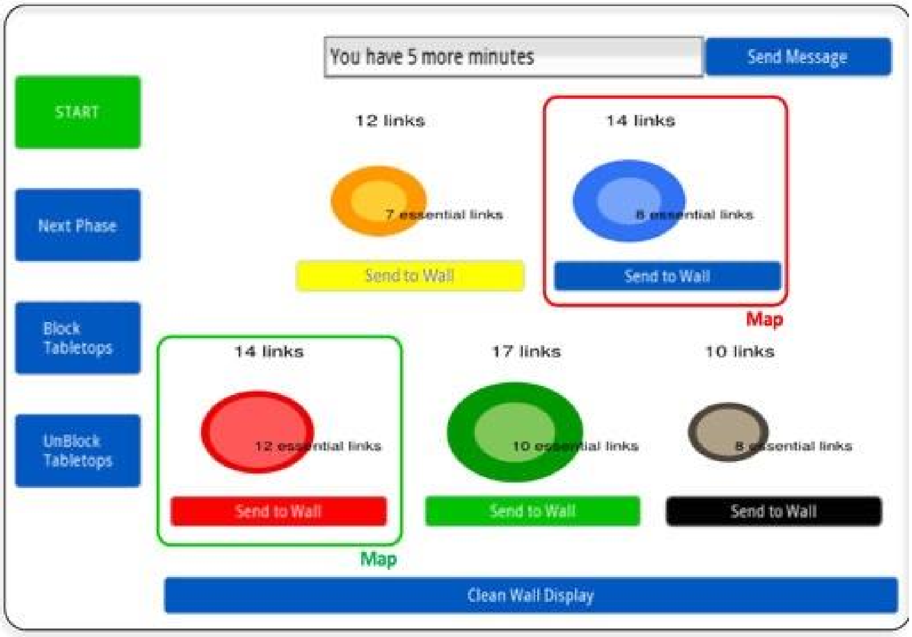


Figure 2.4: *MTFeedback,* an alerting visualisation tool to enhance teacher's awareness. The tool identifies groups with conceptual errors (red square) and groups with no errors (green square) and alerts the teacher (Martinez-Maldonado, Clayphan, Yacef, & Kay, 2015) .

2.5.3 Providing guidance through visual elements

The analytical challenge of visualising complex and heterogeneous data coming from different sources and facilitating the communication of insights have gained attention in past years (Thomas & Cook, 2006). Designing visual analytics and dashboards that effectively guide the interpretation and sensemaking has been the focus of many practitioners. There is a substantial body of Information Visualisation (InfoVis) research since data analytics dominates decision making in many non-educational sectors (Zhang et al., 2012). Research in InfoVis acknowledges that guidance could support users in the exploration and interpretation of data visualisations (Ceneda et al., 2017). Guidance refers to *"error messages, alarms, prompts, and labels, as well as to more formal instructional material provided to help guide a user's interaction with a computer… and to support users of different skills levels"* (Smith & Mosier, 1986, p. 291). This definition highlights the use of ***visual elements*** to guide user's interaction in a computer-supported system. Using guidance to design user interfaces allows users to be informed about their current status, for example, what a user did, where a user is and whether a user successfully completed a task. Likewise, these features of guidance could be applicable to MMLA interfaces (Vieira, Parsons, & Byrd, 2018), where common dashboards includes timeseries plots at different scales (Noroozi et al., 2018). The next section introduces data storytelling as an InfoVis solution to facilitate the guidance and interpretation of complex visualisations, such as MMLA interfaces.

2.5.4 Data storytelling as a technique to facilitate guidance

Schulz, Streit, May, and Tominski (2013) characterised different aspects of guidance to support insights generation and user's decision making such as *orienting* the user to explore the data through visual cues; *directing* the user to explore relevant data by offering predefined alternatives; and *prescribing* a data exploration workflow through a well-defined set of small tasks. A well-known InfoVis technique that endorses these aspects of guidance is *data storytelling*.

Data storytelling (DS) can be defined as an *"information compression"* technique, to help an audience focus on what is most important to communicate via data visualisation (Ryan, 2016). Data storytelling builds on classic InfoVis guidelines (Tufte & Schmieg, 1985) and narrative storytelling foundations (e.g. plots, twists and calls to action; (Lee et al., 2015). Tufte and Schmieg (1985) suggested that a visualisation should incorporate *"graphical excellence"*, meaning that it has *"to reveal data with clarity precision and efficiency"*. In the same way, Ryan (2016) stated that the goal of using visuals is *"to communicate a key message clearly and effectively"*, emphasising the context and meaning through visual enhancements (e.g. changes in size, colours, saturation). While visualisations are intended either to *explore* or *explain* insights, DS is focused on the latter. The goal of DS is to communicate insights through the combination of *data, visuals* and *narrative* (Dykes, 2015). Thus, DS is not applicable to exploratory visualisation interfaces. Instead, DS is aimed at explaining what is happening in the data and why a feature is important. In this context, DS has been widely used by journalists as a means to engage and inform through *"interactive graphics and narratives"* (Segel & Heer, 2010). Current reported uses of DS techniques have focused on helping presenters tell more compelling stories through a clear (written or spoken) narrative supported by data visualisations (e.g. showing to an executive board why they should invest in a STEM program for kids; Knaflic, 2015); communicating scientific data (Ma et al., 2012); and teaching dynamic networks in an educational context (Bach et al., 2016).

Ryan (2016) and Knaflic (2015) have identified a set of *"golden principles"* that can be applied when crafting a story with data. Although both authors use different languages, they coincide in the following DS principles:

DS1. Data storytelling is goal oriented. As discussed above, the data visualisation or dashboard needs to be aligned with a purpose or intention. This can help designers and researchers to have 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, e.g. it is left entirely up to them which filters to apply, which graph type to use, or which region to focus on.

DS2. Data storytelling drives the audience's focus of attention. Visual and narrative elements should be used to create meaning in the visualisation. Adding meaning to visualisations can be accomplished using specific elements (e.g. lines, weight, shapes, size, colours, contrast) to emphasise key aspects to direct visual attention. Narrative text is a salient visual element, providing summaries of visual features.

DS3. Data storytelling relies on choosing an appropriate visual. Certain visualisation techniques work better for certain purposes. For example, line charts can effectively show changes over time (Ryan, 2016). In contrast, Knaflic (2015) dedicates a whole chapter to justifying why pie charts should not be used.

DS4. Data storytelling relies on core InfoVis design principles. Clutter in data visualisations adds visual complexity to the graph and makes it harder to find insights (Knaflic, 2015; Ryan, 2016). Tufte and Schmieg (1985) refer to this as maximising data-ink ratio. *Decluttering* is achieved by removing elements (e.g. unnecessary headers, chart features, borders, grids) that do not add informative value to the graph. Other design aspects are important, for example, Knaflic (2015) proposes the addition of *titles* to drive the intent of the visual, and *captions* to explain important features of the data. General principles related to *alignment, use of colour, shape* and *texture* are information design decisions that can have a strong impact on sensemaking.

Knaflic (2015) describes how to improve graphs by adding or removing visual elements with the aim of supporting data storytelling. These data storytelling elements are explained along with an illustrative educational example. Figure 2.5 (top) presents the typical output of a visualisation tool (Google spreadsheets in this case) that was used to plot a fictional dataset regarding the number of visits per module in an online class during the past week, with the next day projection. Figure 2.5 (bottom) presents the same data visualisation with data storytelling enhancements inspired by Knaflic's work (with additional annotations A–E for expository purposes). A critical insight that can be emphasised is the increased number of visits for Module 4, which is more than the half of visits that other modules, considering that Module 4 is the most important and complex module of the semester. Comparing both graphs, it can be noted how this message is better understood in Figure 2.5 (bottom), without getting lost in exploring the visualisation.

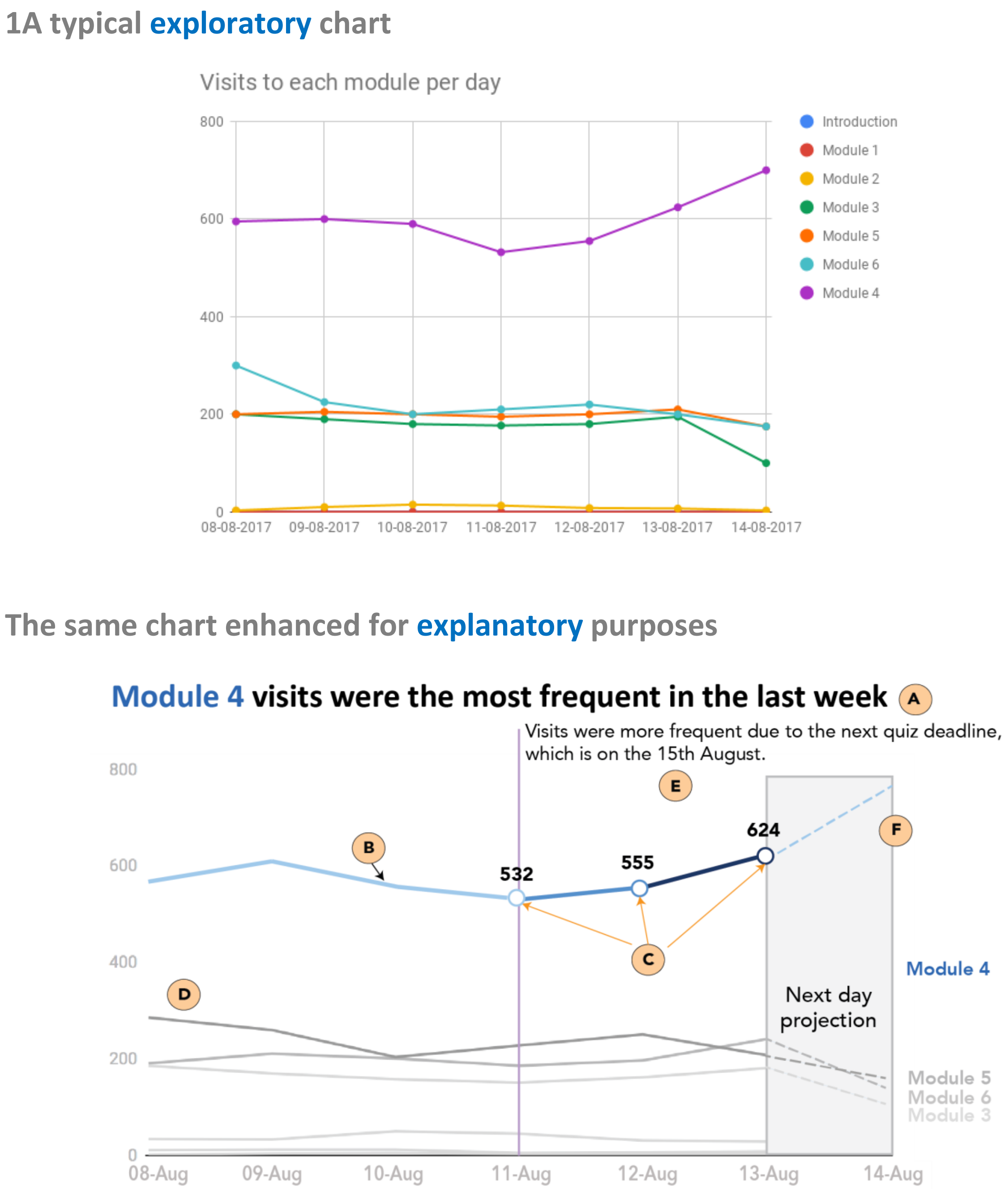


Figure 2.5: *Top*: A typical exploratory chart without data storytelling elements. *Bottom*: An explanatory chart of the same data with data storytelling elements, including: A) a Prescriptive *Title*; B) a *Thicker line* to highlight particular data series; C) only critical *Data points*; D) *Decluttering* by de-emphasising less relevant data; E) Add key contextual data with *Labelled line*; F) a *Shading area*.

For instance, the following data storytelling design elements contribute to communicating one story in this example:

- A) A **prescriptive title** provides a succinct message that is intended to explicitly communicate the main message of the visualisation. Since multiple stories may be told based on the same data, the title may serve to directly communicate the insight that should be sought in the current visualisation.

- B) **Highlighting specific data series** (e.g. by making the line of one data series thicker) can help the audience to focus on the data that is relevant to the current story (in this case, the 'Module 4' data series).
- C) **Highlighting specific data points** and removing irrelevant data points makes explicit which data support the claim in the title.
- D) **Decluttering** by removing grids, eliminating indirect legends, reducing the number of colours used, and removing non-relevant data point markers result in minimising distractions. This also involves deemphasising data series that are not at the centre of the story (in this case, other data series such as 'Module 5, 'Module 6', 'Module 3', appear in light grey).
- E) **Narrative text** added directly to the visualisation in the form of labels can help to gain a better understanding of or explain interesting changes in the data. In the example, the text next to the vertical line pointing at the day '11-Aug' in the horizontal axis ("*Visits were more frequent due to the next quiz deadline…*") explains the increasing trend in the 'Module 4' data series.
- F) A **shaded area** can be used to enclose those data points associated with the same insight. In the example, the shaded area divides the graph into actual data and data projections.

2.5.5 Challenges and gaps

As mentioned above, providing guidance to learning dashboards could potentially help teachers and students in understanding complex learning visualisations, specifically to visualisations that use multimodal data. However, current research has made evident that much more work is still needed to adapt current InfoVis techniques into the design of effective MMLA interfaces. Therefore, some gaps have been identified from the literature reviewed above:

First, there is a relatively small body of research that investigates **the design and validation of dashboards for collocated group work**, even though they provide rich opportunities to coach groups by visualising groups' cohesion, social awareness to foster individual or share regulation, support reflection. While much work has been done in online contexts (e.g. open learner models; Bull & Kay, 2016), where digital traces are easier to capture, more examples of learning dashboards to effectively coach collocated groups are required. Current work in LA research is moving towards introducing Human-Centred Design approaches to support the development of learning analytics (Buckingham Shum, Ferguson, & Martinez-Maldonado, 2019). These approaches encourage researchers and designers to involve teachers and students into the design process of learning dashboards which may lead to a higher adoption and acceptance of the LA solution.

More studies investigating how teachers and students design, interpret and use learning dashboards in collocated groups would allow the generation of effective learning dashboards for monitoring collaborative activities in the classroom and guiding the reflection of groups' skills and strategies.

Second, there is a need for **approaches to contextualise DS principles and elements into educational data**. Current research in LA have started to move towards a more Human-Centred Design approaches to deliver effective data visualisations. Aligning learning design with learning dashboards has been identified as a key issue within the learning analytics community, that requires collective effort (Bakharia et al., 2016; Lockyer, Heathcote, & Dawson, 2013). More work is needed to correctly contextualise learning (e.g. learning design, educational theories, activity) into data visualisations. The critical question that follows is thus: given some generic data storytelling principles and design elements, how can these be explicitly contextualised to educational contexts for crafting explanatory visual learning analytics?

Third, there is a need for **empirical evidence and examples to design MMLA interfaces (i.e. multimodal learning dashboards) that offers a guided reflection using multimodal group evidence targeted to teachers and students in authentic collaboration scenarios**. Displaying low-level data coming from different sources in a learning dashboard may result in convoluted visualisations, hindering the elucidation of collaboration processes and performance, particularly for non-data experts such as teachers and students. It is essential that teachers find relevant information quickly to interpret and suggest changes in students' learning strategies. The challenge is then, encoding and compressing complex and heterogenous data into more meaningful data visualisations to guide teachers and students in the interpretation of such complex data (Vieira et al., 2018) using well-established InfoVis techniques such as data storytelling.

2.6 SUMMARY

Collocated group work activities are important for the future work force and it is of important matter to provide mechanisms to help groups learn to collaborate. Up to now, collocated group work activities have remained ephemeral, due to classroom constraints, which are beyond teaching and learning practices. As it has been noticed over the last few years, there has been growing evidence on how multimodal learning analytics (MMLA) could help capture and analyse group's traces in authentic scenarios where the actual learning is happening. However, the literature suggests that most of MMLA solutions have addressed the complexity of learning in controlled environments. Furthermore, MMLA researchers have put all their efforts to deliver solutions that supports the operationalisation of multimodal data. Yet, little has been done to tackle the next wave in MMLA research, which is related to closing the feedback loop by offering multimodal interfaces to monitor group processes and performance and guide teachers and students into the reflection of their practice to take informed decisions for further improvement.

Designing and delivering multimodal interfaces implies not only the analysis of different means of group data, but also the representation and encoding of low-level data to inform teachers and students meaningful and tailored information of learning processes. This involves three open issues.

The first issue is related to the lack of **approaches to inform the design requirements for multimodal interfaces**. As pointed out in Section 2.3.3, empirical research tackling the design of effective MMLA interfaces is in early stages. Most of the approaches considered for developing MMLA interfaces have not made the design process explicit. Due to the complexity of multimodal low-level data, it is essential to formally define the learning context, type of data sources, analytics and visual representations into requirements that could help designing and developing effective MMLA interfaces. HCD approaches could make this possible by bringing experts, researchers, teachers and students into the design of tailored feedback.

The second issue is related to the **representation and encoding of multiple streams of data into meaningful information**. Low-level multimodal group data per se may not be useful and easy to interpret by teachers and students. For instance, displaying EDA indices in a timeline or plotting all position data points may not support the understanding of collaboration. Thus, providing a higher-level representation of the same data could help in the interpretation and sensemaking of group's processes and performance.

Third, there is a need of **approaches to contextualise the information presented in a multimodal interface**, or more specifically, in a learning dashboard. Research in LA indicates that learning dashboards should be contextualised based on learning theories and the learning design of the activity to reduce misinterpretations produced by one-size-fits all dashboards. Moreover, if the ultimate goal is to provide guidance to teachers and students and communicate insights, it is of important matter to focus the attention on what is really important. Therefore, researchers should consider teachers and students perspectives to inform the design choices of multimodal interfaces. Research in HCI and CSCL suggest that teachers and students should be involved in the whole design process, as main actors.

These three issues emphasised here will be addressed in the following chapters.

Chapter 3: Learning Contexts and Research Methodology

This chapter provides details of the research methodology followed to investigate the research questions outlined in Chapter 1. First, this chapter provides details about the two learning contexts and primary data collection in which the research was conducted. The first learning context concerns nursing team simulations, while the second concerns collaborative database design. Second, this chapter describes how a Design-Based Research approach was adopted to design and validate explanatory multimodal interfaces to support and guide group work reflection.

3.1 LEARNING CONTEXTS AND PRIMARY DATA COLLECTION

This section describes the two authentic learning contexts where this research was applied: *i)* a teamwork nursing simulation and *ii)* a collaborative database design activity. The teamwork nursing simulation context was designed under a *cooperative* learning approach, where teamwork strategies are enacted to complete a task. Clear objectives are defined at the beginning of the activity, specific roles are assigned to team members, and the sum of individual efforts are required to effectively provide care to a manikin patient. On the other hand, the database design activity followed a *collaborative learning* educational approach, where members of a group work share their individual knowledge and experience to create a database design solution. The following sections are dedicated to describing the learning scenario, the data collection sessions, participants and the learning design of both learning contexts.

It is important to highlight that parts of this chapter have already been presented by the author in different publications (see Publications during candidature).

3.1.1 Context 1: Nursing teamwork simulation[5]

Healthcare simulations play an important role in the development of teamwork, critical thinking and clinical skills and prepare nurses for real-world scenarios. Students from the Bachelor of Nursing at the University of Technology Sydney experience many hypothetical scenarios across different stages of their professional development. In these scenarios, students acting as registered nurses (RN) commonly provide care to a patient, who has been diagnosed with a specific condition. Manikins, ranging from newborn to adult, give students the opportunity to practise skills before

[5] Parts of this section have been published in AIED'18 (Echeverria, Martinez-Maldonado, Power, Hayes, & Shum, 2018) and CHI'19 (Echeverria et al., 2019).

implementing them in real life. Simulations are sometimes recorded and played back to students in facilitated debriefing sessions with the aim of reflecting on the skills they need to develop and to identify potential areas for improvement (Green, Stawicki, & Firstenberg, 2018). However, in practice, using the video recordings in authentic classes is challenging due to time and logistic constraints. This suggests that providing actionable and intelligible information can be helpful to support reflection on students' practice.

Two data collection in authentic nursing settings were carried out to address the main research question of this research. *Data Collection 1* was conducted in an immersive simulation room and *Data Collection 2* in a simulated hospital ward-classroom. Both studies followed the framework of UTS curriculum which emphasises patient-centred care (PCC) and teamwork. This means that students must learn not only technical skills, but also develop teamwork skills to enable them to deliver competent, professional care.

3.1.1.1 *Data Collection 1: A controlled simulation scenario*

This data collection was conducted as part of an optional program for nursing students to gain further experience in academic research and healthcare practice, under ethics approval (ETH17-1502) from the UTS Research Ethics Committee.

Participants

Second- and third-year students were invited to participate in an optional simulation, as part of a research program leaded by a Lecturer of the Bachelor of Nursing program. Information about the research project was sent out to students by email. Students who offered to participate in this research received a 20 AUD gift card as an incentive for their participation. In total, nine undergraduate nursing students (6 female and 3 male), aged from 20 to 53 years (avg.: 34 years, std.: 10 years), volunteered to participate in a data collection session. According to their schedule, they were randomly organised into three teams (A, B and C), of four students (2 female and 2 males), three (2 female and 1 male) and two (female) students each.

Physical room configuration

The room was equipped with a control room behind a one-way mirror from which a teacher can control the patient's state and 'voice' via a microphone connected to a speaker located inside the manikin's mouth. Usually, the simulation room is furnished with five patients' beds. Only one bed was prepared for the simulation. Figure 3.3 shows the space near the patient's bed. A monitor, a medicine trolley, a resuscitation trolley, and several tables were used by the students during the simulation.

Learning Activity

A teacher, who was previously appointed as subject coordinator of Integrated Nursing Practice and has taught several simulation courses, designed and adapted a simulation scenario given in the Integrated Nursing Practice subject. The simulation was focused on teamwork and communication in the context of ***caring for a deteriorating patient eventually requiring basic life support***. Prescribed roles and a set of tasks were given by teacher. Four roles for each registered nurse (RN1-4) were determined for this specific scenario, being RN1 the team leader. Table 3.1 lists the tasks to be performed by each RN during the simulation. Also, a high-fidelity patient manikin simulator ("Mr. Lars") was programmed by the teacher to deteriorate over time, dividing the task into two phases. Two researchers, a lab technician and the teacher were present in each session. Three sessions were conducted in the immersive simulation room. The session was planned as follows:

- **Research instruction (5 mins):** Before starting the simulation scenario, the research team explained to students what the research was about. Consequently, students were asked to sign an informed consent form (see Appendix A.2) explaining that all the collected data would be used for research purposes and that no personal information will be shared or used for formal assessment. Then, each student was equipped with a wristband and a location tag.
- **Learning activity explanation (10 mins):** The teacher explained to students the learning objectives and how the scenario will be performed (two phases). Detailed information about

Table 3.1: Tasks to be demonstrated by each Registered Nurse (RN)

Registered Nurse	Tasks
RN1	• Communicate with patient • Conduct pain assessment (PQRST) • Initiate oxygen • Ask someone to get Anginine • Check for danger • Check for response • Send for help (note time) • Clear the airway, head tilt/jaw thrust • Bag the patient
RN2	• Take a set of vital signs • Obtain and administer Anginine (6 rights) • Verbalise precautions of Anginine • Verbalise actions/observations would be appropriate after initial administration • Commence compressions • Verbalise depth and rate of compressions • Count aloud so airway nurse is ready to give breaths
RN3	• Connect 3 lead ECG • Identify rhythm correctly • Complete the MET Form
RN4	• Verbally inform other team members of changing vital signs, improvement or deterioration of the patient • Phone for clinical review using ISBAR • Get resus trolley • Attach AED/follow prompts • Maintain team safety/clear team before shock

the learning objectives and outcomes of this simulation can be found in Appendix A.1. Each student was randomly asked to enact one of four roles (RN1-4) with an associated set of subtasks. RN1 was allocated as the team leader. Depending on the number of students, the subtasks were distributed among the roles. A lanyard was given to each student with a printed list of the actions to be performed. The teacher allocated five minutes to read Mr. Lars handover and to let the team discuss how to provide care to the patient.

- **Simulation Scenario (two phases):** Next, the teacher went to the control room, where both researchers and the technician were located to observe the simulation. The simulation was divided in two phases. **In phase one** the team enacted the roles of registered nurses (RNs) in order to assess and treat the patient, Mr. Lars, who was experiencing chest pain. These RNs, enacting different roles, should ideally communicate with the patient, apply oxygen, assess his pain using PQRST[6] pain assessment method, perform vital sign observations, administer Anginine, connect the patient to a three lead ECG, identify his cardiac rhythm, document appropriately all actions performed by nurses and call to the doctor for a clinical review. **In phase two** students take over Mr Lars's care which is programmed to lose consciousness due to a fatal cardiac rhythm. The team will need to perform basic life support using the DRSABCD protocol[7]. Each session lasted an average of 9.5 minutes (std.=0.7). Phase 1 lasted 5 minutes (std.=0.8) and Phase 2 4.5 minutes (std.=0.4).
- **Debriefing (5 mins):** Once the simulation finished, the teacher conducted a ten-minutes debriefing session. In this debriefing session, the teacher reviewed with students about their performance. For example, the teacher asked students about their role in the team, if they worked together, if the simulation seemed to have helped them to prepare for clinical practice, and if they learnt something from the simulation.

3.1.1.2 *Data Collection 2: A simulated hospital ward-classroom*

The second data collection was conducted in an authentic setting as part of the regular classes of the third-year undergraduate unit, Integrated Nursing Practice. This study also complied with the ethics approval (ETH18-2278) from the UTS Research Ethics Committee.

Participants

Each class typically has 20-30 students working in teams of 4 to 6. Classrooms are equipped with 5-6 patient beds. Only the activity occurring around one bed (either Bed 4 or 5, as displayed in Figure

[6] PQRST stands for: P=Provocation-Palliation; Q=Quality-Quantity; R=Region-Radiation; S=Severity Scale; and T=Timing. Detailed information can be found in https://www.ambulance.qld.gov.au/docs/clinical/cpp/CPP_Pain.pdf

[7] DRSABC stands for: D=Danger; R=Response; S=Send for help; A=Airway; B=Breathing; C=CPR; D=Defibrillation. For more information visit https://www.stjohnnsw.com.au/drsabcd-action-plan/

Figure 3.1: Patient's bed arrangement to simulate a ward hospital scenario in a classroom.

3.1) was tracked and recorded in each class, to allow one team to opt in for the data collection study. Thus, the data collection session involved six teams of 5 students each (teams 1-6), with a total of 30 students (27 female and 2 males) and 2 teachers (A and B). The data collection focused on six 3-hour classes conducted in week 3 of semester 2, 2018 and each teacher taught three classes.

Physical room configuration

The simulation classroom is equipped with a control room behind a one-way mirror from which two researchers were observing and taking notes. Also, it has a trolley and medication shared room located at the back where all students usually get the medicine and implements (e.g. syringes, fluids, ECG machine) to provide care to the patient. In addition, other important spots are located in the simulation room, such as: a) stethoscope zone, b) drugs trolley, c) rubbish, d) recycling zone and e) teacher's desk. Five beds are usually arranged to simulate a ward-like hospital scenario, as depicted in Figure 3.1.

Learning Activity

The coordinator of the unit designed a simulation scenario that involved a patient experiencing an adverse drug reaction to promote the enhancement of teamwork, critical thinking and clinical skills. The learning objectives for this session are listed in Table 3.2 and a full description of the scenario including the manikin setup, patient prompts and the summary of Emergency Department (ED) admission medical notes can be found in Appendix A.3. As listed in Table 3.2, students were expected to perform a set of tasks to provide care to the patient. Additionally, each student was asked (but not required) to play one of 3 possible roles: a team leader, secondary nurses, and the patient. The teacher assumed the doctor role. Two researchers were present in each classroom session while one of two teachers delivered the regular class. The duration of each classroom session was three hours. A total of six sessions were collected for this study. The session was delivered as follows:

- **Learning activity explanation (20 min):** The teacher explained to students the learning objectives, how the scenario will be performed (in teams) and asked students to form teams of five students. After, the teacher read the handover for this particular scenario to all the teams

Table 3.2: Learning objectives and expected tasks for this simulation scenario

Learning objectives
• Primary head to toe assessment (A – G framework) • Respiratory assessment • Pathophysiology of COPD • Oxygen therapy
Tasks for adverse drug reaction scenario
• Hand hygiene • Assess airway is clear • Introduce self to patient • Conduct primary assessment using the A-G framework • Conduct a respiratory assessment • Observe RR and pattern • Listen for bilateral breath sounds • Measure pulse rate and volume • Measure temperature BP and SpO2 • Consider O2 therapy – consider if appropriate delivery device in-situ and flow • Prime IV line and administer IV fluids • Complete Fluid Balance Chart • Review patient clinical charts and medical notes • Consider ECG test • Administer IVAB – Ceftriaxone • Select most appropriate delivery device e.g. rapid infuser or burette set • Document all observations and actions in progress notes

and gave additional instructions to the student who enacted the role of the patient, specifically to give him/her prompts related to the drug reaction.

- **Simulation Scenario (120 mins):** Before starting the scenario, a researcher explained to the voluntary team what data would be collected and that the data will not be used as assessment of the task. Teams provided care to the patient according to the status and prompts given by the manikin and the student acting as the patient. The teacher, who acted as the doctor, was supervising the teams as the activity unfolded. Researchers were in the control room observing the scenario, taking notes and logging the actions using the team observational tool described above. Once the team discovered that the patient suffered an adverse drug reaction, the teacher confirmed this by visiting each team and briefly asking why it occurred and if they had any question. After, teams finished to fill in the documentation charts. Students who voluntarily participated in the data capture signed a consent form (Appendix A.2).

- **Debrief (20 mins):** After the simulation ended, students were asked to gather together to continue the class with a debriefing session of the scenario. The debriefing session allowed students to reflect on team and individual critical thinking and clinical skills. The teacher guided the debriefing by asking questions such as: *how did the team leader exhibit nursing leadership? which aspects of patient care were performed well by the nurses? what did you notice about the patient's behaviour and communication?*

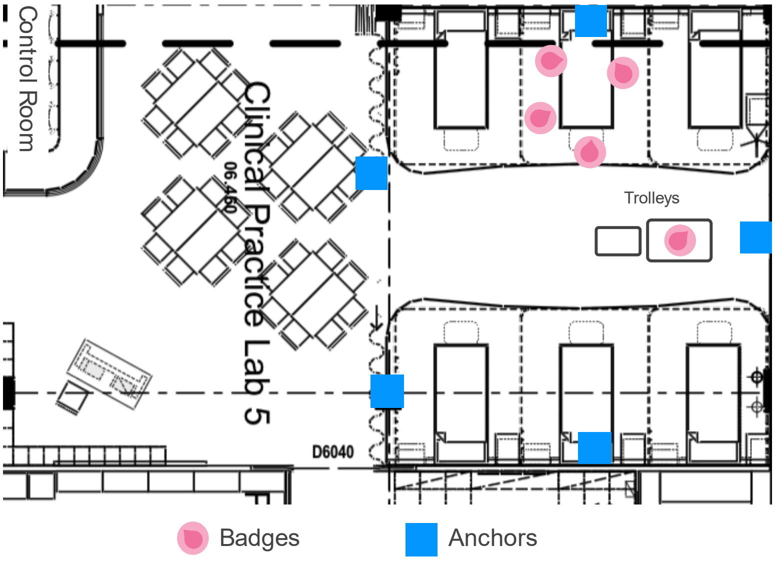


Figure 3.2: Floor plan and distribution of anchors (blue squares) to track badges (pink circles). Five badges were used for Data Collection 1.

- **Closing (20 mins):** Finally, the teacher concluded the session with final remarks and recommendations and presented the agenda for the next session.

3.1.1.3 *Sensors and multimodal data*

The following sensors and equipment were used to capture evidence of collaboration in the physical environment and student's interactions.

- **Indoor localisation system:** People localisation around the manikin was captured automatically through an ultra-wideband (UWB) indoor localisation system[8]. This system is composed of anchors that are used to triangulate the position (as x and y coordinates) of wireless devices in indoor spaces, which are mounted into the walls, and several wearable tags or badges that can be worn by people or be attached to objects to detect movement. The system has an accuracy of 10 cm, which is sufficiently accurate for the simulation activity, and superior to Bluetooth Beacons (accuracy: 1-2 m) or Wi-Fi (accuracy: 3-5 m) solutions (for a detailed review of these technologies see Farid, Nordin, & Ismail, 2013). A Python script run in a MacBook Pro (2015) to configure and calibrate the system; and record the position of every badge. The script saved in a .JSON file the x and y position of each badge identified by an *id* at 2Hz, and the *timestamp* for each record.

  For *Data collection 1*, five anchors were mounted on the walls to track nursing students and objects in the simulation room, one badge was located in the resus trolley and one badge was

[8] Pozyx.io

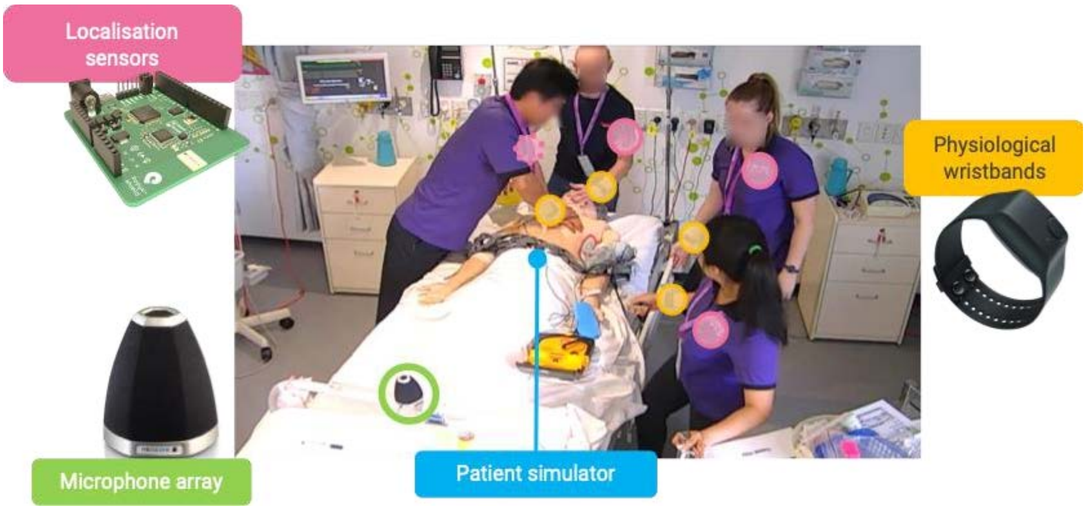


Figure 3.3: Multimodal data collection in a controlled simulation scenario (Data collection 1)

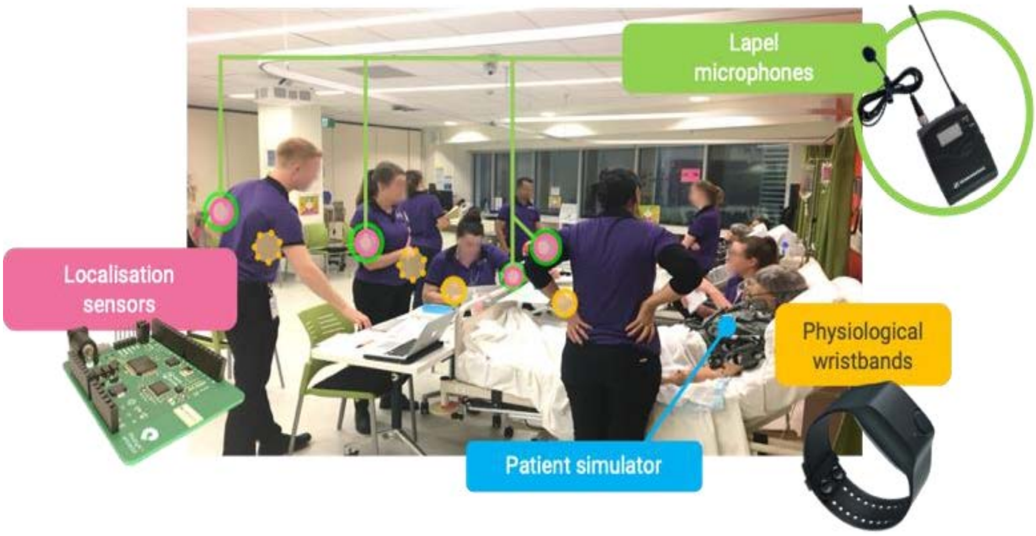


Figure 3.4: Multimodal data collection from a nursing simulation in a classroom scenario (Data collection 2)

located inside an armband worn by each student, as depicted in Figure 3.2. The student enacting the patient was not tracked as it was commonly sitting on a chair at the bed side. For *Data collection 2*, seven anchors were mounted on the walls to track nursing students and the teacher during the simulation and badges were located inside an armband worn by each student. For example, Figure 3.2 illustrates the distribution of the anchors across the simulation room (blue squares) and students and objects (i.e. trolley) being tracked by the system represented as pink dots.

- **Microphones:** For *Data collection 2*, individual lapel microphones (Figure 3.4 – lapel microphones) were located inside an armband worn by each student and the teacher. Four .WAV files were saved at the end of each session, one per student.

- **Patient Simulator:** Some student and patient's actions and status were logged by a patient simulator. For *Data collection 1*, a high-fidelity Laerdal SimMan 3G[9] manikin (Figure 3.3 – patient simulator) was configured and used through the simulation to automatically capture actions such as: placing the oxygen mask, setting oxygen level, attaching blood pressure monitor, reading blood pressure, administering medicine, attaching the ECG device, starting CPR, and stopping CPR. Proprietary Laerdal Software, installed in a CPU located in the control room, allowed to save the actions and the time when they occurred in a .TXT file. For *Data collection 2,* limited actions of the patient were recorded (e.g. checking vital signs) using a Nursing Anne manikin simulator (Figure 3.4 – patient simulator) operated via the Simpad handheld device.

- **Physiological Wristbands:** Physiological wristbands (Figure 3.3 and Figure 3.4 - Physiological wristband Empatica E4[10]) were worn by students. Each wristband includes a photo plethysmography (PPG) sensor to measure Heart Rate (HR) continuously, an electrodermal activity (EDA) sensor to measure skin conductance, a 3-axis accelerometer to detect movement and activity, and an optical thermometer to sense activity.

  For both data collections, each student worn a physiological wristband to capture EDA (at 4Hz) and acceleration (at 32 Hz) streams of data. Data was captured internally on each device and manually downloaded using Empatica Connect software. For each wristband two .CSV files were saved: An EDA.csv file containing the timestamp when the Empatica started to capture data and EDA values; and an ACC.csv file with x, y and z accelerometer values.

- **Video camera system:** All the sessions were video and audio-recorded using three fixed synchronised cameras and several microphones located at the ceiling of the room and managed by B-Line medical software. B-Line medical software was installed in a CPU located in the control room and operated by the technician. Figure 3.5 shows a sample frame captured by the three cameras for Data collection 1.

Synchronisation

Multimodal data collected in authentic settings has to cope with restrictions such as data

Figure 3.5: The camera system located at the ceiling of the room composed of three fixed cameras.

availability (Martinez-Maldonado, Kay, et al., 2019). This means that traces of collaboration gathered from off-the-shelf applications or commercial software are not always ready for data exploitation or interpretability. For instance, the indoor localisation system was connected through a wireless connection in a personal computer, wristbands were not connected to a dedicated server, Laerdal and B-Line medical software were installed in different CPU's with a wired internet connection and only the technician had access to them. Thus, it is worth noting that the diversity of device configurations and software made the synchronisation process a challenging task. Therefore, a web-application tool was developed to manually synchronise each data stream by saving start and end timestamps from all devices before the data collection began, as detailed in the next section. Thus, when a session started, the researcher simultaneously pressed the recording button in a device and clicked a button on the web application to indicate that the device started to record the data. The same process was followed for stopping the data recording. Once that all data was recorded, in a post-hoc task, the researcher saved all files into one computer and a Python script merged all .CSV and .TXT files. Due to the different sampling rates each device had; all data was down sampled to 1 Hz.

Team observation tool

A web-application tool was developed for two main purposes: 1) synchronise and log all critical actions defined by the teacher; and 2) easily visualise the actions performed by a team. The tool was developed using node.js as a back-end server, and angular.js as front-end framework[11]. All data logged from the observational tool was saved in a database and can be exported to a .CSV or .XLS file. In order to manage, synchronise and log actions observed during a simulation scenario, the web-application tool enables an observer to:

- **Manage the sessions being captured, allowing the user to input the place where the session is being recorded and a name.** Internally, a sequential id is created and saved for each session, along with the timestamp, location and name (Figure 3.6 - A). The observer can choose to log actions; add sources and objects; manage the capture; and visualise a timeline of actions. For example, Figure 3.6 – A illustrates five sessions that were logged with the tool, along with a set of options for the observer to manage that session, such as log actions, sources, objects and so on.

- **Allocate the devices that will be used in the data collection, by letting the user select which sensors or devices will be used during the data capture.** The tool allows to associate each session with a set of devices. As can been seen in Figure 3.6 - B, a list of data

[11] Code available at: https://github.com/vanechev/obs-tool

- sources is displayed in the screen along with an *Add* button. If the observer clicks on the *Add* button of a specific data source (e.g. pozyx), the tool will associate the selected data source with the session, and it will show at the bottom of the screen an updated list of the data sources. As

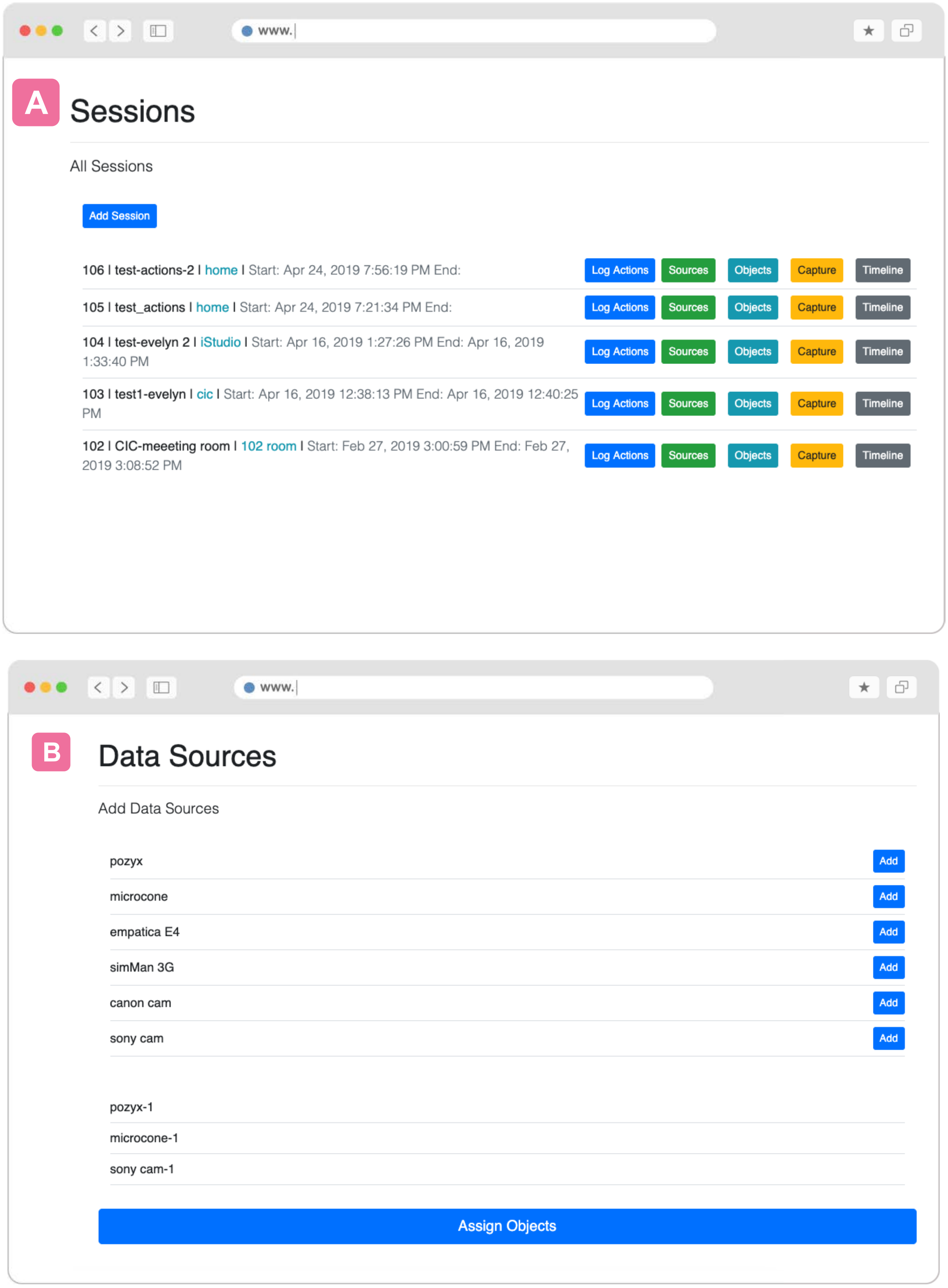


Figure 3.6: The team observation tool. A) List of all the sessions managed by the tool; B) a screen showing the data sources added to a session.

shown in Figure 3.6 - B, three data sources has been added to this session: pozyx-1, microcone-1, and sony-cam1. Also, the tool allows to add several data sources for one device. For example, if the observer requires to track three objects with the same indoor localisation system 'pozyx', she will have to add the data source three times, and internally the tool will increase the number of the data sources (e.g. pozyx-1, pozyx-2, etc).

- **Assign objects (manikin, trolley) and people (e.g. patient, registered nurse) to be automatically tracked or observed.** The tool allows to associate the objects or people that will be tracked, either by a sensor (e.g. pozyx, Empatica) or observed by a human. Figure 3.7 – C depicts the *Objects* screen, where a list of objects and roles is shown with its corresponding *Add* button. Similar to the previous functionality of adding data sources, the tool will add an object to the list of trackable and observed objects and it will display an updated list of the objects. For example, the observer may want to track five registered nurses (RN) and one trolley; and observe one teacher's actions. Therefore, the observer should add five RN (RN1-5), one teacher (teacher-1) and one trolley (trolley-1). In addition, if the object is the same, the tool internally will increment its number. In addition, if the object to be tracked is a nurse or a teacher, it will display two lists of ids corresponding to the badge id (for pozyx, e.g. id: 26886) or the wristband serial (for Empatica E4, e.g. id: A010C7). Figure 3.7 – C displays a screen showing how the observer added six RNs (RN1-RN6) and allocated the localisation badge id to identify the badge worn by each registered nurse.
- **Synchronise the data coming from different devices by saving the timestamp when each device starts its recording, using a centralised-sever to get timestamps.** As explained in the previous section, to synchronise devices or equipment that are beyond the control of the researcher (e.g. Laerdal software) or cannot be connected to a computer (e.g. lapel microphones), the observer should simultaneously press the record button in the device and the start button associated with that specific device in the tool. Likewise, the observer should perform the same actions to stop recording. This will save an entry in the database associating the session, the device being synchronised and the timestamp when the device *started* and *stopped* to record. These multiple timestamps can be used for post-processing analysis to synchronise all data streams.
- **Log actions by allowing the user to select an action-button from a list of buttons and the person or object who is performing the action.** The tool allows the observer to add predefined actions related to the learning design of the activity and associate them with any object of people being observed. The tool will save the timestamp of the action being registered, the name of the action, the person or object, and the session where this action was performed. Figure 3.7 – D displays a list of actions that were previously defined by a teacher. Each action (e.g Anginine) is presented along with the people being observed (e.g. patient1, student1,

student2, teacher1). When the observer is ready to log an action, she should refer to the specific action and click on the name of the person that performed the action. Then, the action will

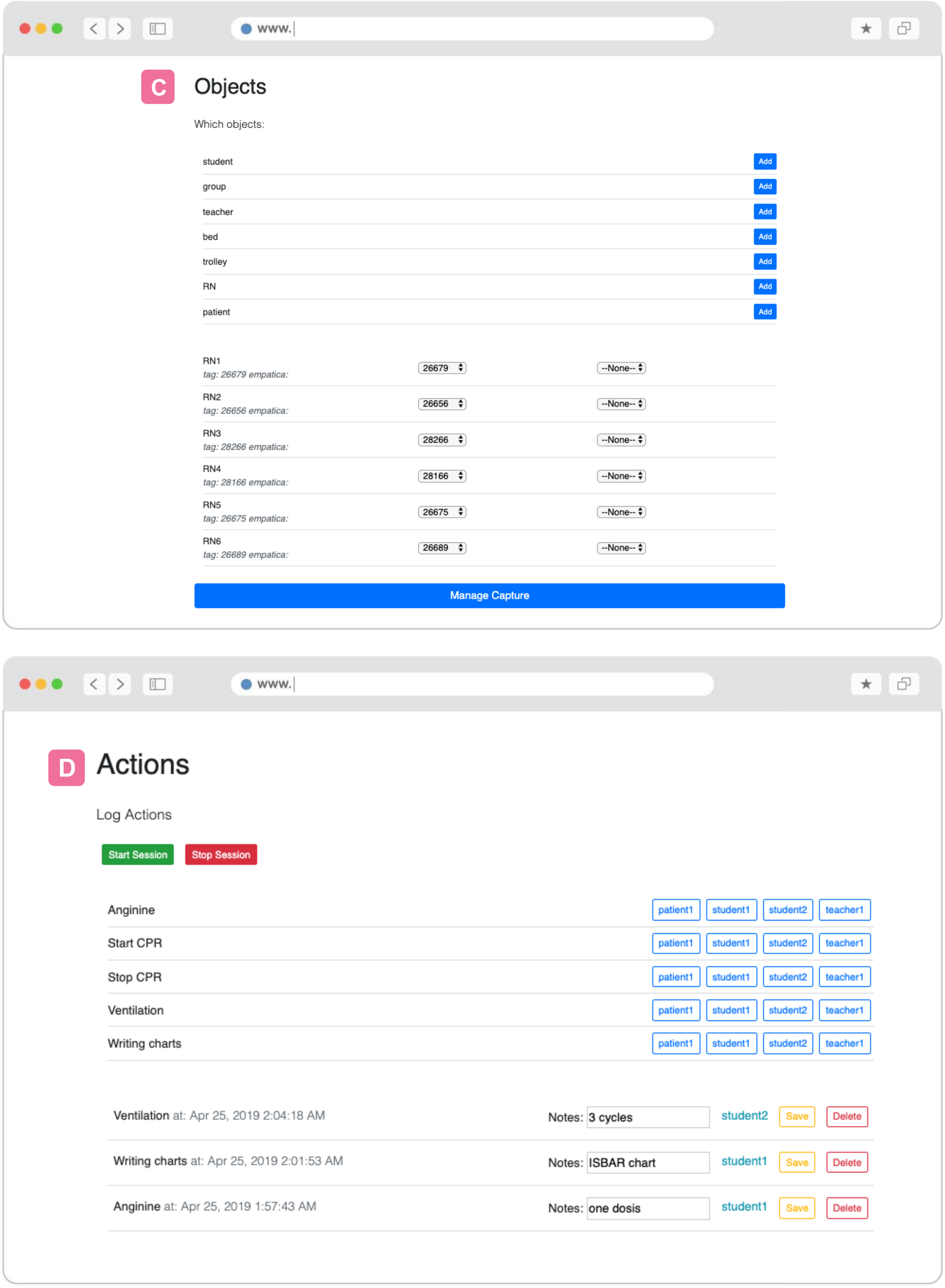


Figure 3.7: The team observational tool. C) a screen illustrating the objects added to a session and its association with a tag or a physiological wristband; and D) the screen where one or multiple observers can log actions performed by students.

appear at the bottom of the screen in an updated list with all the actions that have been recorded until that moment. The observer also has the option to write down and save a note related to an action or delete that action.

3.1.2 Context 2: Database design activity

In a Computer Science (CS) program, regular activities in different courses include topics around the design of database systems through real-life case studies, using an Entity-Relationship diagram as a tool to conceptualise and abstract real-world concepts. Database System Design (DSD) activities are performed as a collaborative activity to promote and support the development of collaborative, communication and critical thinking skills. DSD activities demands designers to present ideas to peers and actively participate and contribute to a final design solution, while working together as an effective group. Usually groups are conformed from three up to five students. In a classroom environment, teachers should simultaneously observe and provide feedback from 5 to 8 groups, while the activity unfolds. In addition, after the activity has finished, students expect to receive feedback from the teacher about the final solution and their performance. As can be noticed, it is challenging for the teacher to provide personalised feedback to groups.

3.1.2.1 *Data collection with students working in a database design activity*

This section describes the data collected in a collaborative environment using a multimodal tabletop system, DBCollab tool, to capture the database design activity. It also describes the participants that were part of the data collection and the task that was designed for this particular purpose.

Participants

Fifteen second- and third-year undergraduate students (2 females), ranging from 21 to 26 years old (average: 21.96 years, stdv.: 2.35 years), enrolled in an introductory Database Systems course (second academic semester – 2017) from ESPOL University, were asked to use the DBCollab tool as part of their regular classes. Ten students were registered in the Computer Science program, and five were from the Telematics Engineer program. Students conformed groups of three members based on their own affinity, ending up with five groups in total.

Learning Activity

After attending a lecture about designing Entity-Relationship models, students were invited to participate in a tutoring session to use the DBCollab tool. The teacher (PhD student), set two sessions (Sessions 1 and 2) to use the DBCollab tool with students and for the research team to gather evidence about their reactions after receiving automated feedback. Students were asked to sign an informed consent form explaining that all the collected data would be used for research purposes and that no personal information will be shared or used for formal assessment. The duration of each classroom session was two and a half hours. One DBCollab system was used; thus, groups were allocated 30-minutes time slots to perform the design activity. The purpose of the session 1 was to let students to familiarise themselves with the tool, so no automated feedback was provided in Session 1. In Session 2, automatically generated feedback was provided to the groups. The data reported here, and that will be used in further chapters, is from Session 2, in which groups performed the following:

- **Collaborative design activity (20 min):** each group was asked to solve the same database design problem using DBCollab. At the end, they presented a final design of the database to the teacher.
- **Feedback (5 min):** each group was asked to navigate through the dashboard and explore all generated feedback (charts from Figure 3.11 and Figure 3.12)
- **Short interview (3-5 min):** a researcher asked students questions related to the automated feedback. Some of the questions were: Why do you think you got that grade for the final design? Do you think this feedback could help you to reflect on the group-work? Do you think this feedback is useful to help you to reflect on the activity performed?
- **Questionnaire about perceptions of the feedback:** Each participant filled out a 5-point Likert-scale questionnaire with six questions: Q1) potential of feedback for both aspects; Q2) usefulness of feedback for both aspects; Q3) usefulness of feedback about social aspects; Q4) usefulness of feedback about epistemic aspects; Q5) usefulness of epistemic process feedback; and Q6) validity of the automated grade.
- **Writing reflection (post-hoc activity):** Three days after the classroom collaborative design activity, each student received all feedback results (i.e. dashboard) by email and each group member was asked to write a reflective text about the collaborative activity, first as an individual reflection and then a shared group reflection.

3.1.2.2 *DBCollab Tool*[12]

A multitouch tabletop system was developed. The system captures multimodal data and automatically generate visualisations that can potentially guide teachers to provide a personalised feedback and support student reflection once the group has finished a database design activity. Figure 3.8 illustrates some students using the tool for creating a database design model (left) and some students discussing about the feedback they received after finishing the modelling activity (right).

[12] Parts of this section have been published in ICCE 2017 (Echeverria, Martinez-Maldonado, Chiluiza, et al., 2017).

The design of the DBCollab tool is based on previous work that focused on defining a set of design features for a multi-touch tabletop system to support argumentation (Falcones, Wong-Villacres, Barzola, & Garcia, 2016) and database design (Wong-Villacres, Chiluiza, Ortiz, & Echeverria, 2015). As a result, the following features were presented in the implementation of the DBCollab tool:

Figure 3.8: The DBCollab tool. Left: three students interacting at the multi-display, multi-touch environment that facilitates the collaborative design of database diagrams. Right: a set of visualisations automatically presented to groups just after completing their task.

- **Structuring the task in sub-tasks:** To help group members clearly understand the design problem the task is divided into six stages: I) individual reading of a short description of the case study; II) group reading and discussion about the short description; III) the actual group design task, where design elements can be added, edited and deleted; IV) group discussion about the final design (looking for errors and small changes, without making any change to the solution, V) adjustments and agreement on the final design, and VI) evaluation and reflection time via the automated feedback dashboard.
- **Providing clue-based instructions:** Information about the case study is presented as small chunks of text to students. During stage III, group members can ask for a "clue" and read it individually or share it with others. This enables group member's interdependence as well as awareness of other's actions.
- **Providing a shared view and highlighting individual contributions:** During stage I, group members can highlight keywords from the short description individually (each member is identified with a colour), to later share these highlighted keywords in a shared screen (stage II). This supports awareness of other's contributions.
- **Providing a collaborative puzzle-like interface:** Design elements (e.g. entities, attributes) are represented as cards to provide flexibility in the design process. Group members can interact with these cards by creating, editing and deleting them during stage III and/or stage V. In addition, each card is identified by a colour, representing a member's contribution to the design.

DBCollab allows the collaborative co-creation of database designs (diagrammatic entity-relationship schema) for up to three students. The software of the DBCollab tool is composed of three applications: a) the tablet interface (see Figure 3.9 – left view); b) the tabletop interface (see Figure 3.9 – right view); and c) the dashboard application (see Figure 3.11). The **tablet application** lets individual members to create, edit cards and send them to the tabletop application through a server. The **tabletop application** allows students to move, link cards and identify group member's actions. All elements created by a student are shown with a different colour (yellow, blue and red). For stage I a snapshot of the tablet device used by group member 1 (red) is observed in Figure 3.9 - left. The student is highlighting words that she considered important from the short description of the case study. During stage III she can *ask for a clue* from three existing clues for this case study (see green button). She can also share the selected clue with group members 2 and 3 (blue and yellow, at the bottom of Figure 3.9, left) by tapping on the grey button *Share clue*. Figure 3.9 (right) shows the tabletop user interface for a group working in stage III. Group members can create and link cards, corresponding to entities and attributes in their data schema. Each card is identified with a different colour per group member (red, blue and yellow).

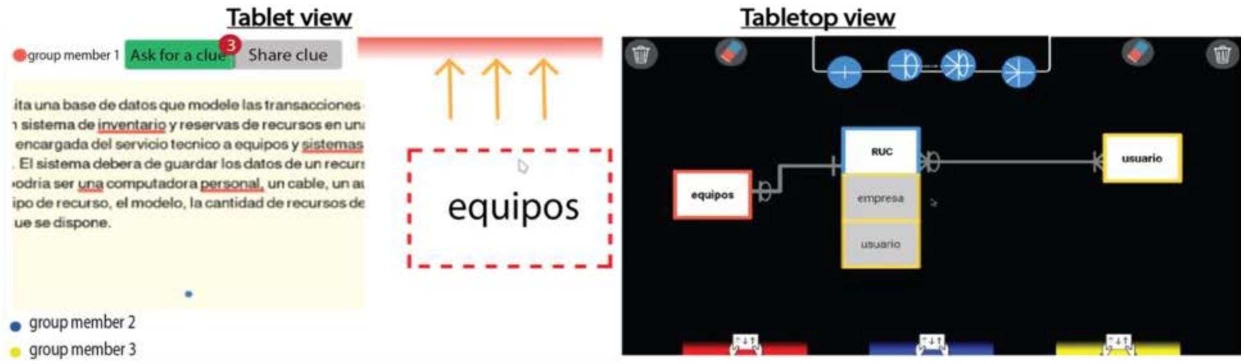


Figure 3.9: The DBCollab design tool. Left: the tablet interfaces used in stage I. Right: the tabletop interface mainly used in stage III.

After the design activity, the **dashboard application** automatically generates both, epistemic (see Figure 3.11, sections A, B, C, D, E) and social visual analytics elements (Figure 3.11, sections F, G, H, I) of collaboration.

Sensors and Multimodal Data

The following sensors and equipment were used to capture evidence of collaboration in the physical environment and student's interactions. DBCollab tool uses interactive surfaces to capture user's input (e.g. writing, moving elements) and sensors to capture speech activity and tracking user's position:

- **Shared interactive surface:** A 60-inch Ideum tabletop (see Figure 3.10, IS1) was used to support simultaneous users' input by touching on the elements displayed at the interface. All actions (e.g. create, move, delete) are timestamped and automatically logged in a .CSV file.

- **Individual tablet devices:** Each student is provided with a tablet device (see Figure 3.10, IS2) to mimic the personal space and create or edit information. Through this personal device, students can add new cards by sending them to the shared view at the tabletop; and ask other students to share or read information. The actions performed in this individual device are also logged in a CSV file.
- **Array microphone to identify individual speech activity:** The microphone array built in a Microsoft Kinect sensor – V.2 is used to identify and record student speech activity around the tabletop (Figure 3.10, S2). The individual speech activity is estimated by using the position of the student around the tabletop, assuming that students do not change their position during the activity. The speech and estimated speaking time by each student are automatically recorded and saved in a WAV and CSV file, respectively.
- **Differentiation of users' touch:** As the tabletop does not differentiate the user touching the interface, a Microsoft Kinect sensor -V.1 (Figure 3.10, S1) is used to track student's hands regarding their position around the tabletop, identifying who touches the table and thus inferring the actions performed by each student (Martinez-Maldonado, Collins, Kay, & Yacef, 2011). The identification of the user is added to the previous CSV file generated with the actions logged by the interactive tabletop.

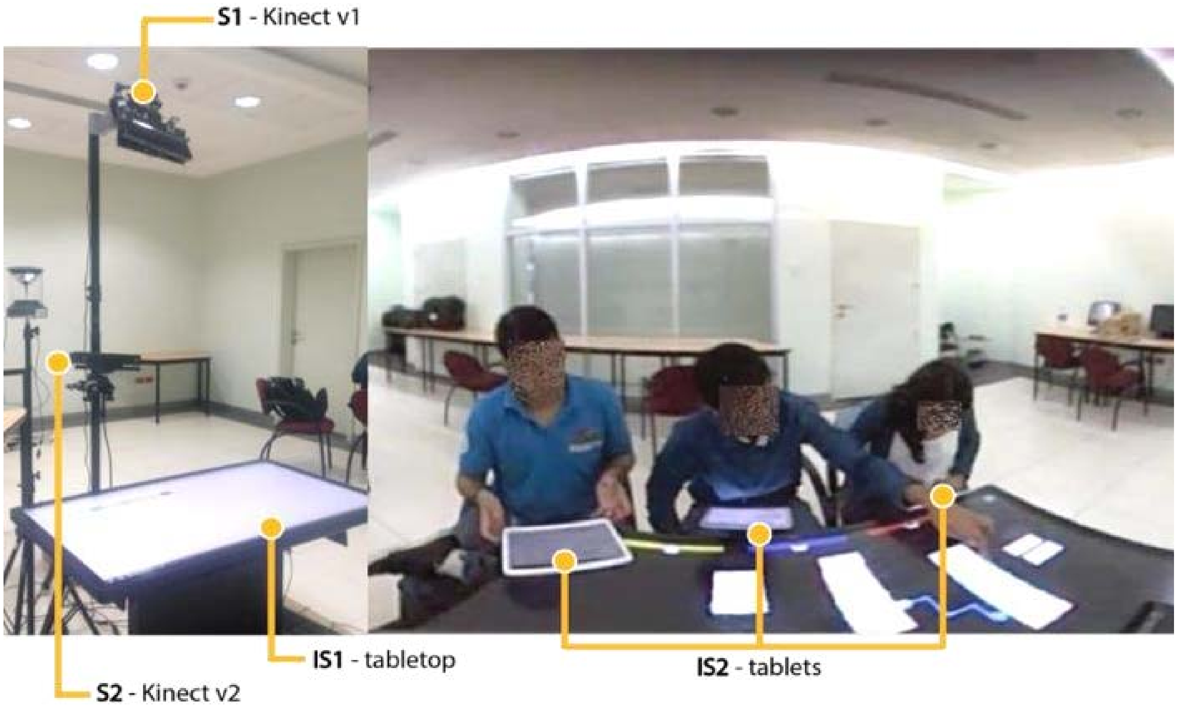


Figure 3.10: Interactive surfaces and sensors used for the implementation of the DBCollab tool.

Visualisations

DBCollab tool automatically generates a student-facing dashboard, to support student's reflection on collaboration and design practice. The dashboard is composed of the following elements.

A. **teacher's solution:** This element shows information obtained from the ideal solution proposed by the teacher, with the purpose of informing groups about the expected goal to be achieved.

B. **group's solution:** This element shows the outcome of the group. In this way, students can compare both, the teacher's and group's solutions. This information is aimed at encouraging dialogue between teacher and group members to discuss discrepancies.

C. **replay:** This feature of the interface allows students to replay the partial design solutions from the beginning to the end of the activity. This feature is aimed at supporting awareness about the epistemic process (e.g. showing how the group approached the task and build their final design, step by step). With this information, group members can reflect on how they were going and if the used strategy was useful to achieve the expected goal.

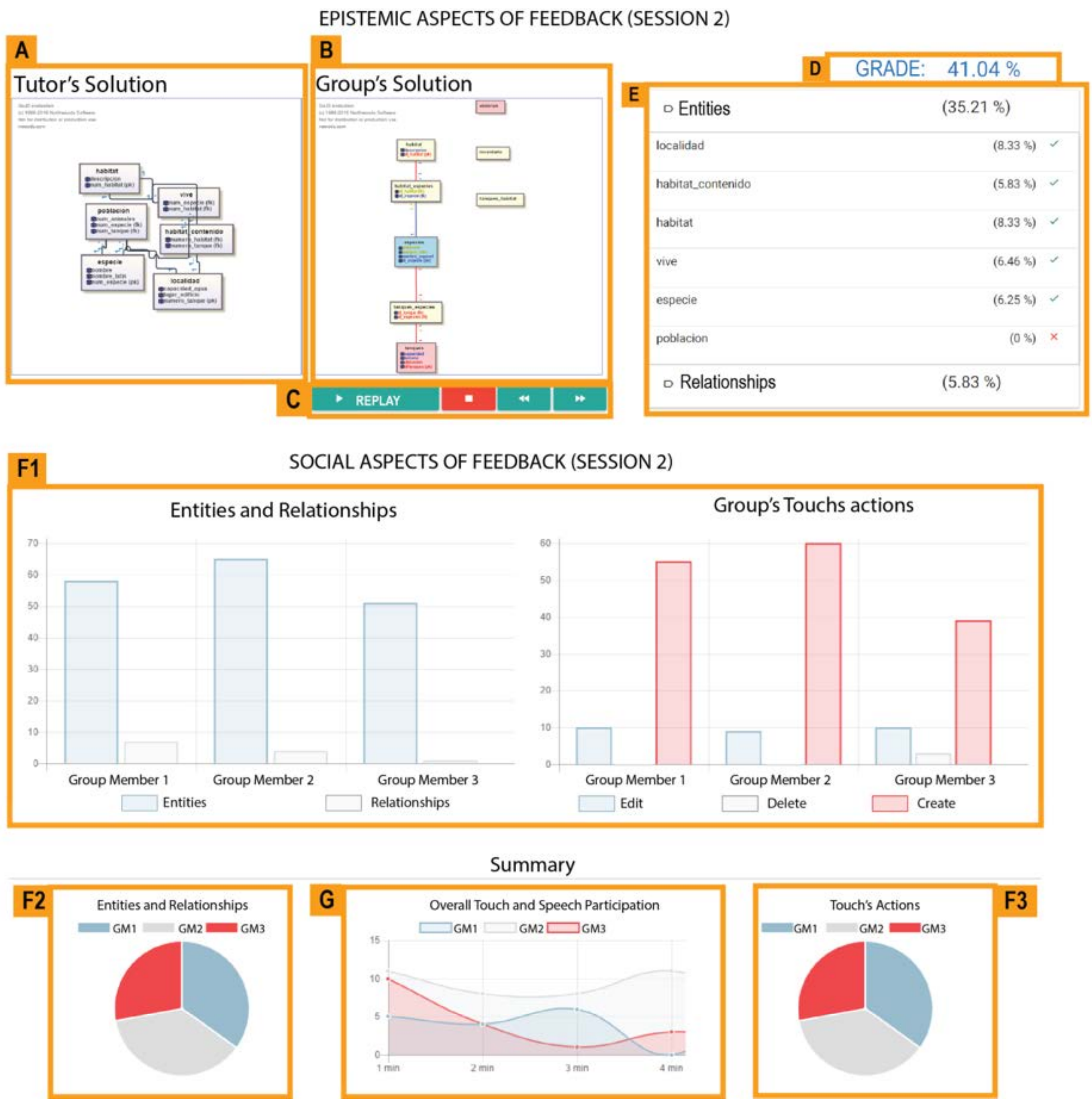


Figure 3.11: The DBCollab feedback tool. Information generated automatically from student's interactions, grouped into epistemic and social aspects.

D. **automatically generated grade:** This information was calculated from the degree of similarity between the teacher's (A) and the group's solution (B). Providing this information to the group can potentially help them to explicitly evaluate their outcome according to teacher's expected goal.

E. **correct and incorrect entities and relationships:** Also, by comparing both solutions (A) and (B), allowed to present correct and incorrect entities and relationships in a detailed list. With this information, group members can reflect on possible misconceptions and mistakes.

F. **entities/relationships and group touch's actions:** (Related to F1, F2 and F3) This information was obtained from touch inputs by counting each time a group member added database elements i.e. attributes and relationships (F1- left, F2) and, counting each time a group member performed a touch action over an element i.e. create, delete and edit (F1-right, F3). Showing this information to the group could provoke reflection about participation at an individual level in the context of other group member's actions.

G. **overall touch and speech participation:** Overall participation was obtained from speech and touch inputs, mapped onto a timeline (accumulated participations vs. seconds). Ideally, showing this information to groups can provoke reflection about periods of high and low interaction. Also, this information can be used to help group members reflect on group interactions and how they should improve participation for a further activity.

The exploration on the effects of providing timely feedback through these visualisations and student's perceptions were reported in (Echeverria, Falcones, Castells, Martinez-Maldonado, & Chiluiza, 2017). In a follow-up study, students reported to prefer a holistic assessment of the collaborative activity (Echeverria, Martinez-Maldonado, Chiluiza, et al., 2017). If the ultimately aim of these visualisations is to support the development of collaboration skills, more specific and temporal information about student's interactions and activity progress towards a correct solution is needed, in terms of social and epistemic aspects. Therefore, another set of visualisations were automatically generated to fulfil these expectations:

H. **participation of each student through the activity:** This chart shows the participation in terms of student's actions with the interactive tabletop and personal tablets per minute (Figure 3.12, H). Each student is represented by a different colour (e.g. red, blue and yellow). The number of actions is given by the counting any action (e.g. create, move, edit, delete any object) per minute. Using this temporal visualisation, the teacher could observe if students were participating equally through the whole activity.

I. **performance of the group through the activity:** This trending line chart illustrates the evolution of the score solution proposed by the group, compared with the teacher's solution (Figure 3.12, I). Entities and relationships objects are represented with green dots and relationships with orange dots. Also, the size of each dot is given by the number of objects that were created per minute. In addition, each data point is labelled with the partial score, calculated according to the comparison of the objects added with the teacher's solution at that

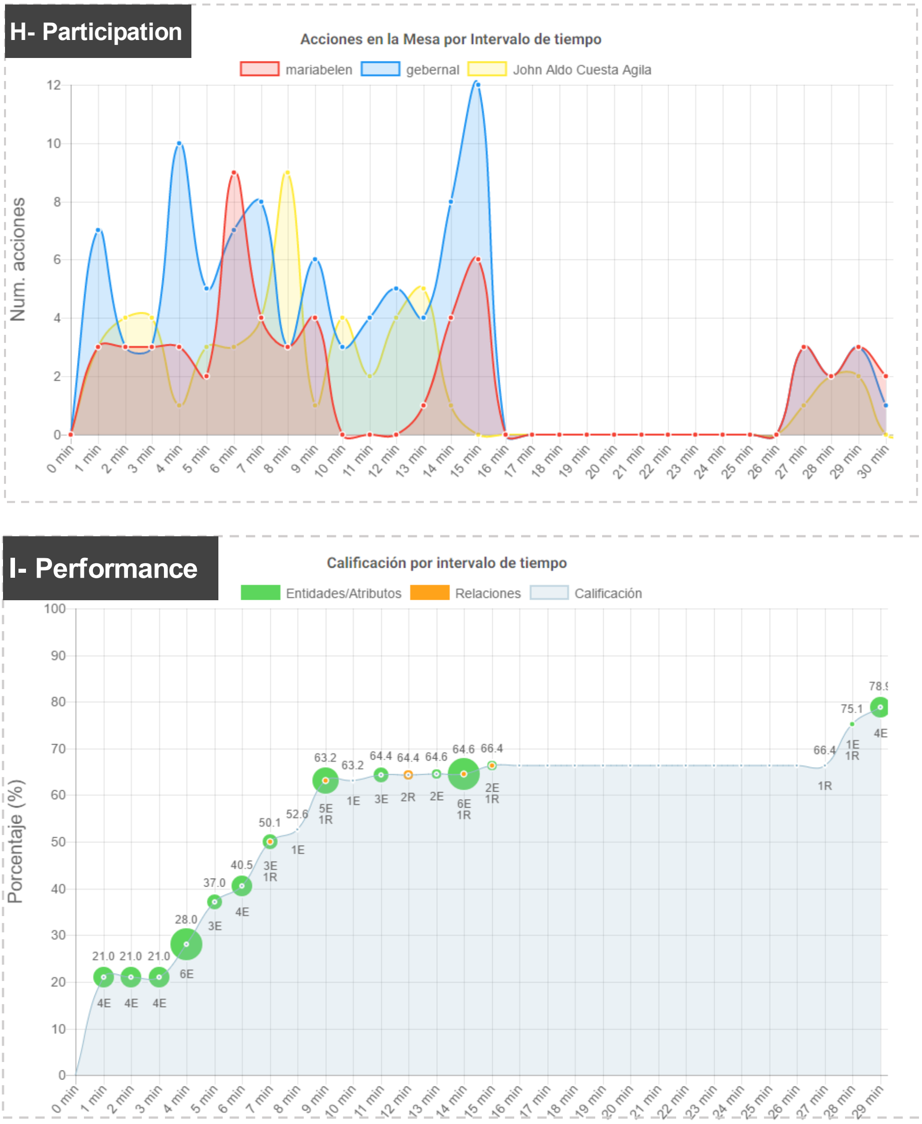


Figure 3.12: A second set of visualisations generated by DBCollab feedback tool. (H) a visualisation showing the participation of each student per minute. (I) a trending line chart with green dots (entities/attributes), orange dots (relationships) and partial scores of the solution per minute.

time. This temporal visualisation also allows the teacher and students to see how the group worked together to find entities and relationships of the solution, and how these contributions affected the score.

For the purpose of this research, the visualisations generated by the five groups will be used in Chapter 6:. A summary of the data and visualizations used across different chapters can be found in Figure 3.14.

3.2 METHODOLOGY

Design-Based Research (DBR) methodology places particular value on the development and validation of innovative, but practical, solutions to real problems by bridging theory and practice (Reeves, 2000). DBR research is suitable for educational contexts because it focuses on the content and pedagogy rather than technology. This means that DBR incorporates theory-based assumptions of learning and practical educational issues into the development and testing of solutions in real classrooms contexts towards the generation of new insights (Reeves, 2005). DBR emphasises the collaboration with different stakeholders and attending to the complexity, authenticity and constraints of real-scenario problems in educational contexts (Cotton, Lockyer, & Brickell, 2009; Reeves, 2000). The DBR process is organised in research cycles and consisted in four phases:

- Phase 1 consists of the **analysis of a practical problem**. Collaboration among researchers and practitioners helps to comprehensively explore and define the problem (Herrington, Reeves, & Oliver, 2009).
- Phase 2 involves the **development of solutions** informed by theory, existing design principles and technological solutions. By analysing current approaches and similar technological solutions from literature, researchers and stakeholders can develop solutions to address the educational problem.
- Phase 3 **implements and evaluates the solution**. Stakeholders and researchers work closely to design the intervention, which is then implemented by the teacher.
- Phase 4 provides opportunities to **document and reflect** on the whole process to inform the redesign or refinement of the intervention and solution, aiming to be implemented in another research cycle (Barab & Squire, 2004).

Depending on the goal of the study, these phases can be repeated several times. DBR contributes to the creation of design principles in a particular context (Edelson, 2002). These design principles do not strive to be generalisable for all cases. Instead, parts of the design principles could be adapted and used in other contexts by researchers and stakeholders. A practical outcome of DBR are products, tools, materials and similar outcomes developed through the phases and the whole process.

Aligned with the main research question of this thesis: **How can explanatory interfaces be designed and used for guiding reflection using multimodal learning analytics evidence?** and with the three sub-questions proposed in Section 1.2, this thesis adopted a DBR methodology to design and implement explanatory interfaces that use multimodal learning analytics evidence. Figure 3.13 depicts the four *cycles* of the DBR process that were followed through this research. It also presents the different prototypes (pink and orange boxes) that were designed and developed through the four cycles (as reported in Chapter 4:, Chapter 5:, Chapter 6:) to produce a final *Capstone Prototype:* an explanatory multimodal interface (described in Chapter 7:). Both quantitative

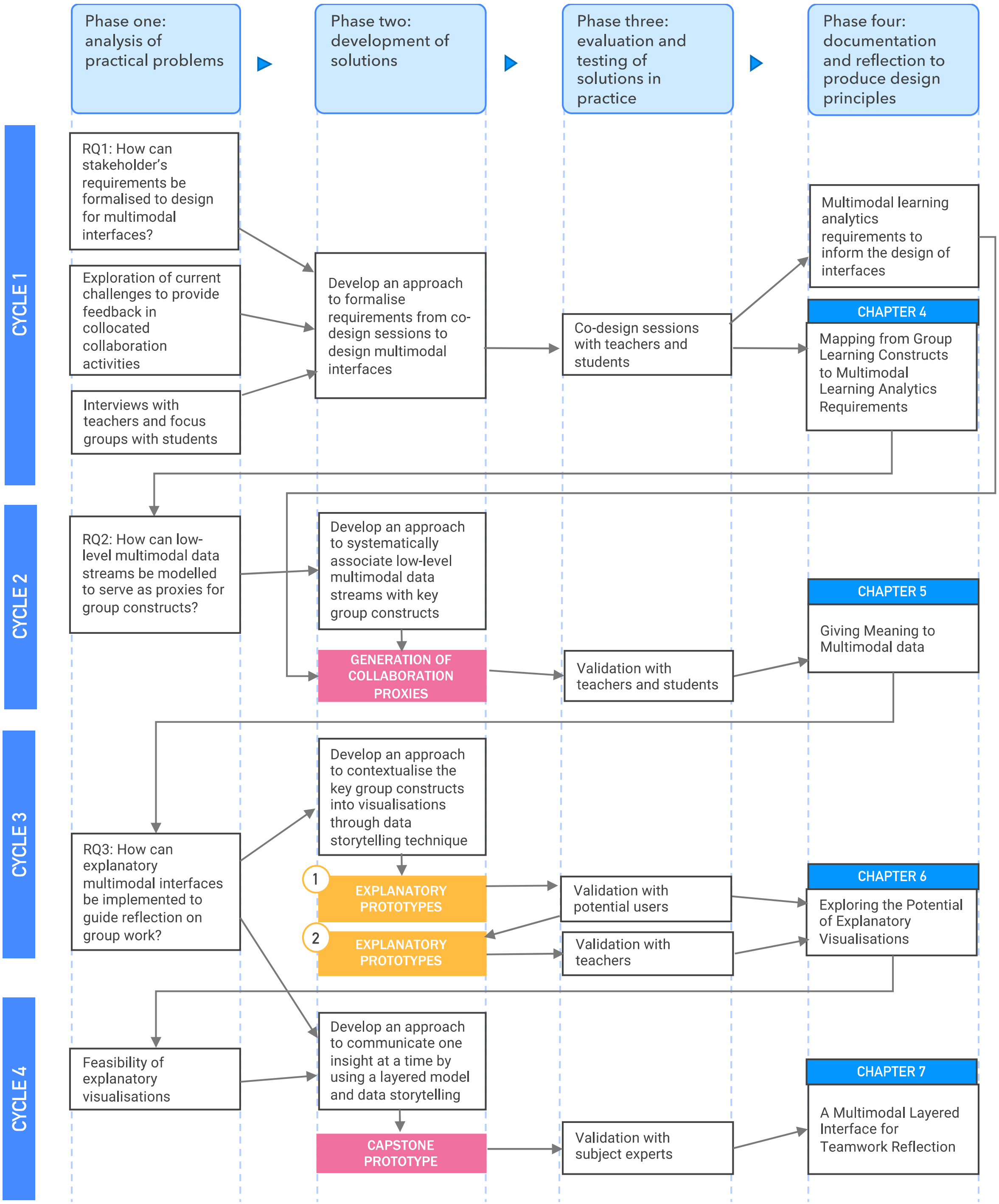


Figure 3.13: Outline of the thesis research design following Reeve's (2000) iterative cycle process.

and qualitative research methods were used to validate the design of the final *Capstone Prototype.* Table 3.3 provides an overview of the research goals, learning contexts, prototypes generated, participants, methods and analysis that were carried out to address the three research questions of this thesis. The following sections describes each iterative cycle in more detail.

3.2.1 Cycle 1

In **phase one**, the first step was to explore current challenges to provide feedback in collocated collaboration activities (Figure 3.13, Cycle 1). The literature review presented in Chapter 2 provided a well-grounded understanding of these challenges and described critical gaps in current LA and CSCL/CSCW work. In addition, interviews with teachers and focus groups with students allowed a deeper exploration of the challenges in the context of nursing simulations (context 1). Researchers actively participated in the activities related to the exploration of the problem with teachers and students. Problems, challenges and potential solutions were derived from the transcriptions of these sessions.

Informed by teachers' and students' perspectives, and the existing literature, **phase two** aimed to address the first goal of this research by developing an approach to formalise requirements from co-design sessions to design multimodal interfaces. Grounded in Human-Centred Design and Learning Analytics methods, this approach consisted of mapping co-design session outputs to the technical, analytical and visual requirements needed to design a multimodal interface.

Phase three validated the developed approach by illustrating its applicability from co-design sessions with 2 teachers and 13 students from the Bachelor of Nursing program at UTS. Co-design sessions included semi-structured interviews, prototyping, fabulation and learning journeys, mapping and voting activities. As listed in Table 3.3 (RQ1, Cycle 1), an inductive content analysis process and a deductive theory-driven analysis was applied to the outputs of co-design sessions. Resulting from this, a mapping template summarising the requirements was developed.

Finally, **phase four** of DBR documented the whole process, which relates to the Contribution 1 of this thesis: A Human-Centred Design Multimodal Learning Analytics (HCD-MMLA) approach and a list of requirements to design multimodal interfaces to support group work reflection. This is reported in Chapter 4:.

3.2.2 Cycle 2

Informed by the reflections reported at the end of cycle 1, the **phase one** of cycle 2 (Figure 3.13, Cycle 2) addressed the second research question by analysing the challenge of combining qualitative insights and quantitative data (RQ2).

Phase two was aimed at developing an approach to systematically associate low-level multimodal data streams with key constructs. This approach was illustrated in the context of nursing simulations (context 1), resulting in a ***set of prototypes*** (Figure 3.13, Cycle 2 - pink box). The design of prototypes was motivated from the requirements documented in the previous cycle. Thus, a set of collaboration proxies were designed and implemented using multimodal data collected in two authentic settings, as described in Section 3.1.1.1 and Section 3.1.1.2.

Phase three included the validation collaboration proxies through semi-structured interviews and focus groups with 4 teachers and 23 students. Table 3.3 (RQ2, Cycle 2) summarises the analysis that was carried out to validate the potential of collaboration proxies. Qualitative video observations and a thematic analysis were carried out to illustrate the value of collaboration proxies in terms of awareness, visibility and accountability.

Finally, **phase four** documented the whole process for this cycle, from the conceptualisation of the approach to the reflections gathered from teachers' and students' perspectives. The outcome of this cycle is related to Contribution 2 of this thesis: A Multimodal Matrix approach. This is reported in Chapter 5:.

3.2.3 Cycle 3

In **phase one**, reflections from cycle 2 informed the development of cycle. This third cycle of the DBR process addressed RQ3, namely, the combination of visual elements with key group constructs to guide the interpretation and sensemaking of visualisations (Figure 3.13, Cycle 3). Thus, based on current works in LA and Information Visualisation principles, **phase two** was aimed at developing an approach to contextualise key group constructs into visualisations through data storytelling (DS) techniques. The approach was illustrated by the generation of prototypes using DS elements in the context of a database design activity (context 2).

Phase three validated a set of prototypes through two students. The first exploratory study validated the design choices of a preliminary set of explanatory visualisations by interviewing five potential users. As presented in Table 3.3 (RQ3, Cycle 3) a thematic analysis from the semi-structure interviews allowed to understand the sensemaking support and the helpfulness of the data storytelling elements.

Findings from the first exploratory study informed the design of a second set of prototypes, which redesigned the explanatory visualisations using the same learning context, as depicted in Figure 3.13, Cycle 3. Table 3.3 (RQ3, Cycle 3) lists the methods followed and analysis performed to validate the explanatory visualisations. A thematic analysis and gaze-behaviour analysis from semi-structured interviews, Likert-questionnaires and eye-tracking data helped to understand how data storytelling elements supported the sensemaking of explanatory prototypes.

Finally, **phase four** documented the whole process of this cycle, from the conceptualisation of the approach to the reflections gathered from practitioners' and teachers' perspectives. The outcome of this cycle is a learning design-driven data storytelling approach and is reported in Chapter 6:.

3.2.4 Cycle 4

For **phase one**, this last cycle is built on the insights from Cycles 2-3 in order to fully address RQ3, attempting to investigate how multimodal explanatory interfaces can be developed to guide reflection on group work (Figure 3.13, Cycle 4).

P**hase two** was focused on developing an approach to communicate multiple 'data stories' by scaffolding one insight at a time via a layered multimodal interface. To illustrate the approach, a capstone prototype was designed and implemented in the context of teamwork nursing simulation (context 1) (Figure 3.13, Cycle 4 - pink box).

Phase three included the validation of the final capstone prototype with eight subject experts through an elicitation process including a guided walkthrough of the prototype, a semi-structured interview and a Likert-scale questionnaire. Table 3.3 (RQ3, Cycle 4) lists the analysis carried out in this phase to validate the capstone prototype. A thematic analysis was performed to explore the value of the layers, validate the rules that were used to map the data storytelling elements with the learning design of the activity, and investigate orchestration opportunities.

Finally **phase four** of DBR documented the whole process of this cycle, including the final approach, which relates to Contribution 3 of this thesis: An explanatory visual layered approach, and the feedback gathered from subject experts. This is reported in Chapter 7:, and further discussed in Chapter 8:.

Table 3.3: Overview of the methods followed to address the research questions and goals of this thesis.

Research Question	Goal	Cycle	Learning Context	Prototype Designs	Validation Studies: Participants	Validation Studies: Methods	Validation Studies: Analysis
RQ1	Formalise stakeholders' requirements for designing multimodal interfaces from co-design techniques outputs.	1	Teamwork simulation scenarios	N/A	2 teachers	Semi-structured interviews Co-design activities	Inductive content analysis process Deductive theory-driven analysis
					13 students	Focus groups Semi-structed interview Co-design activities	
RQ2	Provide a model to systematically associate low-level multimodal data streams with key group constructs.	2	Teamwork simulation scenarios	collaboration proxies	4 teachers	Semi-structed interviews	Video observations Thematic analysis: ▪ awareness ▪ visibility ▪ accountability
					23 students	Focus groups Semi-structed interview questionnaire	
RQ3	Combine explanatory visual elements with group constructs to guide the interpretation and sensemaking of multimodal group interfaces.	3	Collaborative database design activity	explanatory visualisations	5 users	Semi-structured interview Likert-scale questionnaire	Thematic analysis: ▪ sensemaking support ▪ helpfulness of visual elements
				explanatory visualisations	6 teachers	Semi-structured interview Likert-scale questionnaire Eye-tracking	Thematic analysis: ▪ teachers' reactions (original vs. explanatory visualisations) ▪ ease of interpretation and orchestration ▪ data storytelling elements helpfulness Gaze behaviour analysis: ▪ time spent by teachers to inspect the visualisations ▪ Heatmaps of teacher's gaze behaviour ▪ areas of interest (AOI) ▪ clustering analysis ▪ scanning trajectories
		4	Teamwork simulation scenarios	explanatory multimodal layered interface	8 subject experts	Elicitation process Likert-scale questionnaire	Thematic analysis: ▪ value of layers ▪ rule validation ▪ orchestration opportunities

3.3 SUMMARY

In sum, this chapter has described two learning contexts that will support the evidence to answer the main research question: **How can explanatory interfaces be designed and used for guiding teacher reflection using multimodal learning analytics evidence?** It also has presented the rationale for, and implementation of, a Design-Based Research methodology, in combination with iterative design cycles. Table 3.4 summarises the contexts, the educational approach, the goal and number of groups/teams that were part of the data collection.

Table 3.4: Summary of data collection contexts that will be used in further chapters

	Context 1		**Context 2**
	Nurse simulation in a controlled scenario (data collection 1)	Nurse simulations in a classroom scenario (data collection 2)	Database design activity
Educational approach	cooperative learning	cooperative learning	collaborative learning
Goal	Teams working towards provide care to a deteriorating patient	Teams working towards providing care to a patient suffering an adverse drug reaction	Groups working together to solve and propose a database design solution
Participants	9 second-year students 3 teams (4,3 and 2 students)	30 third-year students 6 teams (5 students each)	15 second and third-year students 5 groups (3 students each)

Figure 3.14 summarises the data that will be used in further chapters to validate the contributions proposed in Chapter 1. Chapter 4 will describe how data from context 1 was used to generate prototypes and validate the proposed design approach to map evidence from co-design sessions to multimodal data. Next, data from context 1 will be exploited to validate the modelling approach proposed in Chapter 5. After, data from context 2 will support the exploration of data storytelling elements to create explanatory visualisations. Finally, in Chapter 7, data from context 2 will be utilised to generate a high-fidelity prototype of multimodal visualisations with data storytelling elements.

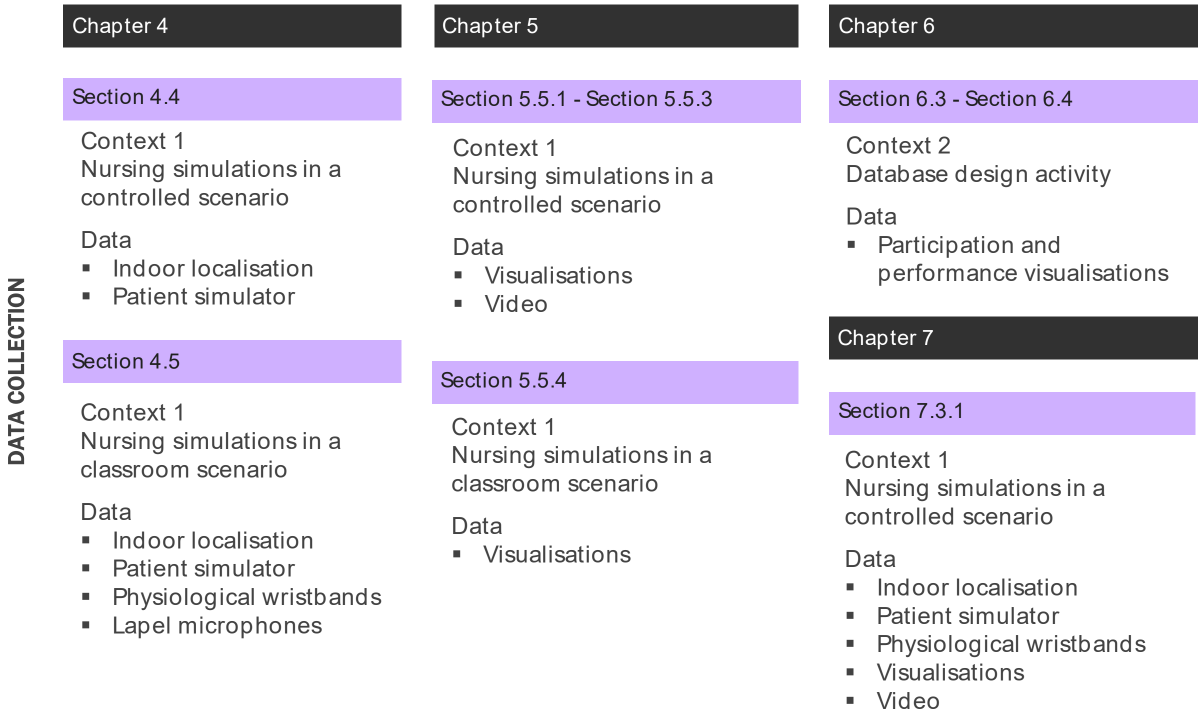


Figure 3.14: Summary of the contexts and data that will be used in further chapters.

Chapter 4: Mapping from Multimodal Data to Group Constructs and Analytics Requirements

This chapter presents an approach for mapping from multimodal data to group constructs and analytics requirements to create meaningful multimodal learning analytics (MMLA) interfaces. Evidence captured during co-design sessions with stakeholders is exploited as a primary source to formalise these requirements. The first part of this chapter presents the motivation for bridging Human Centred Design (HCD) approaches into Learning Analytics (LA). Second, this chapter describes the Human-Centred Design Multimodal Learning Analytics (HCD-MMLA) approach, which is grounded on human-centred design methodologies, theory and, teaching and learning practice. This approach is instantiated in the form of a design template aimed at supporting designers and researchers in creating MMLA interfaces. Third, this chapter illustrates how the staggered HCD-MMLA approach was operationalised in the context of teamwork healthcare simulation. Finally, the chapter concludes with a summary of the findings including *i)* the design requirements for a MMLA interface and *ii)* MMLA requirements in the form of a final template in regards of *teamwork* and *patient-centred care* learning constructs.

4.1 INTRODUCTION

From a human-centred perspective, MMLA solutions should be co-designed with the people involved in the learning activity (e.g. tutors, subject coordinators, students) to focus on untangling the learning activity rather than focusing on the tools that can be used to capture learners data (Martinez-Maldonado, Pardo, et al., 2015; Prieto-Alvarez et al., 2017). However, it is still not clear how researchers or designers can scope and define the requirements for an MMLA interface which should ideally include an extensive understanding of the learning environment, the intended learning design, teachers and students' expectations and educational theory. In the area of LA, current works using HCD methods have proven effective in exploring, defining and implementing potential solutions to support classroom orchestration (Holstein et al., 2017; Martinez-Maldonado, Dimitriadis, et al., 2013).

Figure 4.1 shows the research question, goal, contribution and validation methods addressed in this chapter. This chapter tackles the following research question*: How can stakeholders' requirements be formalised to design for multimodal interfaces?* To address this question, this chapter presents the

Human-Centred Design Multimodal Learning Analytics (HCD-MMLA) approach to map from multimodal data to higher order learning constructs with the purpose of defining the learning analytics (i.e. analytical, visual and technical) requirements that can help create effective multimodal learning analytics (MMLA) interfaces. The formal definition of learning constructs takes the form of a hierarchical mapping of sub-constructs and behavioural markers that can be used to determine the higher order constructs. Following a similar hierarchical mapping approach, the type of analytics or indicators, sensors and potential visual representations needed to characterise a behavioural marker can also be mapped. In subsequent subsections, the operationalisation of the HCD-MMLA approach is illustrated in the context of teamwork healthcare simulations.

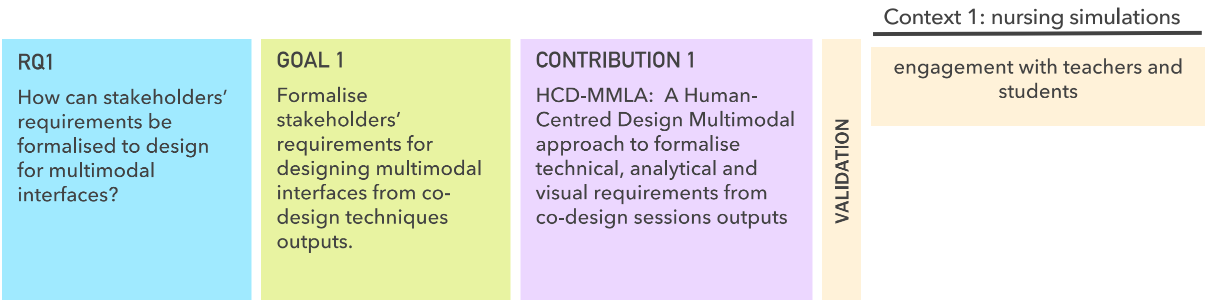


Figure 4.1: Research question, goal, contribution and validation methods addressed in Chapter 4.

The rest of this chapter is structured as follows. Section 4.2 motivates the application of HCD approaches in LA research. Section 4.3 presents the HCD-MMLA approach and its stages to map from multimodal data to higher order constructs aimed at defining the learning domain and analytics requirements for developing MMLA interfaces. This approach is materialised as a mapping template. Section 4.4 illustrates the operationalisation of the HCD-MMLA approach in the context of teamwork healthcare simulations and Section 4.5 reports the results and findings. Finally, a summary of findings is presented at the end of the chapter and recommendations for future work in this particular area of research are addressed in Chapter 8.

4.2 MOTIVATION

This work is motivated by research on Human-Centred Design (HCD) and Learning Analytics (LA) methods (Buckingham Shum, Ferguson, et al., 2019) to establish the requirements for MMLA interfaces considering stakeholders' perspectives.

4.2.1 Human-Centred Design for Learning Analytics

The main focus Human-Centred Design (HCD) is the closer involvement of affected people during the development of product or solution. Giacomin (2014) describes an HCD practice as *"the use of techniques which communicate, interact, empathise and stimulate the people involved, obtaining an understanding of their needs, desires and experiences..."*. HCD methodologies ensure higher acceptance and adoption of a product or solution. The well-stablished ISO standard in *human-centred design for interactive systems* (Standardization, 2010) has defined six key principles that designers should consider in the development of interfaces:

- The design is based upon an explicit understanding of stakeholders, tasks and environments.
- Stakeholders are involved throughout design and development.
- The design is driven and refined by human-centred evaluation.
- The process is iterative.
- The design addresses the whole user experience.
- The design team includes multidisciplinary skills and perspectives.

For instance, the HCD cycle proposed by Norman (2013), established a four-phase iterative process including *observation*, *idea generation*, *prototyping* and *testing* (see Figure 4.2). The *Observation* phase seeks to gather information about the problem to be solved by understanding the people involved and all the surrounding environment where the problem lies. Interviews, focus groups, surveys, participatory design activities and ethnographic studies are some of many other techniques that can be used to collect quantitative and qualitative data to explore the problem in a comprehensive way. The *idea generation* phase aims to make sense of the information gathered in the previous stage by comparing, contrasting and organising the information and look for important aspects, challenges and opportunities to consider in potential solutions. In this phase, designers start to generate ideas and define requirements for one or several potential solutions. Affinity diagrams, journey maps, principles and goals; and personas, are some of the tools and techniques that can be adopted to synthetise the information, extract themes, insights and create concepts that would be implemented through the next stages. In the *prototyping* phase, people involved in the design process create prototypes using the requirements and the core information extracted in the previous phase. Different solutions could emerge on this step to explore all the possibilities and maximise the

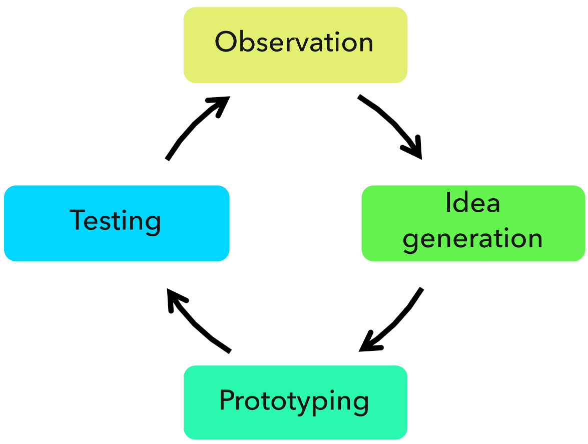


Figure 4.2: Human-Centred Design iterative process proposed by Norman (2013).

potential of finding the most effective one. In this stage, low-fidelity and high-fidelity prototypes can be used to explore the potential and opportunities of the selected solution. Usability and suitability studies are carried out to provide initial insights about people's behaviours and thoughts. Tools and techniques that can be used at this point include sketching, affinity diagrams, storyboards, brainstorming and critique. Finally, the *testing* phase aims to deploy a viable solution and run pilot studies with people. In this phase, a beta version or a more mature version of the proposed solution can be tested. In addition, designers are able to reflect on the whole process based on the lessons learnt on each stage. This will help to refine the solution for the next iteration.

In the LA research field, the same principles and HCD processes should apply when developing solutions to improve teaching and learning practices. Existing research in LA has pointed out that teachers, students and other academic staff should be involved in the design process of learning analytics tools (Prieto, Rodríguez Triana, Martínez Maldonado, Dimitriadis, & Gašević, 2018; Prieto-Alvarez, Martinez-Maldonado, & Anderson, 2018) Some preliminary works have reported that the effectiveness of the deployment of LA tools in authentic classrooms depends on the degree to which educators have been involved in the co-design process (Charleer, Klerkx, Duval, De Laet, & Verbert, 2016; Holstein et al., 2017). For instance, the workflow presented by Martinez-Maldonado, Pardo, et al. (2015) is an exemplar of an iterative process for designing and deploying awareness tools in the classroom, which goes beyond classic iterative development cycles (e.g. Helms, Arthur, Hix, & Hartson, 2006; Sommerville, 2011) by considering the pedagogical requirements and challenges in designing LA solutions. These authors identified five stages towards the development of LA awareness tools, namely problem definition, low-fidelity prototyping, high-fidelity prototyping, pilot studies, and validation in-the-wild – classroom use. Another illustration of applied HCD in LA contexts is the co-design process proposed by Prieto-Alvarez et al. (2017) to design innovations, centred in the inclusion of data, different people and roles. Authors suggested that the definition of data interactions and participants, as actors, during the whole iterative design process may drive important decisions to consider when designing and implementing LA solutions. Thus, the authors proposed an extended version of Ellingsen's (2016) iterative design thinking process, by adding two critical elements in LA research: *actors* of the learning analytics process (i.e. teacher, learner, researchers, developers) and *data interactions* conversations through the whole interaction design process.

The approach proposed in this chapter draws on HCD practices, tools and techniques to formalise the specifications collected from co-designing with stakeholders into requirements to create MMLA interfaces.

4.2.2 Aligning multimodal data with learning constructs

Designing LA interfaces is a complex task where key learning constructs, teachers' and students' intentions, and even educational theory should be carefully considered to communicate data meaningfully. For the case of MMLA interfaces, the mapping from multimodal data to meaningful information that can be useful in a specific the learning domain can be even more challenging, hence requiring a solution to formalise such mapping.

There is a small growing body of research in Learning Analytics and Educational Data Science (LA/EDS) that maps *from observable data to learning constructs*. A learning construct can be defined as "*a concept or idea related to students' behaviours, attitudes, learning processes and experiences that cannot be directly observable but are manifested and observed through the interaction with the learning environment"* (Cronbach & Meehl, 1955). Students are expected to demonstrate different behaviours and skills when exposed to a learning environment (Buckingham Shum & Crick, 2016). For example, in collocated collaborative learning research, researchers and teachers may want to investigate how leadership (i.e. the skill of being a 'good' leader) impacts task and team performance. That is, leadership can be defined as a learning construct that cannot be directly measured, but can be *observed*, either by a human or computationally, for example, in regards of communication interaction (e.g. turn taking) among team members, individual contribution, or the final product or outcome of the activity. In this sense, some studies have explored students' learning behaviours and skills in online learning environments through observable evidence (i.e. clickstreams) to investigate relationships with higher order constructs such as self-regulation (Wise & Hsiao, 2019) and engagement (Fincham et al.; Jovanović, Gašević, Pardo, Dawson, & Whitelock-Wainwright). Milligan and Griffin (2016) demonstrated how students' competencies can be associated with low level logs from click streams in MOOC's to assess students' levels of knowledge. Furthermore, Wise, Knight, and Buckingham Shum (In Press) established the importance of linking higher order constructs to aligned metrics, derived features and low level (automatic or semi-automatic) capturable events in collaborative learning environments.

Another growing body of research have adopted Quantitative Ethnography (QE) as an approach to guide the interpretation of *low-level logs to meaningful information* based on a theory-driven method (Shaffer, 2017). QE is a methodology that combines quantitative and qualitative methods to derive meaningful information from big and complex data. While, quantitative methods provide broad claims to shed light on learning behaviours using large amount of data through statistical and computational techniques, qualitative methods offer an in-depth understanding of the learning environment, its interactions and characteristics to reveal meaningful insights, but at a small scale. Combining the power of both methods, QE aims to link learning behaviours (i.e. constructs) grounded in educational theories with low-level logs (i.e. quantitative data) to help researchers distinguish meaningful nuances of learning. In collaborative learning contexts, QE has proven useful in revealing and modelling temporal patterns from individual and group learners in socio-cognitive learning activities (Csanadi, Eagan, Kollar, Shaffer, & Fischer, 2018). Furthermore, a recent work demonstrated how QE can be operationalised to explain collaboration processes, establish differences between low and high performing groups and predict academic performance in relation to the combination of epistemic and social dimensions of collaboration (Gašević, Joksimović, Eagan, & Shaffer, 2019).

In MMLA research, current initiatives have shown interest on developing conceptual models to map from *multimodal low-level data* to *learning theories*. Worsley et al. (2016) attempted to identify the appropriate mapping from learning constructs to multimodal data by conceptualising the relationship between constructs and multimodal data in terms of *learning constructs, indicators, analytic technique, analytic tool, type(s) of data* and *data capture device*. Similarly, Eradze, Rodríguez-Triana, and Laanpere (2017) proposed a pedagogy-driven approach to frame the philosophical (i.e. bottom-up or top-down approach), pedagogical (i.e. educational theory) and technological (i.e. architecture) aspects of a learning environment to enrich observational data with multimodal data. Specifically, authors defined a four-step process for this purpose, namely: 1) definition of the elements of the learning context, 2) definition of observable indicators, 3) collection of observable events, and 4) analysis and interpretation of results. Worsley and Blikstein (2018) proposed the use of epistemological frames to understand learning by aligning learning theories with epistemic states and multimodal data. Di Mitri et al. (2018) introduced the MLeAM to map from low-level multimodal data to learning theories (i.e. learning labels) by combining human interpretations (annotations) with machine learning and automatic inference. Likewise, Sharma, Papamitsiou, and Giannakos (2019) proposed a *pipeline* to align multimodal data features driven from literature with machine learning techniques aimed at predicting learning events. Although these studies highlighted the importance of mapping learning constructs with multimodal data, a critical point in this research is that most of these approaches were developed to fulfil the challenge of predicting learning processes and performance, (excluding MLeAM; Di Mitri et al., 2018), instead of aiming to design MMLA end-user interfaces. In addition, these approaches do not consider the human perspective (i.e. stakeholders) to connect the learning domain with analytics specifications, which is of relevant matter to inform the design decisions during the development of MMLA interfaces. Wang, Yang, Abdul, and Lim (2019) argue that HCI and AI researchers should enable people to understand, trust and effectively use intelligent systems and interfaces by bridging human-centred and theory-driven approaches to design AI systems.

The approach proposed in this chapter builds on these previous works by formalizing the requirements needed to design tailored and effective MMLA interfaces, forging connections between the learning domain, theory and multimodal learning analytics requirements.

The next section describes the proposed approach, which considers and applies HCD practices and theory-driven approaches to explore and understand problems, opportunities, challenges and the learning context through stakeholders' perspectives and translate these into requirements for MMLA interfaces.

4.3 HCD-MMLA: BRINGING HCD TO DEFINE REQUIREMENTS FOR MMLA INTERFACES

Grounded on well-established HCD practices and approaches to involve stakeholders during the development of MMLA interfaces and ensure a higher viability and adoption, the HCD-MMLA approach aims to formalise requirements for designing MMLA interfaces. Following Norman's iterative process (2013) which consists of *observation*, *idea generation*, *prototyping* and *testing* phases, this approach is unfolded through *observation* and *idea generation* phases. *Observation* phase allows designers and researchers to gather information about the problem to be solved and its context of use by involving and understanding people's needs. *Idea generation* phase aims to generate ideas and define requirements for one or several proposed solutions tailored to people's needs. This approach consists of three stages:

1) a *co-design* stage to elicit problems and challenges in teaching and learning practice, understand the context of use, and explore potential solutions to overcome these problems and challenges;

2) an inductive *mapping* stage to define requirements and link the learning context (i.e. group constructs) with the learning analytics specifications (i.e. data, analytics, sources and potential visual representations); and

3) a *theory-driven mapping* stage to validate these higher order constructs previously defined by stakeholders through relevant literature in educational theory, teaching and learning practices and analytics addressing similar problems.

The output of this approach is the definition of *requirements* (e.g. characteristics of sensors, analytics, and potential visual representations) for MMLA interfaces. Thus, the HCD-MMLA approach is contextualised by exploring stakeholders' needs yielding in a set of requirements; and theoretically motivated by looking at current literature to validate stakeholders' requirements. Figure 4.2 shows a diagram representing the three-stage approach to formalise MMLA requirements by exploiting the evidence gathered from co-design sessions with stakeholders. The figure shows the alignment of the three stages with the first two phases of Norman's iterative design process,

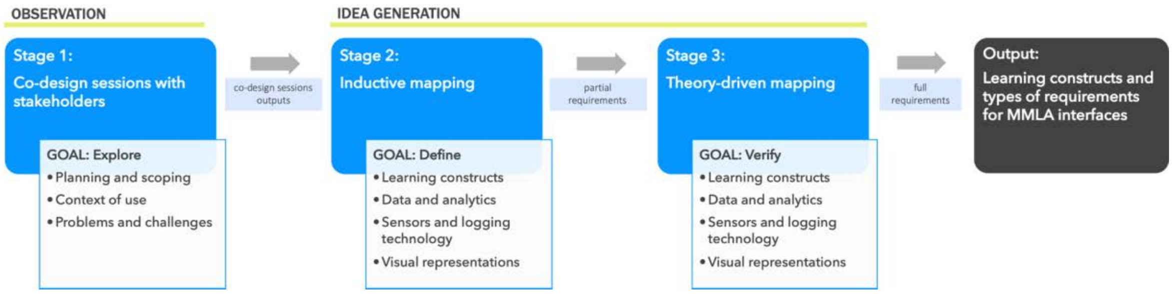


Figure 4.3: Diagram representing the HCD-MMLA method to formalise MMLA requirements.

namely, *observation* and *idea generation* phases, along with the specific goal for each stage.

4.3.1 Stage 1: Co-design sessions with stakeholders

Stage 1 aims to plan and carry out co-design sessions with stakeholders to gather information about the planning and scoping of the solution to be implemented; the context of use to identify potential users, tasks, technical and physical environment; and the exploration of problems, challenges and potential solutions (Maguire, 2001). Co-design techniques such as ideation, card sorting and voting, interviews, focus groups and learning journey maps (Prieto-Alvarez et al., 2017) can be used at this stage to envisage stakeholders' motivations; elucidate the learning context and practices in relation to stakeholders understandings; and explore the opportunities and challenges for developing an MMLA interface. Rapid prototyping can be used to explore what type of data and information stakeholders requires to explore in a visual interface. Outputs of this stage include annotations, transcripts of interviews and focus groups, pictures with a final card sorting, or answers from open-ended questions and questionnaires. These outputs are in the form of unstructured data (e.g. quotes, concepts, ideas).

4.3.2 Stage 2: Inductive mapping

The purpose of Stage 2 is to define a conceptual structure in terms of elements that represents the learning domain and learning analytics specifications (i.e. data sources, analytics, visual representations) More specific, qualitative, unstructured data, coming from the previous step is categorised and clustered through an inductive content analysis (Thomas, 2006) and affinity diagram clustering (Hanington & Martin, 2012) techniques. Content Analysis has been widely used in qualitative research methods that collect open-ended questions from semi-structured interviews, aimed at systematically analysing transcripts and documents using either an inductive or deductive approach. Affinity Diagramming is a widely used design technique to synthesise and capture requirements, concerns and insights by clustering these observations base on "*affinity*" or "*themes*". These "*themes*" help shape the findings related to the elicitation of problems and categorisation of learning constructs.

Inspired by the work presented by Milligan and Griffin (2016) and Wise et al. (In Press), three levels of abstractions were defined to represent the learning domain, namely A) learning constructs, B) sub-constructs, C) behavioural markers. Moreover, to connect the learning domain with learning analytics specifications (Greller & Drachsler, 2012; Worsley et al., 2016), three types of requirements defined, namely D) data/analytics, E) sensors, and F) visual representations. A detailed description on how to define these learning constructs and map those into requirements for MMLA interfaces is presented below.

4.3.2.1 *Defining learning constructs*

Learning constructs can be derived from theory or educational practice in a particular context. In order to systematically define a learning construct, a hierarchical representation of three levels is proposed: A) a ***higher order construct***, which encapsulates a higher student's skill, or behaviour; B) ***sub-constructs***, which describes a construct in more detail, and can be referred as a sub-component; and C) ***behavioural markers***, which are, either human or machine, observable indicators to measure a particular sub-construct.

For example, in a group work activity, teachers may have identified that *teamwork* and *patient-centred care* are two critical higher order constructs that they would like to emphasise and discuss with students to prepare them for real practice. Here, designers and researchers may have identified several constructs from teachers and students' understandings of the learning activity. The *teamwork* higher order construct can be associated with *communication*, *active participation* and *coordination* sub-constructs. If teachers are interested in highlighting *active participation*, some of the observable behavioural markers that they would expect to see in a group work activity would be the *verbal participation* and the *physical movement* for a participant during the activity.

4.3.2.2 *Mapping learning constructs with underlying data and analytics*

Researchers and designers of MMLA interfaces should justify the data and analytics to be collected, and how these are linked with the learning theory (Wise & Cui, 2018). Data sources and analytics may provide evidence to explain observable learning skills and behaviours or, as referred above, learning constructs. Analytics are any *statistical indicator* or *model* that can be extracted from a data source. From example, prosodic features (i.e. such as intonation, volume, tone, stress) can be extracted from audio recordings. These features can be further summarised as a *statistical indicator* (e.g. average tone, average volume) or can be used to create *models* of internal states (e.g. confidence, emotion, and so on) or intentions (e.g. sarcasm, jokes; Ochoa, 2017). If the teacher is interested in observing the *verbal participation* of the group, statistical indicators can be derived, such as the *average time* or the *percentage of time* of the speech activity per student. In short, this step is about connecting evidence (MMLA and multiple streams of data) with learning constructs.

4.3.2.3 *Associating sensing and logging technologies*

Associating the data and analytics with the sensors, devices or any equipment needed to capture most of the learning environment is critical for the design of MMLA solutions to be exploited in the wild, as this can offer the opportunity to evaluate the feasibility of the MMLA solution (Ochoa, 2017). Researchers and designers should analyse the trade-off between the potential intrusiveness of sensors and the authenticity of the learning environment (e.g. behavioural markers associated to cognitive load can be obtained using eye trackers which can be costly or hard to deploy in unconstrained settings). For example, if teachers want to see the *percentage of time* of the speech activity per student, each student should wear a *lapel microphones* or a *microphone array* should be located in the room, depending on the physical learning scenario and how the activity unfolds. This association implies that one or multiple *sensors* could be used to derive one or many *analytics*.

4.3.2.4 *Proposing potential visual representations*

Finally, linking visual representations with the underlying data/analytics and learning constructs helps researchers and designers address the misalignment between visual analytics, learning theory and stakeholders' involvement (Prieto-Alvarez, Martinez-Maldonado, & Anderson, 2018).. From stakeholders' perspective, a set of visual representations may be proposed for a particular context. Information can be presented, for example, using charts (e.g. radial, bar charts) to show summarised data; as a text narrative to give detailed facts; as a network graph to visualise social interactions; or as a timeline with critical events, where the temporality of the information can be inspected at a glance. Here, a *visual representation* can be associated with one or more *analytics*. For example, if the teacher thinks that *verbal participation* should be illustrated as a process through the whole activity, the *speech activity* per student should be visualised in a *timeline*. Also, if the teacher wants to explore student's physical movement enactment during speech activity, *physical intensity levels* can be showed into the timeline, ending up with two layers of information in one visual representation.

Figure 4.4 (inductive stage) illustrates how unstructured data such as excerpts extracted from transcripts, or concepts and ideas gathered from co-design sessions outputs are mapped into: A) learning constructs, B) sub-constructs, C) behavioural markers, D) data/analytics, E) sensors, and F) visual representations. To simplify the explanation in further subsections, these A-F elements conceptualise *learning constructs* and *types of requirements,* and broadly, these will be referred a*s headings;* while excerpts, concepts and ideas will be referred as *quotes*. As depicted in Figure 4.4, grey-coloured boxes stand for a sub-set of mapped headings. Clustered quotes in a grey-coloured box (e.g. sub-construct 1) represents a *theme* that emerged from the affinity diagram and clustering techniques. This *theme* will be related to the specific learning construct or type of requirement. The following stage will complete the mapping of headings by validating and completing the *themes* with relevant literature.

4.3.3 Stage 3: Theory-driven mapping

Once that the partial mapping of *learning constructs* and *types of requirements* has been carried out in the previous stage, more information may be needed to support the ideas and concepts externalised by stakeholders, as these may hold a partial view or a limited understanding of the learning activity. Stage 3 follows a deductive content analysis (Elo & Kyngäs, 2008) by reviewing theory and literature in the specific context of the learning activity (e.g. teamwork, leadership). The

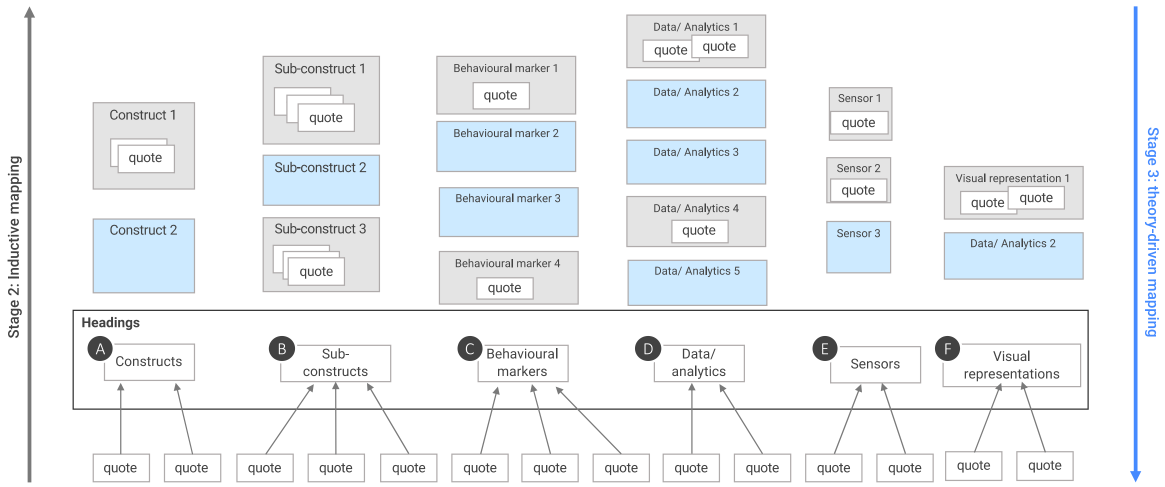


Figure 4.4: Inductive- theory-driven content analysis to define the learning activity and learning analytics requirements considering a set of pre-defined headings. The inductive stage maps the quotes from co-design sessions to headings, and then cluster similar emerging themes (grey-coloured boxes); while the theory-driven stage validates themes from educational theory and other relevant literature (blue-coloured codes).

rationale for the *theory-driven* stage is to back up current learning constructs with theory or any kind of domain expertise, and elicit subconstructs, behavioural markers, that were not explicitly expressed by stakeholders but remained important for describing higher order constructs. Besides this, reviewing relevant theory and literature also fulfils the purpose of re-examining analytics or indicators previously computed in other contexts or in similar research. As shown in Figure 4.4, blue-coloured boxes stand for *themes* that were not explicitly mentioned but retrieved from literature. As a result, a complete mapping of learning constructs and types of requirements is derived from stakeholders' quotes, ideas and concepts (grey-coloured boxes) and literature (blue-coloured boxes).

It is worth noting that, even though the proposed three-staggered approach was conceived to map from multimodal data to learning constructs and analytics specifications using information gathered from co-design sessions with stakeholders, Stage 1 provides an opportunity for researchers and designers to explore the challenges that stakeholders face during a learning activity and define the solution space by describing potential users and general requirements for developing a MMLA interface. Thus, this information should be also part of the definition of the final requirements that should be considered for developing MMLA interfaces.

4.3.4 The HCD-MMLA mapping template

To facilitate the mapping of co-design sessions outputs into requirements for MMLA interfaces, a design template is proposed as the output of the HCD-MMLA. Researchers and designers can configure or adapt this template according to their needs.

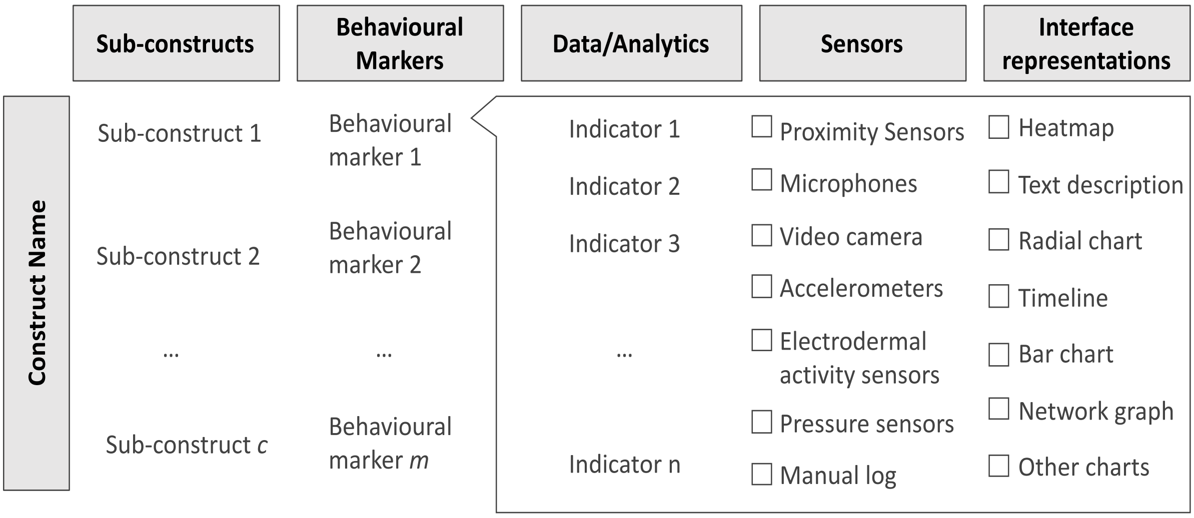


Figure 4.5: Template for mapping user's requirements with multimodal data, analytics and visual representations.

Figure 4.5 shows the mapping template for one construct, with headings, cells and checkbox elements to be filled by researchers and designers, taking the outputs from the HCD-MMLA approach. The heading "Construct Name" should be filled with a construct that was identified in Stage 1. Researchers and designers can define as many constructs as stakeholders have expressed to be critical. Each construct can be characterised by a set of sub-constructs (Figure 4.5, column 1). It is important to note that while sub-constructs may not be explicitly verbalised by stakeholders, learning theories or literature can help to fill in these gaps (as described in Stage 3). Next, behavioural markers should be mapped with each corresponding sub-construct (Figure 4.5, column 2). These can be elicited as the "kinds of evidence" that can be used to assess each construct. After, it is possible to define a set of MMLA statistical indicators or models that can be used to describe each behavioural marker (Figure 4.5, column 3). These indicators can be identified by stakeholders or obtained from the literature (Stage 3). The sensors or mechanisms that can be used to capture such indicators should be mapped with each data/analytics (Figure 4.5, column 4). Finally, it is possible to indicate which representations can help to communicate the data/analytics (Figure 4.5, column 5). In sum, this template can help to identify the constructs, data and technology requirements to design and create an MMLA solution.

4.4 ILLUSTRATIVE SCENARIO

This section illustrates the operationalisation of the three stages of the HCD-MMLA approach through co-design sessions with stakeholders to formalise the requirements for developing a MMLA interface in the context of teamwork simulations in the Bachelor of Nursing program. Healthcare simulations play an important role in the development of teamwork, critical thinking and clinical skills and prepare nurses for real-world scenarios. Preliminary works in this learning context demonstrated the potential of designing a learning analytics solution to support reflection and awareness of practical and teamwork skills (Martinez-Maldonado, Kay, et al., 2019; Martinez-Maldonado, Power, et al., 2017). Section 3.1.1 describes the learning context in detail.

4.4.1 Stage 1: Co-design sessions with teachers and students

Subject coordinators and students from the Bachelor of Nursing Program at the University of Technology Sydney were invited by the research team. Co-design sessions with two subject coordinators and five students' focus groups were designed to address Stage 1 goals (Figure 4.3, Stage 1). Figure 4.6 summarises the participants, tasks and outputs related to the co-design sessions carried out with teachers and students.

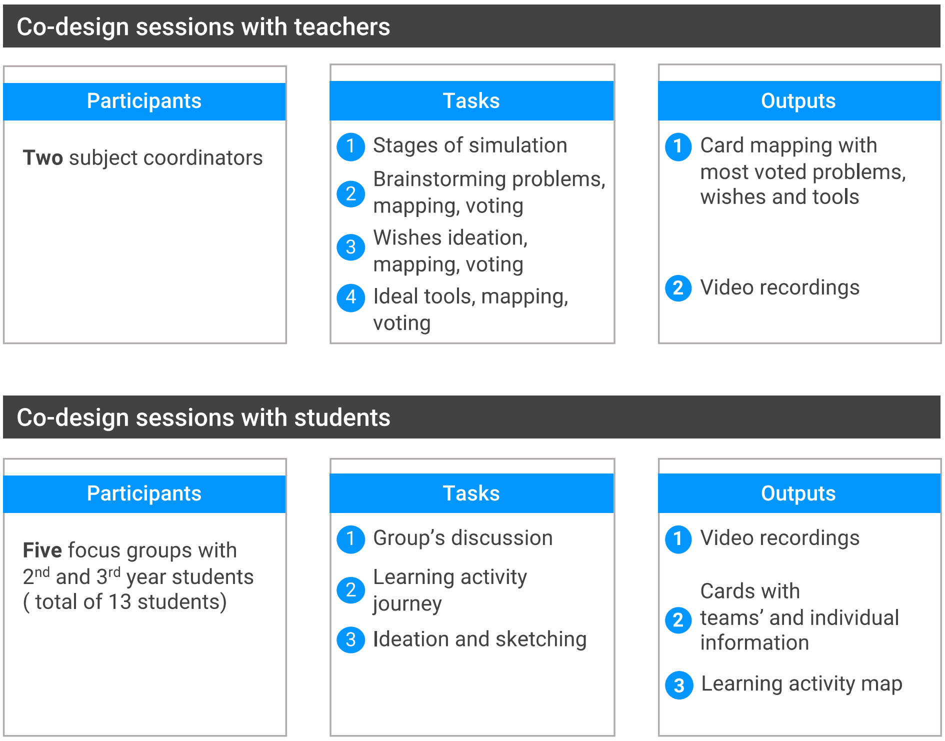


Figure 4.6: Study design for co-design sessions with teachers and students. Co-design sessions with teachers was conducted with two subject coordinators and co-design sessions with students involved five focus groups.

4.4.1.1 *Sessions with teachers*

To gain a better understanding of teacher's requirements in the context of nursing simulations, two subject coordinators (teacher A and B) of the Bachelor of Nursing program at UTS were invited to participate in co-design sessions. These two female subject coordinators usually run simulation-based activities over the semester and have been teaching simulation related subjects for about five years. The interviews were semi-structured and followed a 'think aloud protocol' with different co-design techniques such as *brainstorming/ideation, card mapping* and *voting* (Hanington & Martin, 2012). The *brainstorming* technique has been widely used to explore new concepts and ideas. This technique along with graphical representations of ideas help designers develop and facilitate participant thinking strategies. *Card mapping* or card sorting is another well-known design technique to help participants generate and link concepts, identify terminology or structuring information. *Voting* aims to prioritise ideas and focus on solutions that are more valuable. The interview consisted of four main tasks:

- **Introducing the activity:** Teachers were asked to recognise and write in some cards the stages that are performed during a simulation class. These can be referred as *Stage Cards*. In this way, the aim of this task was to understand how the activity is usually performed and which are the relevant terms that teachers use to name the different stages of the learning activity.
- **Brainstorming problems, mapping and voting:** Teachers were asked to verbalise the *problems* (P) related to their teaching practice in the simulation classroom and write them down on yellow sticky notes (one problem per sticky note) to externalise their thoughts. Next, teachers were asked to map each *problem* with the *Stage Cards*, to identify where the teacher needs more support. Later, to prioritising the design options for teachers, they were instructed to follow a dot voting activity. Teachers were provided with six blue dot stickers and were asked to give priorities to three problems, so teachers could place three votes when the problem priority was high, two votes when the priority was medium and one vote when the priority was low.
- **Wishes ideation, mapping and voting:** Following the approach proposed by Holstein et al. (2017) and fabulation technique (Hanington & Martin, 2012), teachers were asked the following question: *If you have a magic wand and you have infinity wishes to help students in a simulation class, what they would be?* In this way, the aim of this task was to explore any requirement they may have without constraining their ideas to thinking in terms of technical limitations or feasibility of a solution. Teachers were instructed to write their *wishes* (W) down in green sticky notes. Next, they were asked to map *wishes* sticky notes with any identified *problem* (P). Then, they voted again for the three top *wishes* with voting stickers following the dot voting as in the previous step.
- **Ideal tools, mapping and voting:** In order to explore what kind of tools they would expect to have at hand to support the problems and wishes explored in previous steps, teachers were asked to describe and write down in blue sticky notes any technological *tool* (T) they could envisage would help or assist in the simulation class. Next, teachers were instructed to map the *tools* sticky notes with any identified *wish* (W). Finally, teachers prioritised three *tools* with their voting stickers following the dot voting as in previous steps.

Each teacher was interviewed separately. Sessions lasted around 90 minutes. Outputs of these sessions included video recordings and the mapping and voting of problems, wishes and tools of both teachers, which were merged into one mapping. Figure 4.7 depicts a digitalised version of the final output of the *brainstorming/ideation, card mapping* and *voting* activities after merging both teachers' mappings of problems (P), wishes (W) and tools (T). Sections 4.5.1 and 4.5.2 present and summarise the final output of these co-design sessions.

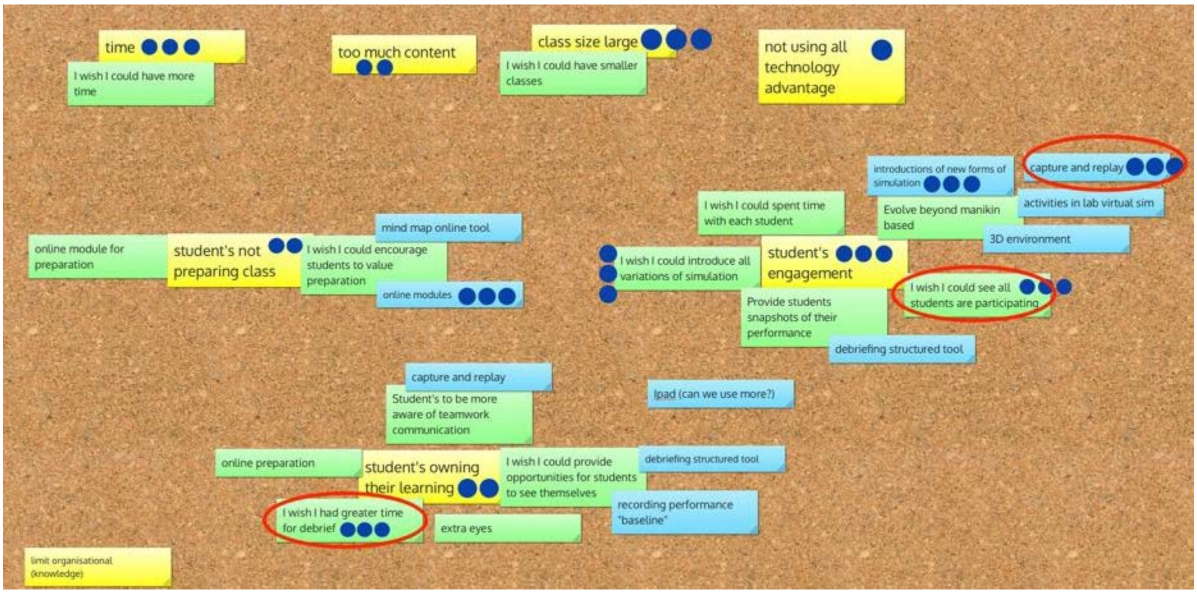


Figure 4.7: Final output of problems, wishes and tools after merging responses from teacher A and B.

4.4.1.2 *Sessions with students*

Five co-design sessions were conducted with second- and third-year students of the Bachelor of Nursing program at UTS. A total of 13 students (10 female) participated in this intervention. Three tasks were designed to investigate students' needs as following the guidelines suggested by Prieto-Alvarez et al. (2017) :

- **Focus group's discussion:** A focus group is a qualitative method broadly used to get opinions, insights and perceptions about a situation, problem or product from a well-chosen group of people (Hanington & Martin, 2012; Prieto-Alvarez et al., 2017). Semi-structured focus group discussions were conducted with students. Students were asked questions related to their perceptions on previous simulations experiences, problems they usually face, what information they would like to receive to improve their simulation experience and when (before, during, after the simulation), potential uses of a debrief tool and data privacy and sharing concerns.
- **Learner journey:** The learner journey technique can help designers to gain insights and understand people's experience through a visual representation of the real-world (Hanington & Martin, 2012). It also helps designers to understand student's activity performed in a learning environment at a higher level (Prieto-Alvarez et al., 2017). To gain insights about students' simulation experience, a template representing the physical arrangement of a typical lab classroom where they perform these simulations was prepared (students' desks, patient beds, medicine room, tutor's desk, see for example Figure 4.8, left). The learner journey template was used to trigger a discussion on particular activities they perform during a simulation,

highlighting specific locations and times (e.g. stages of the activity). In addition, during the discussion, both researchers that conducted these sessions, put different stickers to identify actions and locate where in the classroom: *i)* they have experienced bad and good moods and *ii)* would like to get help or information about their performance (Prieto-Alvarez, Martinez-Maldonado, & Shum, 2018). Then, students were instructed to draw lines to map the trajectory they follow during the class. The output of this task was a map with lines indicating the trajectory they follow during classes, bad and good moods and their location in the room, and the data that can potentially be captured in the simulation classroom.

- **Tool ideation and sketching:** Finally, students were asked to generate and write ideas in sticky notes about the (multimodal) data and information they think would be helpful for their simulation experience and visualise it in a technological tool. The output of this task was a set of sticky notes with ideas of the type of data and information students are willing to explore in a tool (Figure 4.8 - right).

Each focus group session lasted an average of 60 minutes. Outputs of these sessions included audio-recordings, learner's journeys maps, and tool ideation and sketching evidence from each focus group.

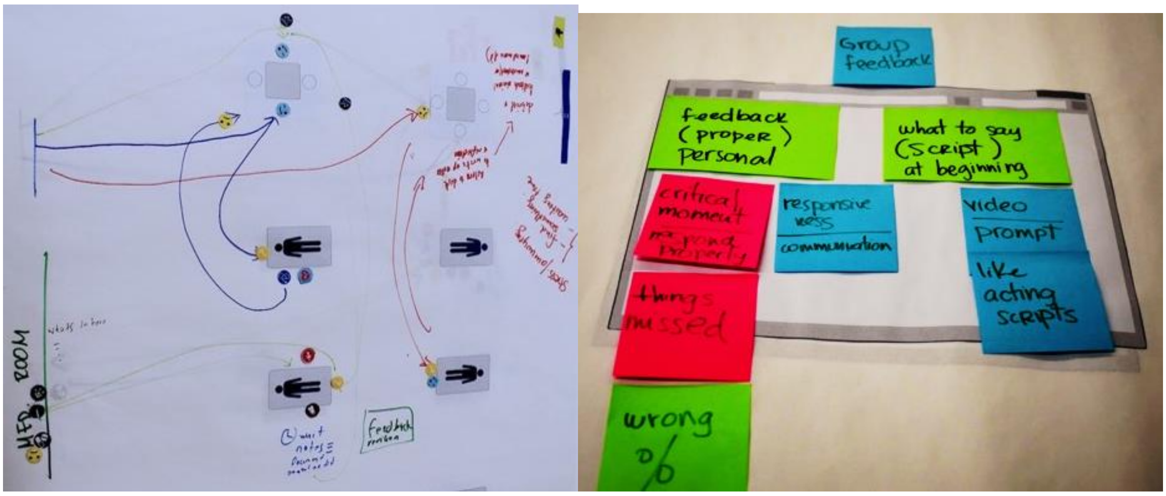


Figure 4.8: *Left*: Learner journey's output from one co-design session, pointing at locations where information can be provided. *Right*: Tool ideation and sketching output from one co-design session with students.

4.4.2 Stage 2: Inductive mapping

Stage 2 is focused on the analysis of the co-design outputs from the previous stage. All video and audio recorded sessions with teachers and students were transcribed and analysed using NVivo 14.

Approximately 7 hours of video were analysed following a Content Analysis and Affinity Diagramming procedures (Hanington & Martin, 2012).

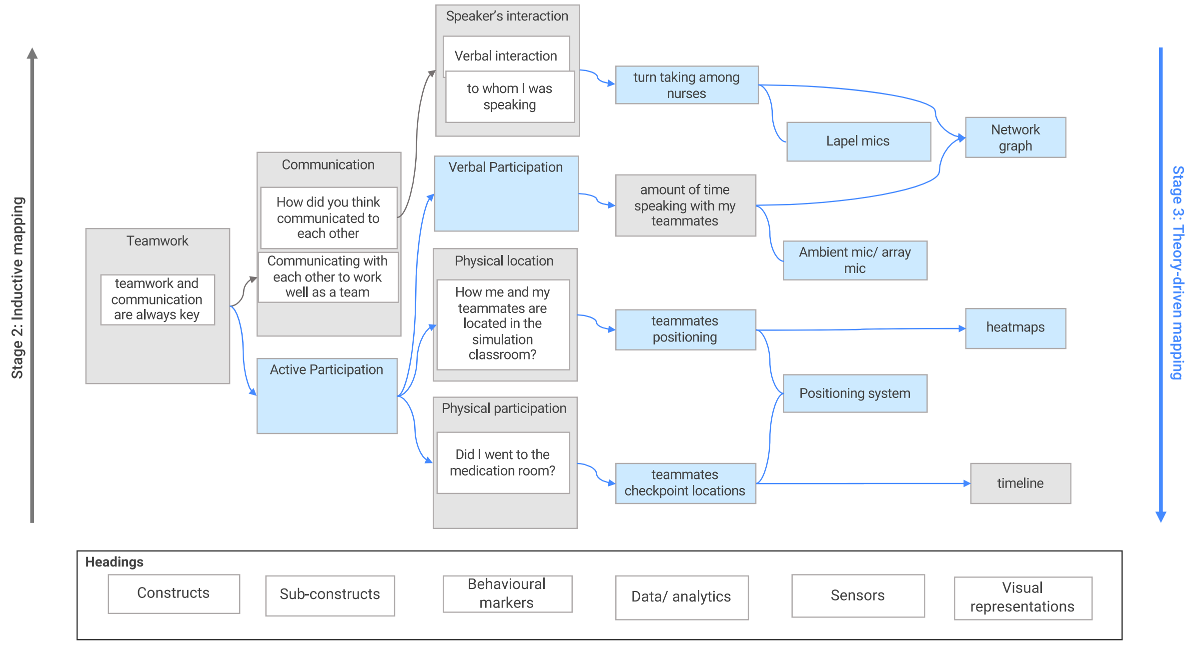


Figure 4.9: Inductive and theory-driven mapping example.

A total of 515 *quotes* were extracted from transcriptions and co-design sessions outputs, where 251 quotes were related to the identification of challenges, purpose and design requirements for developing a MMLA interface and 264 to the definition of learning constructs and learning analytics specifications through the inductive mapping. Only quotes related to the definition of learning constructs and learning analytics specifications were used in the inductive content analysis as described in Section 4.3.2. The other 251 quotes served to explore the challenges and elicit design requirements for a tool to support simulation classrooms. Affinity diagramming was followed for summarising and eliciting these design requirements. Results of this elicitation is presented in Section 4.5.1.

Thus, **264** quotes were grouped into fifty themes using affinity diagramming and then associated with a heading, i.e. ***constructs***, ***sub-constructs***, ***behavioural markers***, ***analytics/indicators***, ***sensors***, and ***visual representations***. The fifty themes were finally distributed into 5 constructs, 4 sub-constructs, 16 behavioural markers, 17 data/analytics, 3 sensors and 4 visual representations.

As can been seen in Figure 4.9, quotes were grouped to define a construct, subconstruct, behavioural marker, and so on. These are represented as grey-coloured boxes. For example, quotes such as *"teamwork and communication are always key"* was used to represent the **teamwork** construct theme. Next, quotes as *"how did you think communicated to each other"* and *"communicating with each other to work well as a team"*, helped to define **communication** sub-construct and to link this with teamwork construct. Furthermore, quotes such as *"information to whom I was speaking"* extracted from a co-design session with students; and *"verbal interaction"* extracted from the tool ideation and sketching; were grouped to indicate **speaker's interaction** behavioural marker and this was linked with communication sub-construct. Another quote *"how me and my teammates are located in the simulation classroom"* was used to indicate **physical location** behavioural marker, however, this behavioural marker was not linked with any sub-construct as teachers and students did not explicitly mention a sub-construct related to physical participation.

4.4.3 Stage 3: Theory-driven mapping

To complete the mapping, the deductive content analysis was followed as described in Section 4.3.3. Relevant literature in nursing education, clinical simulations, teamwork and MMLA was reviewed to validate the information mapped in the inductive stage and support the ideas expressed by teachers and students. In addition, literature also helped elicit information that was not made explicit by teachers and students but was important to define the themes. Figure 4.9 illustrates some themes that were derived from literature and current work in the research area of group analysis and multimodal data, represented as blue-coloured boxes. For example, it can be seen that in the Data/Analytics column, *turn taking among nurses* has been added to the final mapping as to indicate that this analysis helps in the observation of *speaker's interaction* behavioural marker. Also, as mentioned in the previous stage, physical location behavioural marker was linked to **active participation** sub-construct, which was derived from literature in nursing education. As a result of this stage, teamwork and patient-centred care were defined as higher order constructs.

The next section presents the result of the operationalisation of the HCD-MMLA, consisting in the definition of design requirements and a final requirements' template with a detailed explanation of the complete mapping.

4.5 RESULTS AND FINDINGS

This section presents the results of the operationalisation of the HCD-MMLA approach by 1) eliciting design requirements for a MMLA interface (i.e. debriefing tool); and 2) defining the learning constructs and learning analytics requirements for a MMLA interface in the form of a template resulting from the mapping process.

4.5.1 Eliciting teachers and students' design requirements for supporting simulation classrooms

After analysing 251 quotes using Affinity Diagram technique, the following themes emerged in relation to teachers' and students' issues and needs in simulations classrooms.

Helping to orchestrate teams during simulations: Teachers wanted a debriefing tool to support the orchestration and monitoring of the across many teams working at the same time, where some of them may need more attention than others. For instance, teachers found challenging to observe students' engagement and participation during the simulation classroom. Teacher A indicated that it is difficult to monitor all students, as she is *"just one person going around each team"*. She also expressed that it is hard to give *"one-to-one teaching experience"* to all students. Likewise, students indicated that one tutor is not enough to "*supervise everyone at the same time"* (P3, focus group 1). Teachers also wanted to *see* all students' participation through this tool so they could monitor efficiently all teams.

Making sure all students have the opportunity to get personalised feedback. Teachers wanted a *debriefing structured* tool as they found challenging to give personalised feedback to help students own their learning in a limited time. Teachers and students noticed that it is a demanding task for the tutor to give each student personalised feedback. One focus group rose this concern arguing that *"it is hard for teachers to give feedback to every student due to large classes"* (P1, focus group 5). Furthermore, students discussed that while teachers are willing to give feedback, most of the feedback is given when the opportunity comes. This was expressed by one student in focus group 3: "*if you ask to the tutor, she will give you answers about a certain procedure. But it is not like hey, this is your feedback*" (P3). Students found that summative feedback during examination is often the most valuable and key information they get to improve their learning. This was expressed as follows: *"In OSCE [Objective Structured Clinical Examination], tutors provide feedback, proper feedback"* (P3, focus group 3) and *"…in OSCEs you don't forget the things you do wrong. When I've done something wrong and got feedback, I haven't forgotten it."* (P2, Focus group 4). Teachers emphasised that all simulations are structured in a way that they end with a debriefing session, following best teaching practices for nursing education. However, due to limited time and large classes, it results impossible to give each group or student a detailed feedback. Teacher B expressed that tutors *"have a clear direction for the debriefing"* and opening up the opportunity to discuss all incidents will not be beneficial for students.

Giving information about team's performance and students' misconceptions. Teachers wanted to provide students with information about their performance with the support of a *debriefing structured* tool. Likewise, students found important to know *where* exactly they are making errors and struggling on their practice, as well as receiving positive feedback. For instance, students mentioned that they would feel more motivated when tutors highlight actions or procedures that went *wrong* (e.g. *"if you're doing something wrong, they could tell you, this is wrong"*; P3, Focus group 3) or went right (e.g. *"if we're doing a physical action, I want to be told I'm doing it right"* ; P2, Focus group 2). Students valued the preparation and time dedicated to debriefing sessions. However, they argued that tutors often lack of awareness on students' performance, as it was discussed by focus group 4, where students said that *"if teams or students need more specific feedback on performance, tutors are not always able to watch all students do a procedure"* (P1).

Giving timely feedback, either real-time or post-hoc. Teachers and students also discussed *when*, during the simulation, a tool may be helpful to support the orchestration of teams. Two divergent discussions emerged in relation to this. Two focus groups argued that presenting *prompt* *and immediate feedback* (as the activity unfolds) may help students to be aware on possible mistakes or missing procedures. Students from focus groups argued that if the information is given at the end of the simulation, they may not remember what they actually did. This point of view was illustrated as follows: *"I think I'd prefer to get feedbacks and remediation when there's a problem that arises rather than toward the end when the whole thing has become more blurred."* (P2, focus group 1); and *"I think a prompt, immediate feedback would help you remember what is going on during the simulation. Then after a while you get the feedback; you can't really remember what actually you did."* (P4, focus group 5).

On the other hand, some students suggested that the feedback should be offered immediately after the simulation (post-hoc), during the debriefing, as it would help them to reflect on their performance, without being distracted. This was expressed by a student as follows: *"I think I don't really want to be distracted by things that are happening right now. I want to focus on what I'm doing and then after that see if I did that [procedure] at that point."* (P3, focus group 1). Teachers also shared this perspective, arguing that debriefing sessions allow a timeframe to discuss and give feedback to students. Therefore, using a tool that shows team's performance and processes *"would be really helpful for students to see themselves"* (Teacher A) as it would guide a deeper reflection using evidence.

Making team's behaviour and performance visible. Both teachers and students agreed that if a tool is going to be used during simulations, it should make team's behaviour and performance *visible*. Teacher A expressed that *"anything that could capture team's performance in any way is important, to then run a discussion. It's having that technology to be able to capture and play straightaway"*. Also, Teacher B mentioned that *"a debriefing tool may help students to reflect on their practice, because they can view themselves"*. The debriefing tool would allow nursing students to make sense of the complete simulation experience. Often, students tend to forget what they did and if they did it well. Thus, *"capturing their body positioning, or their movement, or, you know, if they're leaning over the patient, or there's no moving .... all those things would help them [nurses] visualise, to complete the picture. I think this experience would be useful"* (Teacher A). Likewise, students commented that a debriefing tool would help them to get an additional perspective. This was mentioned by one student as follows: *"when you are in the simulation you can't see where you're positioned, you can't see how you're talking because you don't think about that. So, it would give me perspective"* (P1, focus group 4). Another student suggested to playback the information that were recorded to make sense of their conversations, procedures performed, and stress levels during the simulation: *"I think the most useful would be having the playback and then seeing where we're going with our conversations and then having what the mannequins respond to. If we're doing the procedures right, if we're hitting the right points that we need to look at. Even like stress level on the side would be pretty funny, I reckon."* (P3, focus group 4). Two focus groups discussed the potential of a debriefing tool to provide visibility on team's participation and reflect upon this information*: "I think it [MMLA interface] is a tool for improvement as well. This would make sure that people put the effort in the simulation and make those lazy people do more."* (P1, focus group 4); *"It would definitely be interesting seeing the amount of time speaking or explaining something to the patient"* (P1, focus group 2)*;* and *"also the interaction with teammates. With both of them. It would be good knowing if you have been clear"* (P2, focus group 2).

Being careful with data sharing. Opinions differed as to whether the information presented in a debriefing tool should be used for reflection at an individual, team level or the whole classroom. Some students suggested that the MMLA interface should be used as a ***personal*** feedback. Sharing information with other classmates would make students feel uncomfortable. For example, two students from different focus groups said: *"Personal would be better, because it's so embarrassing if it is in front of others"* (P2, focus group 4), and *"Everyone would just laugh if they saw each other or themselves and get embarrassed. So maybe it'd be good to have it personal"* (P2, focus group 3).

By contrast, other students considered that information could be shared at a ***team*** level, as they trust their team members. One student said that *"Usually we trust the team we choose to do simulations with. Whenever I'm in a lab, at the table I sit with, I'm okay with them knowing if I'm doing something wrong, we work together as a team. So, it's a team environment, so we take team criticism"* (P1, focus group 4).

Students in all focus groups agreed that the information presented in a debriefing tool can be shared and compared with other teams and the ***whole classroom*** as long as it is de-identified. Several students expressed this as follows: *"as long as is de-identified information, it is good."* (P2, focus group 3); *"I don't have any problem to share my information, as long as my name is not mentioned"* (P1, focus group 2); and "*If it is not identified, it might be ok"* (P1, focus group 4). Students also mentioned that they are willing to share their information with others, if there is a clear outcome or it helps to improve the learning experience (e.g. *"If other people are going to learn from it then it is good";* P2, focus group 3).

Offering this information to ***subject coordinators*** would help to raise awareness and detect difficulties and issues timely. In this way, students could have a better learning experience and satisfaction. One student mentioned: *"I would share this with coordinators, to make sure that I'm getting the same amount of education. Last semester I found my simulations were very poorly run, and I wasn't happy with what I was getting out of them, very not happy"* (P1, focus group 2). Other student commented on how this evidence could help tutors and subject coordinators to take a step back, plan and reinforce student's learning based on common errors found in practice: *"I think it would actually be good feedback for teachers and course organisers if they have everyone's data. Because, if everybody is missing the same steps .... it's something they can address in the next lesson or certainly in future semesters"* (P1, focus group 4). Furthermore, they would have the opportunity to detect flaws in simulations and propose other ways of training (e.g. extra labs). This was expressed as follows: *"throughout all of the student cohort there are going to be quite a few who have got the same issues. If this is the case, we would have set up practice labs for this skill. So, teachers would recommend that you go to it."* (P2, focus group 4).

Offering a reflection or coaching tool for teams. Teachers and students wanted a tool that could serve as an aid to guide discussions and potentially benefit individual and teams to reflect on their practice. Both teachers and all focus groups agreed that a *reflection* tool may support students' reflection on their practice and performance. The majority of students emphasised that this tool could potentially be used as evidence to check the right or wrong procedures followed during the simulation and their willingness to reflect upon these procedures for the next time. Students commented on this as follows: *"if the tool could track the events like safety procedures, I could know which I didn't follow"* (P2, focus group 1); *"I'll take a look [to the tool] and see if I'm doing really good in that area"* (P3, focus group 4); and *"For the next time I would remember what are the wrong things and I'll change my approach"* (P2, focus group 4). One student suggested that tool may help her to reflect on her *communication skills* with the patient. She explained this as follows: *"I do think that, even I don't like hearing my voice, I would get some good feedback from it [the tool]. When you do hear yourself talk someone, you realise if your tone of voice is the adequate or not"* (P1, focus group 4).

In addition, both teachers and most focus groups commented that they would need a *guided reflection* from the tutor, besides the information provided in a reflection tool. Students argued that reflection tool may provide opportunities to tutors to trigger discussions around the actions or errors performed during the simulation. One student said: *"If we do something wrong, she will… probably trigger a question, because she got the experience, so she can see what we've done wrong there."* (P1, focus group 3). Another student commented: *"Let's suppose that you have to take the medicine, but you didn't at that point or something like that. The tutor would ask "what do you think [about the missing action]? If you are having problems you can be given additional assistance"* (P3, focus group 1). Similarly, the reflection tool could guide discussions toward supporting the development of skills needed in professional practice. Teacher A supported this idea as follows: *"So, if we can capture what they were doing, and then once they're all together we can ask those specific questions on what you want… you know, what you think is important for them to take notice, for example teamwork and communication, which are always key."* And Teacher B commented that: *"This tool would be good to use in the debrief, because that's when you've got them all together. So, you could use this in the debrief session ad get them to think. Well, what do you think happened here? Get them to analyse the situation."*

Offering a training tool for tutors. A small number of comments were expressed on the development of a tool to training tutors. This tool could help raise awareness on teaching practices by comparing the actual learning design deployed in a class with the expected learning design. One teacher said that: *"this tool would be good for training the tutors. If we're just using it for training the tutors to tell them don't spend 15 minutes talking at the beginning, just… Because I'd rather they spent that time in here [doing a debrief]. So, you could have a tutor training session and show this during the tutor training session*" (Teacher B). Also, this tool could support the design of better simulations by providing information about difficulties on simulations across several classes. Another student commented: "*If the staff see that certain skills take more time to acquire, maybe they can give more time to that particular session. It gives feedback to the staff how well they're presenting so they can also improve their presentation.*" (P2, focus group 1).

Offering an assessment tool for students. Another topic that emerged was the development of an assessment tool by comparing team's performance with a baseline or expected performance, using a scale, a marking scheme or a check list. Three focus groups made comments to explain this as follows: *"I guess I'd measure it against what I thought the performance was. So, if I felt I did good and then that said I did rubbish, it'd obviously not be very uplifting for you."* (P1, focus group 4); *"maybe like a scale. If you achieved all the marks of the actions that needed to be done, then you have a higher scale of marking maybe on some things. Like, one out of five, so you had five because you got them all, but if you missed them, you could have lower."* (P2, focus group 2); and *"I'd like a marker to be in red or highlighted that something missed or done incorrectly. Because generally, you think you're doing the correct thing, so if it was just 70% correct, you wouldn't know what part of the procedure you did incorrect if it wasn't specified."* (P1, focus group 4). [baseline] One student mentioned that the tool could be used to assess their communication skills: *"I think to assess how we speak. How we're speaking to the patient…If we're using medical terms or just terms that they can understand or… Because you might not realise it but if you listen to it afterwards, say I could have said that better or I could have answered that question better that the patient asked, or I could have been more respectful or compassionate or something. Or I forgot to wash my hands, I didn't notice"* (P3, focus group 1).

In sum, this evidence suggests that teachers and students are aware that more opportunities for students to receive feedback are needed during simulation classes to foster their clinical and teamwork skills. However, these feedback opportunities have not been provided due to limitations enforced by administrative challenges, that usually most universities and programs face, such large class size and time to manage heaps of content in the courses. It is worth noting that these issues described above are not criticising the fact that administrators should be aware and act upon these problems. Instead, the exploration of challenges motivates the potential impact that a learning analytics solution may have in helping teachers and students to reflect on their practical and professional skills. Hence, this exploration presents a great opportunity to validate the potential of enriching these simulations with analytics, that could support teachers' orchestration and provision of detailed and personalised feedback to help students during their reflection time on their practical and professional skills.

The next section describes how teachers and students defined the simulation activity and what kind of information they are keen to explore, in order to empower their learning experience.

4.5.2 Defining the learning constructs and Analytics Requirements

The HCD-MMLA approach followed in the previous sections concludes with the mapping from co-designing with teachers and learners to higher order learning constructs, related to group work, and learning analytics requirements that are relevant in the context of nursing simulations. The final mapping resulted in the definition of two higher order constructs (HoC), namely *teamwork* and *patient-centred care*. In a higher-level a HoC is defined in terms of one or many sub-constructs; a sub-construct is described by one or many behavioural markers; and one behavioural marker is determined by one or many data/analytics.

Next, a description of each HoC is presented as follows: findings from co-design sessions outputs with teachers and students are described first (Stage 2), followed by literature reviewed to validate these findings or suggest sub-constructs, behavioural markers or data/analytics (Stage 3).

4.5.2.1 *Teamwork*

Figure 4.10 shows the template corresponding to teamwork construct, the sub-constructs, behavioural markers and data/analytics that are relevant to define teamwork as a higher order construct, along with a suggestion of sensors and visual representations to design a MMLA interface.

Teamwork is an essential skill to practice in simulations and is highly linked with the course aims and objectives for professional development. Teacher A expressed that simulation scenarios are ultimately designed to help students to ***improve their teamwork skills***. Teacher A stated this as follows: *"it's all about teamwork and about working together to try and plan care for the person* […] *teamwork and communication are always key things for them to notice".* In addition, in all focus groups students argued that they would ***prefer and value effective teamwork*** in simulations. One student explained this as follows: *"You learn it [teamwork skills] in uni[versity] but I think when you're on clinical, it's very important because you're working with professional nurses and you're then working in a team"* (P1, focus group 3). Another student mentioned that *"Sometimes you can't do things by yourself, you need a second nurse"* (P3, focus group 5). Connecting this evidence and current teamwork strategies in clinical settings, literature suggests that nurses should be taught to recognise critical events to externalise clear teamwork skills (i.e. how to work together) aiming at promoting higher team performance (Lerner, Magrane, & Friedman, 2009; Miller, Riley, & Davis, 2009; Salas et al., 2005).

HoC	Sub-constructs	Behavioural markers	Data/Analytics	Sensors	Visual representations
TEAMWORK	Closed-loop communication	Speaker's interaction	Turn-taking among nurses	☒ lapel mics ☐ array mics ☒ video camera ☒ accelerometers ☒ electrodermal activity sensors ☐ depth camera ☐ RFID tags ☒ indoor localisation tags	☐ heatmap ☐ radial chart ☒ timeline ☐ bar chart ☒ network graph ☐ Other charts
	Active participation	Verbal participation	Speech activity		
		Physical movement	Physical intensity levels Distance travelled		
		Nurses actions	Timestamped actions		
	Coordination	Verbal interaction	Turn-taking among nurses		
		Embodied interaction	Time spent on each zone, as individual and as team Time in close proximity with other nurses Zones occupied by each nurse		
		Joint actions	Sequence of actions per nurse		
	Leadership	Planning	Speech activity Social network analysis of speech interaction Graph analysis of leader proximity interaction Time spent near desk/bed footer		
		Task delegation	Logged sequence of actions per nurse		
	Stress Management	Arousal changes	Changes in arousal levels		

Figure 4.10: Defining teamwork construct (HoC) and analytics requirements. Grey boxes are themes that emerged during the inductive stage (Stage 2) and blue boxes are themes that were drawn by relevant literature (Stage 3).

Sub-construct: Closed-loop communication

Closed-Loop communication is a key sub-construct of teamwork. This was highlighted by Teacher A and B as follows: *"Teamwork and communication are always key. You know, how did you think you communicated with each other?"* (Teacher A) and *"[Students] need to be discussing and communicating with each other in order to work well as a team"* (Teacher B). Team members should be aware of their communication as a team, and how they interact each other in order to monitor and exchange information (e.g. *"I want to know to whom I was speaking"*; P1, focus group 2 and *"I would like to see the verbal interaction with my peers"*; P3, Focus group 2). According to Manser (2009), communication is the most frequent expressed issue during critical interventions, and a strongest predictor of surgical errors, therefore it is of critical matter to guarantee a good communication among team members. Similarly, Lerner et al. (2009) expressed that the lack of communication introduced medical errors during patient care due to the incomplete information exchanged by clinicians. Also, as argued by Salas et al. (2005), closed-loop communication is one of the three coordinated mechanisms for developing effective teamwork strategies, as it facilitates the exchange of information among team members and continuously monitor and anticipate other's needs. In short, both literature on teamwork and healthcare education support the idea illustrated by teachers and learners in terms of how important is to facilitate and teach communication skills in effective teams.

Behavioural marker and Data/Analytics: Closed-loop communication between nurses can be observed through **speaker's interaction**. Speaker's interaction can be considered a nonverbal cue that can be obtained from social interactions, by identifying to whom a person is speaking (Ochoa, 2017).

Motivated by several studies that have demonstrated the feasibility of detecting speakers interaction in co-located group work activities (Kim et al., 2008; Martinez-Maldonado et al., 2011; Mohan et al., 2018), a potential marker to describe speaker's interaction behaviour is by computing the **turn taking interaction among nurses**.

Sub-construct: Active Participation

In regard to active participation as a sub-construct to characterise teamwork strategies, teachers stated that "*Participation is pretty much everything*" (Teacher A), and that they would like to *"observe which team members are active participants"* (Teacher B). Students mentioned that a way to know the participation of the team is through *"the amount of time speaking with my teammates, discussing the information of the patient"* and *"movement, as I can know who was moving into certain part of the room to do a specific task"* (P2, Focus group 3). These quotes suggested that active participation could allow teachers and students be aware of who is really contributing to solve the task and who is active in their role. For instance, Bland, Topping, and Wood (2011) proposed that simulations are to be used as an integrated dynamic process of active participation and reflection to develop teamwork competences. Moreover, Riebe et al. (2016) argued that understanding teamwork participation could lead to improve the perception of distributive justice among team members and increase the team's satisfaction and performance. Overall, these studies support what teachers and learners said about participation in teams of nurses. Active participation could improve team's satisfaction and enhance team members' awareness of others' involvement and contribution to the task.

Behavioural marker and Data/analytics: Active participation can be observed through the **verbal (or speaking) participation** from a team member while helping the team to solve a particular task (Ketrow, 1999; Mehrabian, 2017). One student stated that: *"It would definitely be interesting seeing the amount of time spent speaking or explaining something"* (P2, focus group 2). Another student expressed some interest in observing the verbal participation *"between teammates"* (P1, focus group 2).

A number of studies have determined verbal participation in groups by measuring the total time of **speech activity** through the whole task (Bachour et al., 2010; Kim et al., 2008; Martinez-Maldonado, 2014; Mohan et al., 2018; Müller et al., 2018) or as a cumulative indicator (e.g. indicator of symmetry) aiming at promoting a more balanced participation, influencing social dynamics, detecting social behaviours and facilitating collaborative learning (Dich et al., 2018; Sturm et al., 2007). Hence, following similar same approaches as previous works, it could be feasible to extract the speech activity of each team member, either to represent their active participation during the whole simulation, or during particular stages of the simulation.

Behavioural marker and Data/analytics: Furthermore, team's active participation can be observed in terms of students' **physical movement.** Students mentioned that knowing their physical movement or transitions across the room along with the actions and roles they performed, would help them to be more informed on their team participation. It is expected that nurses actively move across the room, as they are exposed to high-stakes environments (Chappel, Verswijveren, Aisbett, Considine, & Ridgers, 2017). Depending on the task, nurses should move across the room to get instruments, provide care to the patient, grab medicine (e.g. walking to grab a mask, standing next to the patient). Hence, the physical movement of a nurse could determine if she was engaged or participating during the simulation.

Several analyses found in the literature may be used to measure the team's physical movement. As suggested by Petrosoniak et al. (2018), the **distance travelled by a nurse** may help to explain team physical movement and provide insights about team role and task allocation effectiveness, which may be difficult to observe at glance in high-stake scenarios. Moreover, according to Chappel et al. (2017), measuring nurses **physical activity levels** might help researchers to understand activity patterns and engagement in activities during hospital shifts. Usually, nurses engaged in light, moderate and high physical intensity levels (e.g. standing, walking) during their shifts, engaging in nursing care tasks such as completing documentation, providing care to the patient, administering medication, or in non-care tasks such as leisure-related activities.

Sub-construct: Coordination

Although this sub-construct was not explicitly mentioned by teachers and students during co-design sessions, literature on teamwork highlights its importance. Current literature in nursing education and teamwork suggest that coordination strategies are important when coaching effective team strategies, as it would maximise the performance of the team (Lerner et al., 2009; Riebe et al., 2016; Rosen et al., 2008). In the framework proposed by Kolbe, Burtscher, and Manser (2013), authors stated that coordination can be partially understood by observing explicit action coordination behaviours, for example by giving clear instructions to directly coordinate joint actions.

Behavioural marker and Data/analytics: Ethnography research in clinical settings has stressed that coordination strategies involve not only **verbal communication**, but also embodied strategies with tools and patient's location (Manser, 2009). Also, the study by Salas et al. (2005) proposed that effective teamwork strategies are exhibited on those teams that have clear understanding on their roles (social), tasks (epistemic) and resources (physical).

Verbal interactions can be measured by extracting the **turn-taking among nurses**. Similar to the analysis explained above to measure speaker's interaction, turn-taking among nurses is calculated by identifying who is talking to whom. In addition, what matters in this sub-construct is the potential recognition of verbal coordination strategies over time. Therefore, it is desirable to calculate the turn-taking among nurses by considering the temporality of the activity, to further illustrate the verbal interactions of each nurse associated with critical incidents of the simulation (Rosen et al., 2008).

Behavioural marker and Data/analytics: Embodied interactions allow teams to construct awareness and coordinate tasks during critical scenarios (Zhang & Sarcevic, 2015). As argued by Zhang and Sarcevic (2015), each role should be strategically located around the bed to guarantee timely and effective accomplishment of the critical task being performed. For instance, one student highlighted the importance of embodied interactions referring to the strategic location of a nurse when performing a critical task as follows: *"Location is important for some procedures. In CPR, it's actually quite important. There should be someone at the top of the bed and there's meant to be someone doing something on the other side. For CPR I think it will be quite useful, because I remember when we did our simulation, I was behind the headboard and I forgot to remove the headboard. That was my job and I was trying to bend over, because I was in charge of the masks."* (P3, focus group 4). Embodied interaction behaviours can be determined by estimating several indicators about nurse's physical presence in specific zones of the room and their interaction with other nurses or equipment (e.g. stethoscope tray, medication trolley) during critical tasks.

Preliminary works have attempted to provide analytics based on the position of nurses in the room to then measure the ***time spent on each zone of the room*** (Kannampallil et al., 2011; Olguin, Gloor, & Pentland, 2009b)**.** Similarly, other studies have demonstrated the feasibility to detect the ***time in close proximity to other team members*** (Isella et al., 2011; Olguin et al., 2009b).

Behavioural marker and Data/analytics: Joint actions or coordinated actions are usually measured by capturing the actions performed by each team member. While this behavioural marker was not explicitly stated by students, teachers made its own interpretation of coordinated actions. For instance, one teacher mentioned: *"during the simulation, nursing students would perform actions in parallel to effectively provide care to the patient"* (Teacher B). During the enactment of the simulation, it is expected that actions and interactions occur together, over time; that is, sequential actions and interactions may describe the team's behaviour and performance (Schecter, Pilny, Leung, Poole, & Contractor, 2018). Joint actions are expected to occur in high-stakes environments, in order to assure patient safety, either in well- or less- structured simulations (Miller et al., 2009;

Rosen et al., 2008). A recent work has shown the potential use of a computer-supported annotation tool to record in-class group learning behaviours and support group and teachers' awareness and reflection (Kharrufa et al., 2017). Therefore, following this same approach, teachers and/or researchers could observe and log ***sequential actions at a specific time*** during the enactment of the simulation, becoming the indicator that describes the joint actions among nurses.

Sub-construct: Leadership

Leadership also emerged as a relevant sub-construct to characterise teamwork strategies. Teacher A stated that during the simulation students are encouraged to *"not just to think critically, but as a leader too"*. She explained this as follows: *"If somebody is in a team leader position, she should be empowering their other team members to start thinking and doing things in a certain way"***.** Literature suggests that team leadership has been considered a key factor that warrants the success of the team. A good team leader is the one that guides, plans and suggests possible solutions (Salas et al., 2005). Moreover, leadership in a nurse's team must be assertive, as this allow the team to ensure patient safety (Miller et al., 2009).

Behavioural marker and Data/analytics: Teachers and students identified two main behavioural markers that could support the observation of leadership, namely **planning** and **task delegation**. For instance, **planning** was explained by Teacher B as follows: *"In a simulation scenario, students are expected to work together to try and plan care for the person"*. Similarly, students mentioned that *"the team leader will ask you to do things to provide care to the patient"* (P3, focus group 3) and *"the team leader is going to give you commands"* (P2, focus group 1). This evidence is supported by Salas et al. (2005) argument, where authors stated that team leaders should specify *"team member's roles and responsibilities"* (Salas et al., 2005, p. 34) at an initial stage of the activity, to support team effectiveness over time.

Analytics that support the measurement of planning behaviours of team leaders could be related to social and physical interactions among nurses or physical presence specific zones in the room that unveil leadership strategies. For example, it is expected that the team leader demonstrates higher **amount of time of speech activity**, as she is encouraged to give directions, guide other nurses verbally, ask for information and communicate with the patient. In current work, it was suggested that **social network analysis of speech interaction** (Gašević et al., 2019) could help team and teachers to reflect upon expected behaviours of team leader's interactions with other nurses. Regarding physical interactions, one student from focus group 3 suggested that *"a digital map of where the leader has been"* could help team to understand leadership patterns in regards of their physical positioning which can be contrasted with literature suggesting that a **graph analysis of leader's proximity interactions with nurses** (Isella et al., 2011) may help to describe and understand team leader's interaction patterns. Moreover, Teacher A stated that *"it is expected that* *the leader would be positioned at the end of the bed, giving instructions to others"*. As argued by Zhang and Sarcevic (2015), physical location of nurses in a hospital room is critical to guarantee the completion of a task in high-stakes scenarios. Authors stated that the team leader should *"stand in the back"* to monitor the work of other nurses. Therefore, this evidence suggests that the ***time spent near the bed footer*** would serve as a marker to detect planning and leadership strategies.

Behavioural marker and Data/analytics: Task delegation was also pointed out by students as a behaviour that leaders should demonstrate: *"the team leader should be delegating what someone else needs to do"* (P2, focus group 2); and *"there would be one team leader who delegates the task"* (P1, focus group 1). Salas et al. (2005) recommended that a team leader should be able to synchronise and coordinate individual member's contributions. Therefore, it is expected less actions performed by the team leader, as she should demonstrate the ability to delegate and monitor what other nurses are doing. One way to externalise this behaviour is by **logging actions performed per role or nurse**.

Sub-construct: Stress Management

Stress management sub-construct emerged from students' co-design sessions. Teachers indicated that usually nurses get stressed during simulations, and it would be helpful to get evidence of these states for a guided reflection during debriefing. Students indicated that they would like to know *"when they got stressed"* (P2, focus group 1); *"if the stress could affect their performance*" (P4, focus group 5); or if students are more prone to *"make mistakes due to stress"* (P2, focus group 5). Literature points out a growing interest in measuring and analyse stress episodes to make nurses aware of related events that could cause high stress levels aimed at helping nurses reflect on their practice and support the development of coping strategies for improving future experiences (Müller, Rivera-Pelayo, Kunzmann, & Schmidt, 2011). Researchers and educators are also interested in examining connections between the role of emotions (e.g. stress) and learning performance (Duffy, Lajoie, & Lachapelle, 2016).

Behavioural marker and Data/analytics: Arousal changes may help to detect stress episodes of nurses during the activity. While this behavioural maker was not explicitly mentioned by teachers and students, literature suggests that different levels of mental arousal (stress) are connected with learning performance (Broadhurst, 1957), arguing that there should exists a balanced level of stress to lead to an optimal learning performance.

An increase in electrodermal activity (EDA), specifically, is typically associated with changes in arousal states, commonly influenced by changes in emotions, stress, cognitive load or environmental stimuli (Darrow, 1964). When a change in the **level of arousal** (peaks) is produced, physiological responses are activated in the body (e.g. increasing sweat production, heart rate, and blood pressure) (Dawson, Schell, & Filion, 2007).

Sensors

Previous research in multimodal and group analysis have reported the use of lapel mics to capture and analyse speech interaction among group members (DiMicco et al., 2007). Video recordings from multiple cameras have also been exploited to capture group's verbal and non-verbal interactions such as physical movement and coordination (Müller et al., 2018). RFID accelerometers were utilised to detect group's physical activity and proximity interactions in a building environment (Ward et al., 2017). Electrodermal activity sensors have been used to monitor self-regulation by deriving sympathetic arousal as a marker of challenge and engagement (Malmberg et al., 2018). Indoor localisation tags may provide opportunities to track physical and zone proximity in group settings (Feese, Burscher, Jonas, & Tröster, 2014; Mohan et al., 2018).

Visual representations

There are many visual representations and dashboards that have been developed to support the understanding of group behaviours and performance (Liu & Nesbit, 2020). For instance, network graphs may be useful to display social interactions and detect a balanced participation (Kim et al., 2008) or leadership strategies (Feese et al., 2014). Also, a timeline representation may illustrate temporal individual and group's behaviours and visually detect collaboration patterns such as verbal and physical (touch) participation (Al-Qaraghuli, Zaman, Olivier, Kharrufa, & Ahmad, 2011) or visualise proximity interactions in firefighting teams (Feese et al., 2014). It is worth noting that these examples do not provide an exhaustive list of all possible visualisations that could be effectively used to illustrate teamwork strategies. Rather, these examples were selected to guide the selection of a visual representation that could be properly adopted to this context and further evaluated with stakeholders.

4.5.2.2 *Patient-centred care*

Figure 4.11 lists the subconstructs and behavioural markers that are relevant to define patient-centred care as a higher order construct, along with the data/analytics, sensors and visual representations that are required to design a MMLA interface.

Teachers and students expect the development of a patient-centred care skills, where patient safety and care is the key aspect of the activity. While this HoC was not explicitly stated by teachers nor students, substantial body of research in healthcare education recommends the incorporation of a patient-centred approach in current programs, with the aim of coaching nurses regarding their communication, problem-solving and psychomotor skills, tailoring these skills towards a human connection between healthcare professionals and patients (Mead & Bower, 2000; Weston & Brown, 2003).

HoC	Sub-constructs	Behavioural markers	Data/Analytics	Sensors	Visual representations
PATIENT-CENTRED CARE	Communication with the patient	Exchange of information	Speech activity with the patient Content analysis during patient's assessment	☒ proximity sensors ☒ lapel mics ☐ array mics ☒ video camera ☐ accelerometers ☐ electrodermal activity sensors ☒ depth camera ☐ RFID tags	☒ heatmap ☒ radial chart ☒ timeline ☐ bar chart ☒ network graph ☐ Other charts
		Confident speech	Speech pace: average speech rate		
			Tone: average pitch per nurse Speech style: detection of filled pauses Speech prosody: detection of emotion markers in voice		
		Embodied interaction	Body posture (relaxed, leaning to patient) Position of nurses (time standing at the bed side, near the patient) Care to the patient (actions)		
	Time management	Responsiveness to patient's critical moments	Time a nurse took to first attend the patient Time to complete the task		

Figure 4.11: Defining patient-centred care construct (HoC) and analytics requirements. Grey boxes are themes that emerged during the inductive stage (Stage 2) and blue boxes are themes that were drawn by relevant literature (Stage 3).

Sub-construct: Communication with the patient

Students indicated that that the communication with the patient should be critical for developing patient-centred care skills. This was mentioned as follows: *"I would like to know how we're speaking to the patient...If we're using medical terms or just terms that they can understand...* (P3, focus group 1). According to Kurtz, Draper, and Silverman (2016), communication is a core clinical skill that needs to be taught with the same formality as other core skills. Authors reported that training in communication skills for healthcare students should address the verbal content (what they communicate) and the communication process (how they do it) with the patient.

Behavioural marker and Data/analytics: Observing how nurses **exchange information** in a patient-centred care approach involves the examination on what nurses say to discover the patient's status, by asking questions or giving information to the patient. Student P1 from focus group 2 expressed that she would like to observe when she asked the patient about the pain (e.g. *"if I am using the correct words to assess patient's pain"*). Furthermore, understanding how nurses communicate with patient implies the analysis of non-verbal skills, their assertive behaviours to gather information and how nurses respond to patient's concerns.

Content analysis has been explored in online settings, by automatically analysing text from chat interactions (e.g. Leshed, Hancock, Cosley, McLeod, & Gay, 2007). Similarly, manual analysis on the content of group meetings transcripts have been analysed by using linguistic features to categorise verbal expressions (e.g. Neubauer et al., 2016).

Behavioural marker and Data/analytics: Confident speech features were mentioned by students during co-design sessions. For instance, the communication process that nurses should demonstrated when assisting a patient could be determined in terms of clarity of communication - *"it would be good knowing if you've been clear, you're speaking clear"* (P3, focus group 2); speed of speech- *"sometimes I speak too fast, but in my head, I speak okay"* (P1, focus group 1); and confidence - *"you may want to know if you're nervous to speak or you're just confident"* (P1, focus group 2). Literature related to the analysis of non-verbal communication suggests that confident speech could support the awareness of these markers expressed by students. For instance, **speech tone**, **pace** and **style** are some indicators that can be associated with clear, fast and assertive speech (Mehrabian, 2017).

Previous research in multimodal analytics has shown promising ways to measure speech tone, pace and fluency from audio streams. For example, some works have associated the **speech tone** by calculating the average speaking energy (Kim et al., 2008; Ward et al., 2017) or the average volume modulation (Olguin, Gloor, & Pentland, 2009a) from individual audio files captured in face-to-face collaboration activities. Other studies analysed the **speech pace** by extracting the average speaking speed (Kim et al., 2008) per participant in a group meeting setting and; the speech rate, articulation rate and average syllable duration from a person during an oral presentation setting (Luzardo, Guamán, Chiluiza, Castells, & Ochoa, 2014). Another study has shown how the **speech fluency** of a person can be measured by inferring the presence of filled pauses through formants in speech (e.g. Audhkhasi, Kandhway, Deshmukh, & Verma, 2009). Also, some recent works have explored the estimation of **speech prosody** features through speech-based emotion analysis from participant's voice in group work activities (Müller et al., 2018; Ochoa, Chiluiza, et al., 2018).

Behavioural marker and Data/analytics: Although **embodied interaction** was not explicitly mentioned by teachers and students during co-design sessions, literature suggests its importance to characterise non-verbal communication with the patient. According to Stein-Parbury (2013), encouraging patient-interaction is important as this will lead to an increased patient satisfaction. It also demonstrates nurses listening skills to understand patient's status and physical and psychological presence.

Inspired by Egan' model (2013) to develop effective helping skills, nurses may demonstrate interest on the patient's by exhibiting an attending posture with the patient. For instance, **body posture** was mentioned during co-design sessions. Teacher A stated that *"anything to show their [nurses] body positioning, or their movement would be useful to complete the picture about the simulation; for example, if they're [nurses] leaning over the patient, or there's no movement"*. The **position of nurses** when interacting with the patient, especially at the **bed side** was also noticed. One student mentioned that: *"Having feedback on patient interaction, bedside manner would be good, as everything happens at the bedside"* (P3, focus group 1). And, Teacher B explained the importance of nurses being at the bedside as *"they [nurses] have got all the information at the bedside, the progress notes and the medical record. So, they've got the history of the patient as well. So, we're trying to make it real."* Finally, linking **actions** related to the provision of care to the patient (e.g. taking vital signs) with embodied interaction behaviours would allow to support the detection of right or wrong procedures. For instance, one student complemented nurse's position (e.g. next to the patient) with the action of taking patient's vital signs to indicate that, if both occurred in an extended period of time, nurses would be checking the heart rate due to a previous error or dealing with an unexpected patient's behaviour. She said that: *"if we're taking vital signs and stay with the patient like for five minutes, it might happen that we're counting heart rate two times because we didn't get it right, it's fast, or maybe we're checking something else."* (P1, focus group 3). Literature in nursing education suggests that body posture, proximity and body movement (e.g. facing the patient at the side of the bed, showing an open posture, leaning forward, maintaining eye contact, and demonstrating a relaxed posture) add another layer of non-verbal cues that nurses should effectively demonstrate when attending a patient (McCormack & McCance, 2016).

Sub-construct: Time Management

Teachers and students made explicit comments about actions that should be done timely according to patient's status. Teacher A mentioned this by proving an example of an action that should be performed during a specific period of time in the simulation: *"it is reasonable if [nurses] managed to pick [ECG lecture] earlier in the simulation. It would have been better if [nurses] would thought of if quicker, as the patient needed to be monitored"*. Also, one student expressed *that "if it took me five minutes to take the vital signs, next time I'll do it promptly"* (P2, focus group 3). These quotes indicate that showing time management strategies are important to guide the reflection of nurses' best practices.

Behavioural marker and Data/analytics: Time management could be observed through nurses' **responsiveness to patient's critical moments.** One student mentioned that during the simulation she takes *"a lot of time drawing up the medicine"*, which made her think that sometimes she is *"too slow"* (P2, focus group 5). Another student reflected on moments during the simulation that nurses should be done things quickly, but instead *"people are just standing there, quiet, without responding"* (P2, focus group 3). These examples and literature suggest that high responsiveness to patient's critical status or being aware during the simulations are behaviours expected from nurses to exhibit proper time management strategies and high performance (Malec et al., 2007).

For instance, the **time a nurse took to attend the patient** after a critical status and the **time the team took to complete the task** may help detect team responsiveness (Kusunoki & Sarcevic, 2015). These time indicators may be used compare team performance with some assessment criteria drawn from teachers' expectations or general guidelines for providing care to a patient. For example, teachers would expect patients to take vital signs within a time frame of 2 minutes, however, the actual performance was around five minutes. Therefore, showing the actual performance in regards of the responsiveness for providing care to the patient could help team members to think reflectively about things that caused this low performance.

Sensors

Similar to the previous HoC, lapel microphones, recordings from video cameras and proximity sensors could be exploited to capture and analyse patient-centred care skills. However, due to the authenticity of the simulation classroom, background noise may hinder the detection of contextual speech information (e.g. speech content analysis), which is relevant for determining effective exchange of information with the patient. Previous work has used more sophisticated sensors, such as eye-tracking glasses to detect eye contact interactions (Schneider, Sharma, et al., 2018). Other works have determined eye contact interactions by approximating eye gaze from body postures extracted from video recording in a more controlled scenario (Müller et al., 2018; Ochoa, Domínguez, et al., 2018). However, eye contact can be challenging to detect in the context of simulation classrooms, as classrooms are not equipped with enough video cameras to cover all beds and interactions. Depth cameras could be used to detect body postures of students, as it has been demonstrated in controlled learning scenarios (Echeverria, Avendaño, Chiluiza, Vásquez, & Ochoa, 2014). Nonetheless, depth cameras could not be the best technological solution in nursing simulation scenarios, as it is difficult to track and identify all nurses' interactions with the patient (Martinez-Maldonado, Power, et al., 2017).

Visual representations

Heat maps can be used to visually detect if nurses were near the patient (i.e. hot spots) during the simulation activity (Martinez-Maldonado, Power, et al., 2017). Radial charts have also demonstrated to be useful when comparing several dimensions (Liu & Nesbit, 2020). For this context, radial charts may serve to summarise the position of nurses in relation to patient's proximity zones. Similar to the previous HoC, network graphs could be used to illustrate patient-centred skills by highlighting or including patient-nurse interactions. As mentioned above, the intention of these examples is only to illustrate and guide the selection of visual representations for further evaluation with stakeholders.

4.6 SUMMARY

This chapter described a **Human-Centred approach for mapping from multimodal data to group constructs and analytics requirements for designing MMLA interfaces**. A mapping template was proposed as the output of the approach. The template can be adapted to researchers and designers' needs. This work should be seen as the first step to generate Human-Centred MMLA innovations that focus on exploiting the power of sensing and analytics technologies in close alignment with pedagogical and learning needs. This chapter illustrated the application of the HCD-MMLA approach in the context of simulation classrooms. Co-design sessions with teachers and students helped to elicit design requirements and define the learning constructs and requirements for a MMLA interface.

The requirements for designing MMLA presented in this chapter provide first insights into the creation of MMLA interfaces that are tailored according to the pedagogical intentions explicitly stated in the learning design, teachers' and students' expectations, both educational and domain-specific literature, and other technical needs.

Summarising the results from the elicitation of design requirements for a MMLA interface (Section 4.5.1), teachers and students wanted a tool to support debriefing. Usually, teachers and students have a debrief session after the simulation, and this debriefing can be done at an individual-level, team-level or at a classroom-level. Teachers, students and subject coordinators are the main users of this tool.

Teachers and students suggested the development of:

- A tool for teachers to guide the reflection with students during debriefing, especially in large classes, where it is challenging for teachers to observe all teams and students at glance and provide personalised feedback at the same time.
- A tool for students to help them reflect on their practice by making visible teamwork and patient-centred care skills.
- A tool for subject coordinators to support evidence towards the justification of the learning outcomes of the activity; to plan better simulations or to train tutors.

In addition to this, the debriefing should contemplate the following features expressed by teachers and students:

- *Visibility.* The solution should support visibility of team's behaviour and performance.
- *Data privacy and sharing.* The solution should be used as a personal or team level. If the tool is going to be use at a classroom-level, then the information should be de-identified to avoid potential disagreements on data sharing.

As described through the definition of HoC's in Section 4.5.2, several sensors and equipment are needed to capture the simulation environment. Some analytics were identified to computationally detect nurse's behaviour in a team activity. These analytics, along with the sensors specifications and potential visual representations could be used by researchers as an opening point to deploy MMLA interfaces in this context. It is worth noting that a particular tool may not fulfil all the analytics requirements presented in this study. Instead, with the appropriate design and evaluation, it is expected that one or several MMLA interfaces may support debriefing sessions.

While this chapter has provided a conceptual approach to define learning constructs and requirements for a MMLA interface, the next step in this research is to provide a computational model to systematically map low-level multimodal data streams to key learning constructs. In practice, not all low-level data can be meaningful or critical, if the final aim is to give personalised and actionable feedback to face-to-face groups. Therefore, a well-structured method is needed to link the underlying data with more meaningful information.

Chapter 5: Giving Meaning to Multimodal Group Data

This chapter introduces an approach to model and give meaning to multimodal group data. The previous chapter tackled the challenge of mapping multimodal data from evidence collected in design sessions to learning constructs, analytical, technical and visual requirements for providing automated feedback to groups. This chapter moves forward and addresses the challenge of modelling key aspects of group work activities by enriching quantitative data streams with the qualitative insights needed to make sense of them. In response, first, this chapter describes the Multimodal Matrix approach, grounded both theoretically (in the physical, epistemic, social and affective dimensions of group activity), and contextually (using domain-specific concepts). Second, this chapter demonstrates how the Multimodal Matrix approach can be used to generate visual proxies of collaboration, derived from the physical, epistemic, social and affective aspects of the activity. Third, the validation of these proxies is presented. Teachers and students from the Bachelor of Nursing at the University of Technology Sydney were part of this validation.

5.1 INTRODUCTION

This chapter introduces an approach to give meaning to multimodal group activity data. Inspired by the metaphor of *social translucence* (Erickson, Halverson, Kellogg, Laff, & Wolf, 2002; Erickson & Kellogg, 2000), the approach proposes a vision to make evidence of collaboration "translucent", a term that will be explained. In doing so, it is important to emphasise that besides the *social* realm, collaboration also involves *epistemic*, *physical* (the use of tools, devices and the space) and *affective* aspects. The approach is built on the notion of Quantitative Ethnography (Shaffer, 2017) which seeks to give meaning to large scale quantitative data (such as generated by sensors) based on qualitative insights into the human contexts that give rise to the data (e.g. domain knowledge, theory). The approach is illustrated in the context of simulation in nursing, as described in Chapter 4 (Section 4.3). Simulated scenarios are representative of situations in which it is critical for team members to reflect upon evidence on different aspects of their activity.

Figure 5.1 summarises the research question, goal, contribution and validation addressed in this chapter. This chapter addresses the following question: *How can multimodal data be modelled to serve as proxies for group constructs?* To this end, this chapter illustrates how ideas from quantitative ethnography can help reveal important aspects of collaboration, from low level logs, to meaningful higher order constructs, in the form of the Multimodal Matrix approach. Four prototypes are presented as exemplar 'proxies' of collaboration from the elaboration of the Multimodal Matrix. The validation of these proxies with teachers and students is reported at the end of this chapter.

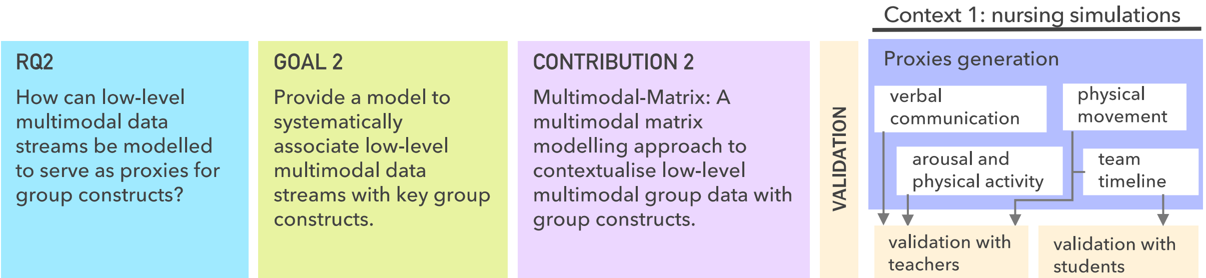


Figure 5.1: Research question, goal, contribution and validation methods addressed in Chapter 5.

5.2 THE MULTIMODAL MATRIX APPROACH[13]

Quantitative ethnography (Shaffer, 2017) is a method that lets researchers use statistical methods on fieldnotes, interviews and other kinds of qualitative data which has been mainly applied to build epistemic and social networks (Gašević et al., 2019). Inspired by this concept, the approach introduces grounding quantitative data in the semantics derived from a qualitative interpretation of the context from which it arises. A conceptual data representation is proposed, named the *multimodal matrix* (Figure 5.2), comprising the following conceptual elements: *dimensions of collaboration, multimodal observations, segments,* and *stanzas.*

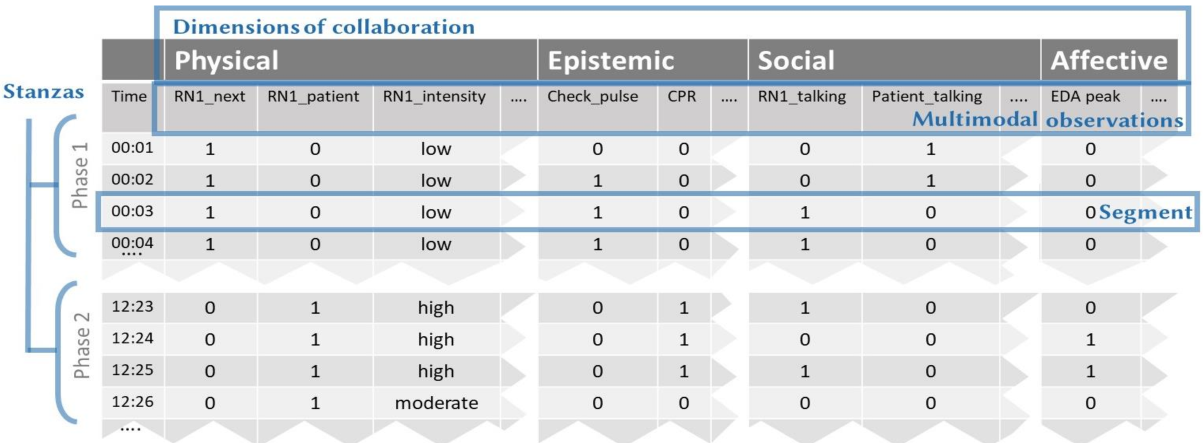


Figure 5.2: Multimodal Matrix representation

Dimensions of collaboration. These explain the complexity of group activity (groups of columns in Figure 5.2). As an illustration, the dimensions from the ACAD framework (introduced in Section 2.4.1) were selected (plus the affective dimension) as a set that can be considered. While not all dimensions need to be considered in every single multimodal study, having a systemic view of the

[13] Parts of this chapter have been published in CHI'19 (Echeverria et al., 2019) and ICQE'19 (Buckingham Shum, Echeverria, & Martinez-Maldonado, 2019).

key aspects of group activity can help researchers and designers to justify the emphasis of some kinds of connections over others (Hutchins, 2010) and to provide meaning to multiple sources of data used together.

Multimodal Observations. Each modality of data can be coded into one or more *kinds of information* called multimodal observations (columns in Figure 5.2). Each can be associated with a dimension of collaboration. For example, data obtained from the discourse or task-related actions would be associated with the epistemic dimension; communication data with the social dimension; logs of tool/space usage with the physical dimension; and physiological data the affective dimension. Some observations may span more than one dimension, but, most likely are associated with one dominant dimension. For example, dialogue *content* would be associated with the epistemic dimension, but quantitative features such as *turn-taking* with the social. This is a modelling decision, depending on researchers, designers or stakeholders' perspective. Critically, each column should only contain one kind of information and only one term can be used in each (ontological and terminological consistency; Shaffer, 2017, p. 129) p. 129). These columns are where each stream of data is coded into meaningful information based on theory, the learning design or pedagogical intentions, expert knowledge, domain knowledge. For example (see Figure 5.2, column 4), instead of simply logging raw accelerometer data, the data is encoded categorically (low/medium/high) in terms of a nurse's *physical intensity* while performing a sub-task. The meaning ascribed to the data would obviously be quite different for people in a meeting, or children playing outdoors.

Segments. Based on *quantitative ethnography* (Shaffer, 2017), segments are the smallest units of meaning considered for analysis (rows in Figure 5.2). The information contained in a row depends on the context. For example, in a highly qualitative analysis (e.g. discourse analysis) each line could correspond to an utterance. In multimodal analytics cases, many things would be happening in between one utterance and another (e.g. gestures, changes in eye-gaze, changes in physiological states). Each line might instead represent a time window (e.g. 100 milliseconds, 1 second; see second column in Figure 5.2) or critical incidents in the activity. As before, all information relevant to each part of the sample (e.g. a small window of time) should be in the same row and all rows should contain the same kind of information (evidentiary and ontological consistency; Shaffer, 2017, p. 167).

Stanzas. Segments can be grouped according to criteria to show meaningful relationships. In discourse analysis, a stanza might correspond to a number of utterances before or after a particular incident. In collocated group work, a stanza might correspond to well defined phases in the collaborative task (e.g. see rows grouped by phase in Figure 5.2). As shown in the next section, grouping the rows according to meaningful task phases gives meaning to the multimodal data streams. In collaborative learning tasks, these phases are sometimes made explicit in the learning design (e.g. see collaboration scripts; Rummel, Spada, & Hauser, 2009).

It is worth noting that the Multimodal Matrix approach has been introduced as a conceptual data representation. This approach could guide researchers into the implementation of an organised collection data structure (e.g. database, files). Information in the multimodal matrix can be populated from data gathered manually, semi- or fully automatically.

The next section illustrates the application of the approach to give meaning to multimodal group data in two teamwork settings.

5.3 ILLUSTRATIVE STUDIES

The two data collection studies in authentic nursing settings were used to illustrate the application of the Multimodal Matrix approach. Data collection 1 was conducted in an immersive simulation room (see Section 3.1.1.1), and data collection 2 in a simulated hospital ward-classroom (see Section 3.1.1.2). Both were conducted within the framework of the university's curriculum which emphasises *patient-centred care* (PCC) and *teamwork* — group constructs that were defined in previous chapter (see Section 4.5.2). This means that students must learn not only technical skills, but also develop communication skills to enable them to deliver professional care. The Multimodal Matrix modelling was applied to the multimodal data collection 1 and 2. Then, four collaboration proxies were designed and implemented using the information modelled in the Multimodal Matrix.

In the next section, the application of the approach is described, along with a detailed description on how to map low-level data streams to meaningful group constructs.

5.4 GIVING MEANING TO MULTIMODAL GROUP DATA

Each data stream captured by the sensors and devices in the data collection studies was encoded into columns in the multimodal matrix based on meaning elicited from teachers and students, the learning design, or literature (see Section 4.5.2). The multimodal observations used in the studies, and their relationship with the dimensions of collaboration are depicted in Figure 5.3, and described as follows:

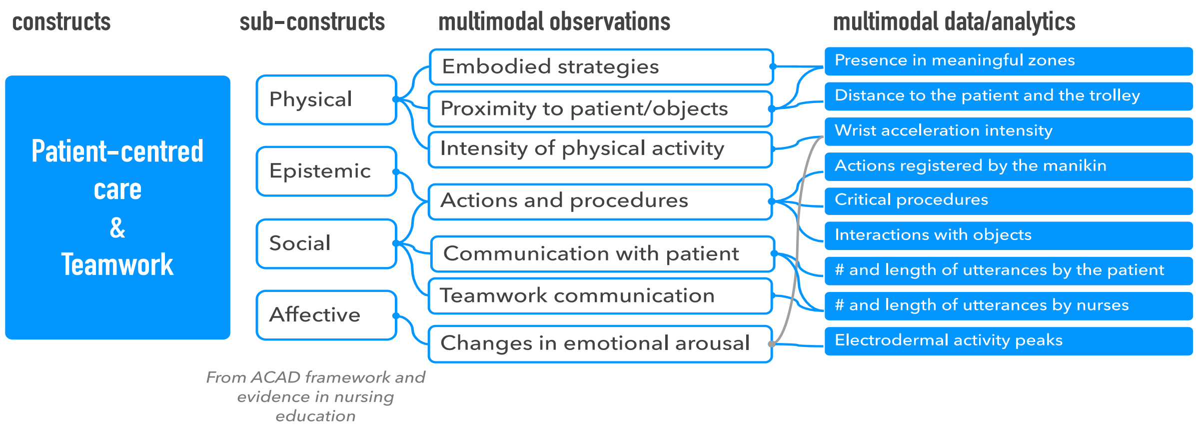


Figure 5.3: From multimodal logs to constructs in healthcare scenarios.

5.4.1 Embodied strategies and proximity

Embodied strategies during high-stakes teamwork scenarios are critical in healthcare education (Martinez-Maldonado, Echeverria, Santos, Dos Santos, & Yacef, 2018). Nurses are expected to be positioned in strategic areas when it comes to an emergency. Based on interviews held with nursing teachers, as reported in Section 4.5.2, and related work (Zhang & Sarcevic, 2015), five meaningful zones were identified which are commonly associated with a range of actions nurses perform (see Figure 5.4): i) *the patient's bed*, for cases in which nurses were located on top of or very close to the patient; ii) *next to patient*, for cases in which nurses were at either side of the bed; iii) *around the patient*, for cases in which nurses were further away from the bed, from 1.5 to 3 meters away of the bed); iv) *bed head*; which is an area where a nurse commonly stands to clear the airway (colloquially known as *bagging*) during a CPR procedure; and v) *trolley area*, for cases in which nurses were getting medication or equipment (a localisation badge was attached to the trolley).

Proximity data (for zones i and v) and localisation data (for zones ii, iii and iv) was automatically encoded into these meaningful zones. Five columns per nurse were added to the multimodal matrix representation: *RN(#).patient, RN(#).next, RN(#).around, RN(#).bagging* and *RN(#).trolley*. Each row has a value of "1", if that zone is occupied by a nurse, or "0" otherwise. The association of the zones at a specific period of time is mutually exclusive (e.g. [0,0,0,1,0] for a nurse in the bed head 'bagging' zone).

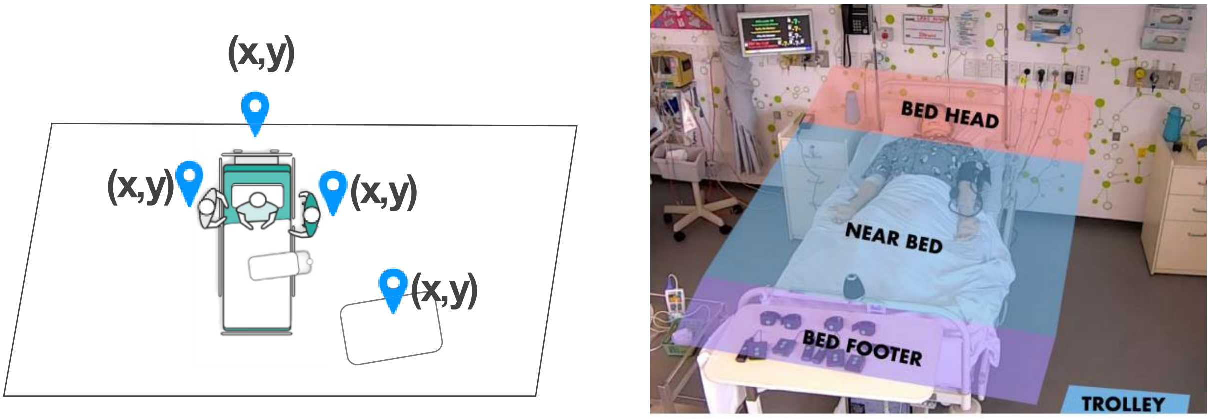


Figure 5.4: Translating data points (x,y) into zones of interest around the patient

5.4.2 Intensity of physical activity

Nursing literature (Chappel et al., 2017) indicates that nurses' physical activity may impact patient safety. Nurses' physical activity varies from light (e.g. walking, talking, manipulating medical tools) to moderate (e.g. performing a CPR) intensity. Three levels of physical intensity were defined, namely *static*, *light* and *moderate,* where the highest intensity corresponds to performing CPR (as can be depicted in Figure 5.5). For example, Figure 5.5 (left) depicts the EDA, BVP and accelerometer data plotted into a timeseries chart, with a shaded red area indicating that the nurse was performing chest compressions to the patient at that specific interval. Inspired by related work (Martinez-Maldonado et al., 2018) to determine the level of intensity, a moving average filter was applied on the raw acceleration data streams from the wristbands to remove signal noise (window sample = 32). Then, through visual inspection of these data, thresholds were defined for each physical intensity level taking values corresponding to CPR activity as the maximum level of intensity. One column per nurse *RN#.intensity* was added to the matrix, containing a value of 1, 2, or 3 for static, low, and moderate intensity, respectively in each cell.

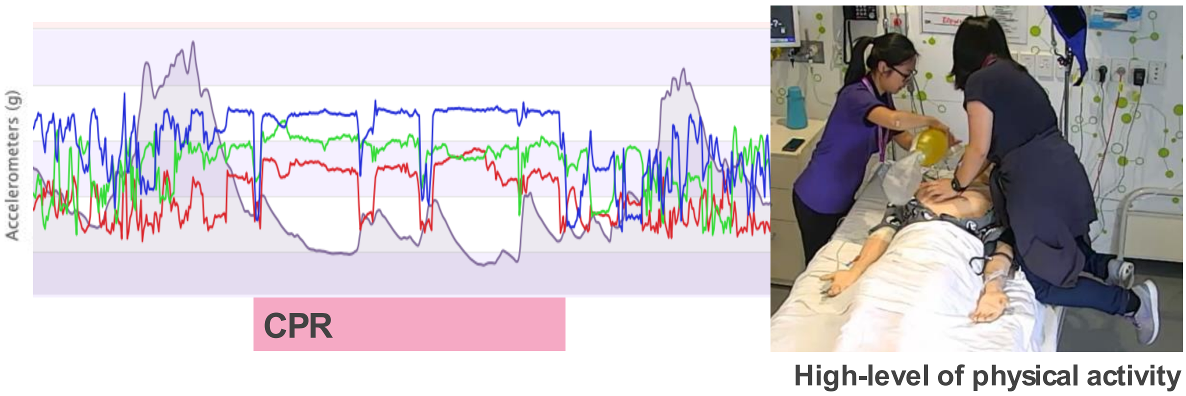


Figure 5.5: Translating accelerometer data (left) from a physiological sensor data into a high-level of physical activity (right).

5.4.3 Actions

Based on the learning design, a set of expected actions and procedures was defined and matched with the actions that could be captured, either automatically by the manikin or manually logged by an observer. In the matrix representation, actions were encoded into a column called *action* with a keyword per action performed and who performed it (if this information was available).

5.4.4 Speech activity and interaction

Verbal communication clearly plays an important role in the management and coordination of patient care. As indicated by teachers and students in Section 4.5.2., it is critical for students to understand how they are communicating with each other and with the patient in order to reflect and improve their communication skills. Nurses are encouraged to coordinate tasks, anticipate actions, and report information to the medical team in order to construct awareness (Zhang & Sarcevic, 2015). Non-verbal metadata about speech, in the form of speech onset/offset, may shed light on the team performance (Mohan et al., 2018; Sturm et al., 2007). Hands-free speech activity via a microphone array was captured in Study 1. Given the nature of the simulation, noise generated by clinical instruments made it hard to obtain clean streams of voice that could be automatically analysed through speech detection algorithms. For this reason, lapel microphones were used in Study 2. In the visual proxies presented in the next section, the dataset of Study 1 was reconstructed by manually transcribing and synchronising the video recording. As a result, one column per nurse and for the patient was added to the multimodal matrix (as *RN(#)_talking* and *PT_talking* respectively) to indicate the presence (1) or absence (0) of speech per second. In addition, one column per nurse and for the patient was added to indicate who was listening (*RN(#)_listening, PT_listening)*, if they were in close proximity to the person speaking (recall that in these trials, some actions were logged by a student observer or researcher as described in Section 3.1.1). For example, a row [0,1,0]; [1,0,1] (assuming columns correspond to R1, R2 and R3 for speaking and listening actions respectively) means that *RN1 and RN3 are listening to RN2.*

5.4.5 Electrodermal activity (EDA) peaks

It has been identified that physiological data can be effectively used to aid nurses in recalling confronting experiences in order to develop coping strategies (Müller et al., 2011). Also, as expressed by teachers during co-design sessions (Section 4.5.2) usually nurses get stressed during simulations, and it would be helpful to get evidence of these states for a guided reflection during debriefing. An increase in EDA, specifically, is typically associated with changes in arousal states, commonly influenced by changes in emotions, stress, cognitive load or environmental stimuli (Darrow, 1964). When a change in the level of arousal is produced, physiological responses are activated in the body (e.g. increasing sweat production, heart rate, and blood pressure; Dawson et al., 2007). Peaks were automatically identified using a tool called EDA Explorer (Taylor et al., 2015).

A peak in skin conductance was defined by a minimum increase of 0.03 µs, as suggested in the literature (Braithwaite, Watson, Jones, & Rowe, 2013). As depicted in Figure 5.6, EDA Explorer identifies peaks (green vertical lines) in the skin conductance signal. A column called *EDA peaks* was added to the multimodal matrix. Each cell contains a value of "1" when a peak was detected (green vertical line) or "0" otherwise.

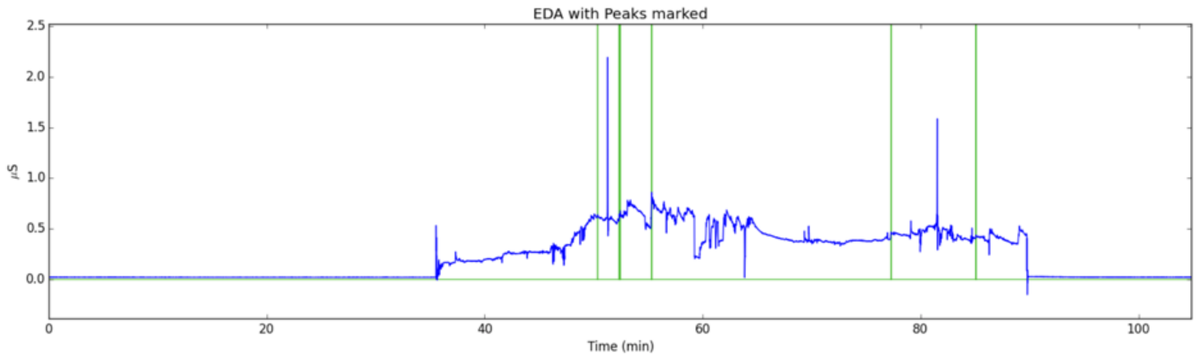


Figure 5.6: EDA peaks from the EDA signal of a student acting as a Registered Nurse (RN).

5.5 VISUAL PROXY PROTOTYPE DESIGNS

This section presents exemplar visual proxies for different aspects of collaboration, which can be generated to help make collaboration "translucent" (Erickson et al., 2002; Erickson & Kellogg, 2000). Proxies 1-3 present information from *Data Collection 1* (Sec. 3.1.1.1), focusing (respectively) on (1) how nurses communicated with each other and with the patient; (2) how they occupied the simulation space; and (3) how they may have experienced physiological arousal. Proxy 4 presents the critical actions performed by nurses during the simulation in *Data Collection 2* (Sec. 3.1.1.2).

For each proxy, its design, the dimensions of collaboration made visible, and the columns in the multimodal matrix used to model the proxy are explained in the next subsections.

5.5.1 Proxy 1: Verbal Communication

This proxy aims to depict to what extent the care provided by the nurses during the simulation was "patient-centred". This proxy integrates the *talking* and *listening* columns from the multimodal matrix. Segments are grouped into two stanzas corresponding to phases 1 and 2 of the simulation. In phase 1 it is expected that nurses engage in conversation with the patient who is still conscious. In phase 2, the patient goes into cardiac arrest and needs CPR. The nurse performing the CPR is encouraged to count aloud the chest compressions and coordinate with other nurses to synchronise the airway clearance and defibrillation.

5.5.1.1 *Design*

The requirement is to design a visualisation that represents teamwork and patient-care centred higher order constructs in relation to the communication with the patient and teamwork communication (Figure 5.7, top), using the speech activity and interaction data. Figure 5.7 (bottom) shows six sociogram-based proxies of social interaction for teams A, B, and C, over phases 1 and 2. This proxy resembles previous social proxies in online and collocated settings (Bachour et al., 2010; DiMicco et al., 2007; Erickson et al., 2002; Kim et al., 2008), using a typical undirected network representation. Each node in the proxy represents the nurses and the patient in the simulation. The size of each node represents the amount of speech activity, while the thickness of the edges denotes the number of verbal interactions (utterances) between people. This visualisation was generated automatically using an online visualisation tool[14]

HoC	Sub-constructs	Behavioural markers	Data/Analytics	Sensors	Visual representations
PATIENT-CENTRED CARE & TEAMWORK	Communication with the patient Teamwork communication	patient's speech activity nurses' speech activity turn taking among nurses and patient	overall time of patient's speech activity overall time of nurses speech activity speech transitions per nurse and patient	☒ array mic ☒ video camera	☒ network graph

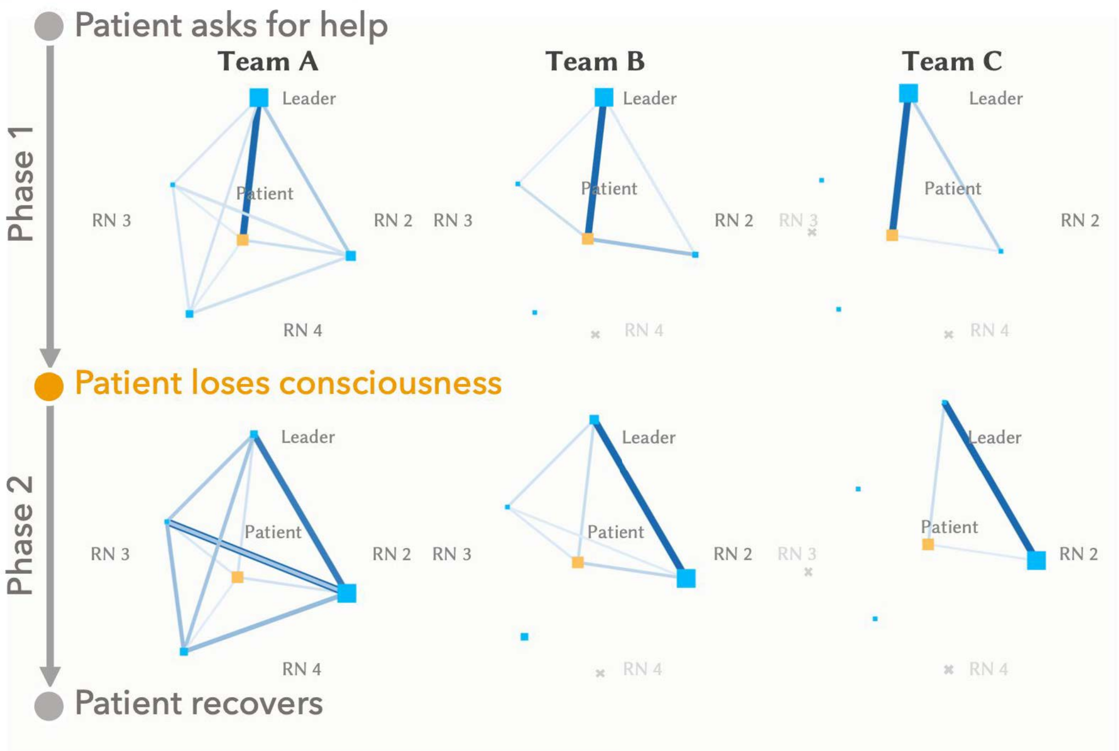


Figure 5.7: ***Top:*** **The upper table illustrates the requirements for designing a visualisation to represent patient-centred care and teamwork higher order constructs in relation to the verbal communication with the patient and teamwork communication.** ***Bottom:*** **Social proxies as visual representations. The orange node represents the patient and blue nodes the nurses. Edges represent verbal communication among these.**

5.5.1.2 *Observations on resulting proxies*

The social proxies of phase 1 (Figure 5.7, bottom) suggest that all three teams established patient-centred communication, with at least one nurse interacting with the patient in each team. Most of the communication was mediated by the team leader while other nurses remained almost silent (depicted by the small size of other nodes). The only exception is some conversation among RN2, 3 and 4 in team A. By contrast, once the patient lost consciousness, the communication dynamics in each team change completely. The proxies for phase 2 (Figure 5.7, bottom) show that RN2 (who was prescribed with the CPR lead role) dominated communication in all teams. Interestingly, members of team A were more communicative than those in team B. For instance, it can be noticed that all nurses from team A interacted with each other to some extent, while in team B, RN3 had less speech activity and interaction with other nurses. To illustrate this behaviour, two excerpts in teams A and B are shown in equivalent episodes during phase 2 in which team members had to coordinate clearing the airway, performing the CPR, and one nurse should hand over the CPR procedure to another nurse.

The first excerpt suggests that members of team A were coordinating their activity, verbally communicating what each will do next (a good practice in nursing). This is quantitatively presented in the proxy of team A (Figure 5.7, bottom-left) as edges between all nodes, with the thickest edges connecting RN2 with other nurses. RN3 did not count aloud during the CPR as depicted by the small node in the proxy. The second excerpt shows how only two nurses in team B talked to each other. From the videos, it was confirmed that the other nurse (RN3) was standing away from the bed, either waiting for instructions or just observing the situation. For the case of team C, given that there were only two nurses, it was expected to only see communication being led by the nurse performing the CPR (RN2).

Excerpt 1: Nurses in Team A communicating effectively	
1	RN2 ⇨ Leader: Put the head up.
2	Leader ⇨ RN2: one, two (giving oxygen to the patient)
3	RN2 ⇨ Everyone: I am going to do one more… twenty-one, twenty-two, twenty-three … *(doing CPR and counting aloud)*
4	RN2 ⇨ RN3: You take the next round please.
5	RN3 ⇨ RN2: Ok!
6	Leader ⇨ Everyone: one, two *(giving oxygen to the patient)*
7	RN4 ⇨ Everyone: Guys, I am going to start, I am going to do the defib now.
Excerpt 2: Nurses in Team B communicating less effectively	
1	Leader ⇨ RN2: I am going to check the airway.
2	RN2 ⇨ Leader: …and I will need this one *(pointing to the aging mask)* …so, should I start?
3	Leader ⇨ RN2: Yes!
4	RN2 ⇨ RN1: one, two *(doing CPR and counting aloud)* …twenty-nine, thirty
5	Leader ⇨ Everyone: one, two *(giving oxygen to the patient)*

To summarise, this proxy suitably segmented the data temporally (e.g. before/after a critical incident in order to expose changes), and highlighted behaviours at both individual and team levels which can help provoke reflection by participants or coaches.

5.5.2 Proxy 2: Localisation and proximity

This proxy aims to depict whether the embodied strategies that nurses enacted during the high-stakes scenario were *physically* patient-centric. The proxy integrates indoor localisation and proximity data encoded into meaningful zones columns in the multimodal matrix. Segments were also grouped into phases 1 and 2. The expectation from a "patient-centred care" perspective is for at least one nurse to remain most of the time in close proximity to the patient who is asking for help.

5.5.2.1 *Design*

From the design requirements elicited for representing patient-centred care higher order construct in relation to teams' embodied strategies (Figure 5.8, top), a physical localisation and proximity proxy was designed. Figure 5.8 (bottom) shows the physical localisation and proximity for the three teams in data collection 1. This proxy resembles a typical state diagram representation where the size of state (circle) represents the time that each nurse spent in each meaningful zone (e.g. the node *patient* indicated that the nurse was in very close proximity or on top of the patient). The edges represent the number of transitions from one zone to another. This proxy can be generated per individual or for the whole team. This visualisation was generated automatically using an online visualisation tool[15].

5.5.2.2 *Observations on resulting proxies*

The proxies of the teams during phase 1 (Figure 5.8, bottom) suggest that members of teams A and C were closer to the patient compared with team B. For team B, nurses were mostly *around* and further away from the patient, also showing more transitions between the *next* and *around* zones. Interestingly, the evidence from the social proxy (described above) suggests that, whilst all teams were assessing the patient's symptoms, nurses in team A were more engaged with the care of the patient by occupying a closer distance and communicating more with him and among themselves.

Nurses in teams B and C were closer to the patient than team A in phase 2 (see large orange nodes in Figure 5.8, bottom). Members of team A occupied the space *next* to the patient to a greater extent. The analysis of the videos explained this behaviour. For team A, RN2 and RN3 performed the CPR technique *next* to the patient (Figure 5.8, left), a suboptimal posture that may result in a poor CPR (Oh, Chee, Lim, Cho, & Kim, 2014). By contrast, teams B and C performed CPR on top of the patient or kneeling on the bed (Figure 5.9, centre and right), postures associated with better quality CPR.T

---

[15] www.kumu.io

This information was triangulated with the CPR information recorded by the manikin and included in the multimodal matrix. CPR scores from the three teams revealed that while none performed a good CPR in terms of compression depth and hand positions, the compression rate was appropriate for teams B and C (>100/min). The proxy of team C also shows a small node for the *bagging* zone in Phase 1, when one of the nurses should have been clearing the airway, as other teams did (see Figure 5.9, left and centre). This issue was raised by the teacher during the debrief.

In sum, this proxy can help to make visible how team members made use of the physical space. This can assist teachers to discuss in detail how certain clinical procedures were performed.

HoC	Sub-constructs	Behavioural markers	Data/Analytics	Sensors	Visual representations
PATIENT-CENTRED CARE	Communication with the patient	Embodied interaction	overall time spent with the patient overall time spent next to the patient, bagging area, around the patient and trolley zone	☒ proximity sensors	☒ network graph

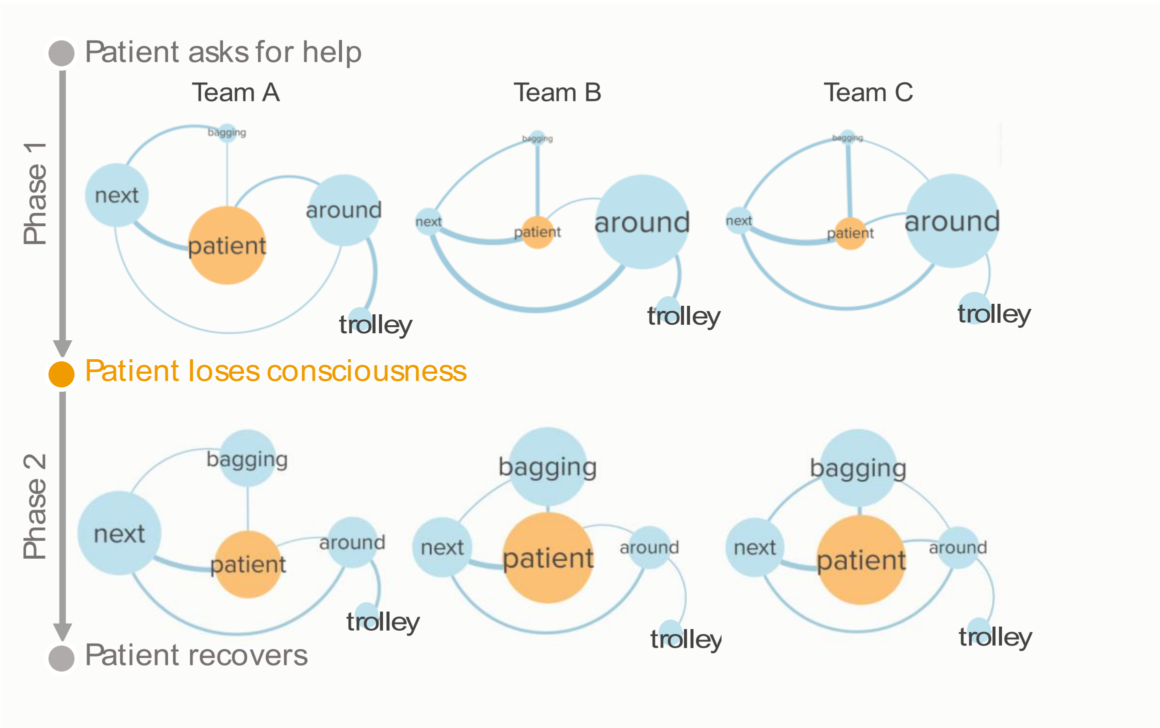


Figure 5.8: ***Up:*** **The upper table illustrates the requirements for designing a visualisation to represent patient-centred care higher order construct in relation to embodied strategies.** ***Bottom:*** **Physical proxies as visual representations. The orange node represents nurses' proximity with patient and light blue nodes other meaningful zones.**

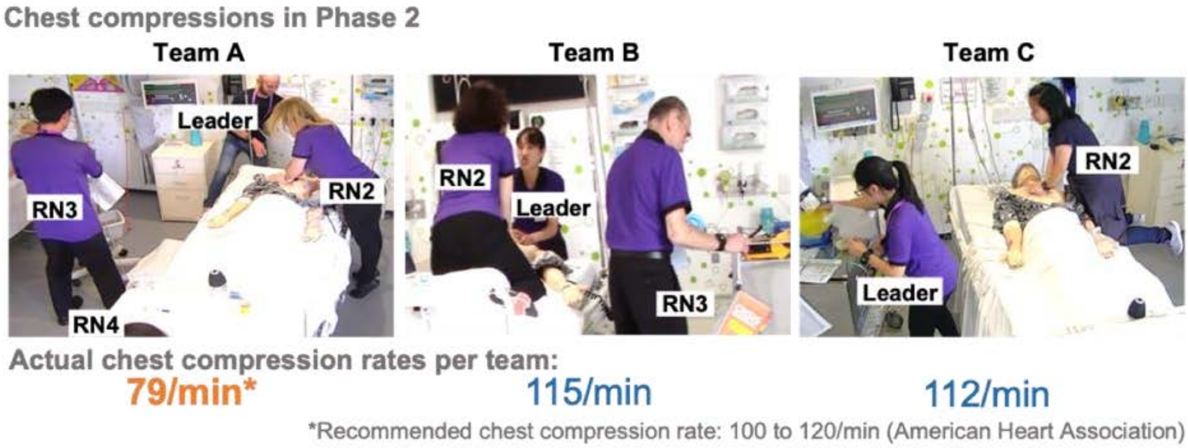


Figure 5.9: Three different ways in which nurses performed chest compressions: by the bed (team A), on top of the bed (team B), and on top of the manikin (team C).

5.5.3 Proxy 3: Physical activity and arousal peaks

This proxy aims to help nurses reflect on *affective* traits that they may have experienced, based on detected peaks in their EDA (Ruiz-Robledillo & Moya-Albiol, 2015), to consider potential coping strategies (Müller et al., 2011). The proxy draws on both EDA peaks and levels of physical activity, since high physical activity may decrease the reliability of EDA modelling (Pugh, Oldroyd, Ray, & Clark, 1966). Thus, by triangulating EDA with wristbands' accelerometer data, and whether the nurse was performing CPR (captured by the manikin) it is possible to attribute certain EDA peaks to intense activity, rather than to authentic sources of arousal.

5.5.3.1 *Design*

From the design requirements to represent patient-centred care and teamwork higher order constructs in relation to physical activity and stress (Figure 5.10, top), a physical activity and arousal proxy was designed. Figure 5.10 (bottom) shows exemplars of this proxy using data captured in Study 1. Each row relates to a team member (team leader and RN2-4). Orange dots represent the EDA peaks that may be associated with students experiencing arousal in certain moments during the simulation. The bar segments represent three levels of physical intensity (static, light and moderate) by different shades of blue (from light to darker). The two red bars indicate the beginning of phases 1 and 2. This visualisation was generated automatically using a Python script and a customised version of the Vis.js library [16]

[16] https://visjs.org

HoC	Sub-constructs	Behavioural markers	Data/Analytics	Sensors	Visual representations
PATIENT-CENTRED CARE & TEAMWORK	Active participation	Physical movement	static, light and moderate physical intensity from raw acceleration data	☒ accelerometer sensor ☒ electrodermal sensor	☒ timeline
	Stress management	Arousal changes	peak detection from skin conductance data		

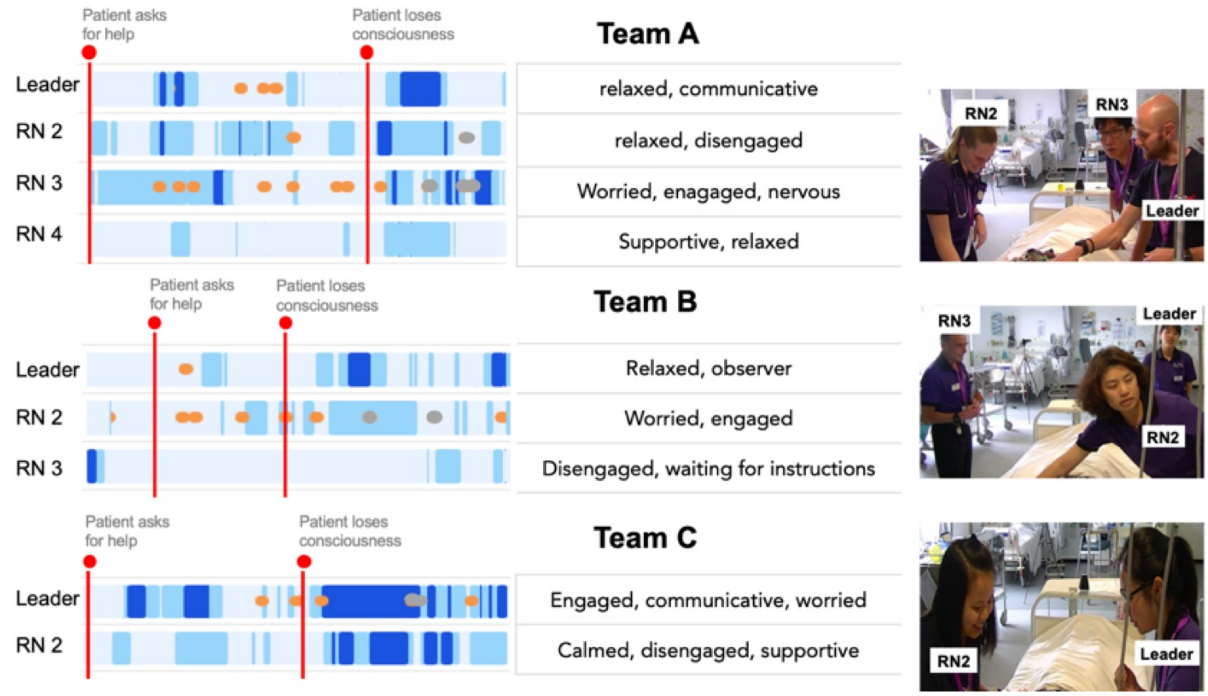


Figure 5.10: ***Top:*** **The upper table illustrates the requirements for designing a visualisation representing patient-centred care and teamwork higher order construct in relation to physical activity and stress.** ***Bottom:*** **physical intensity and arousal proxies. Physical intensity is reflected by different shades of blue segments in the timeline (darker = more intense). Arousal peaks are represented as orange and grey dots.**

5.5.3.2 *Observations on resulting proxies*

These proxies indicate that over the course of the whole simulation, nurses either experienced increased arousal, or very low or none. For example, in team A (Figure 5.10, top-left proxy) RN4 does not show any EDA peaks (orange dots), the team leader and RN2 show a few, but RN3 has many. A similar situation emerged in teams B and C where only one team member experience peaks (RN2 and the leader respectively). The measure of high physical activity (especially present in phase 2) seems to be helpful in accounting for peaks that may be influenced by increase motion during the CPR.

To explain the information provided by this proxy, the videos were qualitatively analysed. A researcher (the author of this thesis) took the timestamps where the arousal peaks were identified and used them to navigate through the video and annotate what students were doing and how engaged students were. This analysis revealed that the nurses with more EDA peaks displayed signs of *engagement, worry,* or *anticipation.* By contrast, nurses with fewer or zero peaks were either *calm* (e.g. team A – RN 4), *disengaged* (team B – RN3), or *laughing* with peers (team A – leader, RN2, RN4; team C – RN2). In team A (Figure 5.10, top-right), RN2 and leader nurses were chuckling, whilst RN 3 (the nurse with most EDA peaks in this team) looked focused. For team B (Figure 5.10, middle-right), the leader and RN3 appeared to be calm and relaxed; whilst RN2 looked very concentrated and worried about the patient. In team C (Figure 5.10, bottom-right), RN2 was chuckling, whereas the leader seemed to be focused.

In sum, the automated detection of EDA states, and its further filtering using information about physical activity to dismiss potential false positives, may be helpful during the debrief for students to reflect on their own performance, or for teachers to open a conversation about nurses coped with different challenges. This is further discussed in Section 7.5.2.

5.5.4 Proxy 4: Team timeline

This proxy aims to make visible the order and timing of the epistemic actions and procedures performed by nurses. It was deployed in the six authentic classroom sessions (data collection 2) for supporting post-hoc reflection. This proxy utilises the *actions* columns from the multimodal matrix.

5.5.4.1 *Design*

From the design requirements elicited in Section 4.5.2, a timeline proxy was generated (Figure 5.11, top), including the actions, the time when the actions were enacted and who performed each. Figure 5.11 (bottom) depicts an example of a timeline of team A. The timeline contains one line per role. The red lines delimit the beginning of phase 1 and phase 2, and each action performed by nurses appears as a blue dot. This visualisation was generated automatically using a Python script and a customised version of the Vis.js library[17].

[17] https://visjs.org

HoC	Sub-constructs	Behavioural markers	Data/Analytics	Sensors	Visual representations
PATIENT-CENTRED CARE & TEAMWORK	Coordination	Joint actions	Sequence of actions per nurse	☒ manual logging ☒ manikin data	☒ timeline
	Leadership	Task delegation	Actions per nurse		
	Time management	Responsiveness to patient's critical moments	Time the team took to attend the patient		

Figure 5.11: *Top*: The table illustrates the requirements for designing a visualization representing teamwork time management higher order construct. *Bottom*: The team timeline as an epistemic proxy, depicting each nursing student's actions for a team during the simulation.

5.5.4.2 *Observations on resulting proxies*

The team timeline depicted in Figure 5.11 (bottom) illustrated the actions performed by team A from data collection 1. It can be observed that RN1 had fewer actions during the first phase (i.e. patient ask for help) compared to RN2 and RN4. With no contextual knowledge this might make RN1 look rather passive, but in fact, since they are in the team leader role, this could be interpreted as the leader appropriately delegating actions to other RNs. This was validated by observing this behaviour in video recordings, where RN1 was asking for information, guiding and monitoring other nurses' actions. This information is further corroborated with a similar behaviour described in proxy 1 (RN1 engaging in lots of communication with the patient).

RN2 was attending to the patient most of the time during phase 1 (e.g. taking vital signs, giving medicine) while RN4 was documenting all of the patient's status and information given by other nurses. Regarding RN3, it can be observed in Figure 5.11 that just one action was performed during phase 1, complementing the information given by proxy 1 (less communicative) and proxy 3 (high arousal, worried, nervous).

Turning to phase 2, all RNs were actively participating by providing CPR to the patient who had just suffered a cardiac arrest. This demonstrates good coordination as all RNs performed joint actions. The information from the team timeline also suggests that this team responded quickly to the patient's critical status, taking less than two minutes to take the vital signs after the patient asked for help (phase 1) and less than thirty seconds to start compressions after the patient lost consciousness (phase 2).

5.6 VALIDATION WITH STUDENTS AND TEACHERS

In addition to the multimodal data collection 1 and 2, interviews were held with teachers and students to validate the collaboration proxies described above. Participants demographics, methods and analysis performed from the interviews are summarised in Table 5.1.

5.6.1 Participants

In addition to the multimodal data gathering for data collection 2 (see Section 3.1.1.2 for details), each team was invited to participate in a 30-minutes semi-structured follow-up interview a week after their class (week 4, winter 2018). Out of the 30 students, 23 students (3 males) voluntarily participated in these sessions in which they explored one of the collaboration proxies (proxy 4 presented in Section 5.5). Also, four teachers (T1, T2, T3 and T4) were invited to participate in a 60-minutes semi-structured interview to explore the four collaboration proxies. T1 teaches Medical-Surgical Nursing course; T2 is Lecturer and Simulation Manager in the Faculty of Health at UTS; T3 was coordinator of Medical-Surgical Nursing and is currently a Research Fellow; and T4 is the Subject Coordinator and teaches Medical-Surgical Nursing course. All teachers use simulation scenarios as part of their courses.

5.6.2 Methods

Interviews with teachers and students were conducted to understand how non-data experts envisage the use of these proxies to afford collaboration translucence (Zolyomi, Ross, Bhattacharya, Milne, & Munson, 2018).

Students participated in a 30-minutes semi-structured focus groups interviews, while each teacher participated in a 60-minutes semi-structured interview, both using a think aloud protocol (Hanington & Martin, 2012). The semi-structured interview was divided into two tasks:

- The *first task* consisted on the exploration of the proxies. Focus groups visualised and explored proxy 4 in a 60-inch TV screen connected to a laptop running Mac OS 10.13 and mounted in a conference meeting. Students were asked to recall their previous practice by exploring the information presented in proxy 4. Likewise, teachers were asked to explore proxies from all teams (A, B, C) using printed versions to get familiarised with nursing roles and to confirm or reject their initial expectations about team's performance.
- The *second task* corresponded to answer questions using three key concepts from social translucence: 1) *visibility*, in which students and teachers were expected to re-construct the

activity based on the proxy and students were asked if they could identify their own actions represented in it; 2) *awareness*, in which students were asked whether they could reflect their practice individually and as a group and teachers were asked if the proxies could be used for students to reflect on their practice; and 3) *accountability*, in which students and teachers were asked about who else should have access to the proxy and for what purposes.

Table 5.1. Study design overview: participants demographics, methods, proxies and analysis

Participants	Demographics	Method	Collaboration proxies	Analysis
Students	**n** = 23 (3 males) **avg. age**= 27.22 **stdv. age**= 8.69 **min. age**= 20 **max. age**= 49 **current year enrolled:** third	30-minutes semi-structured focus group interview: - initial exploration - questions about collaboration translucence -questions about usefulness	Data collection 2: • Team timeline	Collaboration translucence: • visibility • awareness • accountability
Teachers	**n** = 4 (1 male) **avg. years expertise**= 15 **stdv. years expertise** = 10.23 **min. years expertise**= 7 **max. years expertise**= 30	60-minutes semi-structured interview: - initial exploration - questions about collaboration translucence -questions about orchestration opportunities	Data collection 1: • Verbal communication • Localisation and proximity • Physical activity and arousal peaks Data collection 2: • Team timeline	

Table 5.1 shows students' and teachers' demographics information, the methodology followed during the interview with teachers and students, and the type of collaboration proxies that each of them were exposed. Video recordings of focus groups and teachers' discussions and were gathered and organised for subsequent analysis. All video recordings were transcribed using a transcription service.

5.6.3 Analysis

Statements of interest from teachers' and students' discussions, and actions from the exploration of proxies were grouped according to collaboration translucence key concepts, namely 1) visibility, 2) awareness, and 3) accountability (Niemantsverdriet, Broekhuijsen, van Essen, & Eggen, 2016) using NVivo Software (Table 5.1).

5.6.4 Results from focus groups with students

This section reports observations from the transcripts under the three principles of collaboration translucence: *visibility, awareness* and *accountability*.

5.6.4.1 *Visibility*

In all the six teams, students individually and collectively reconstructed the simulation following the timeline from left to right. This behaviour endorses the *visibility* property of the proxy as students could easily identify themselves by connecting the timeline events with their own experience. For example, one student stated the following: *[Alice] was the patient, you were the leader and I am RN2 because I was preparing the fluids"* (team 3, RN2). In some cases, the proxy helped students to recall and discuss actions they performed during the simulation. For example, a member of team 1 asked "*did we check the pulse?"* (RN3) whilst inspecting the timeline. Another student, pointing at the action "check pulse", replied: *"yeah, we did it"* (RN1).

5.6.4.2 *Awareness*

Students highlighted what they thought were correct and incorrect instances of teamwork performance (e.g. coordination, leadership, time management). Students in teams 1, 2 and 5 recognised they reacted fast to the most critical event (patient's allergic reaction) whilst students in four teams agreed that in future sessions, they should improve their reaction time (e.g. stop the IV fluids straight away). Students in teams 1, 2, 3 4 and 5 indicated that they were "*coordinated*" and "*working as a team*", by pointing to the timeline where similar actions were done, or actions were grouped. Seeing that they were all consulting documents at once, a student remarked: *"It's interesting to see that we were all looking for information at the same time. That might show that we discussed and worked as a team. If you would see that there's people that are looking for information at different time, that wouldn't make sense"* (team 1, RN3).

Some students in team 2 performed their own 'multimodal fusion', reading meaning into combinations of actions. For example, two students associated actions with roles across time, as follows: *"it seems like a lot was done in clumps, you [RN3] were talking to the patient, looking for information while others were doing the observations, that seems practical to me"* (the patient); and *"while RN4 and RN2 were doing the fluids I was staying with the patient. It is good to step back and look at what each person was doing, one thing at the same time, I think it shows you how your worked as a team"* (RN3). As another example, students in three teams reflected on their poorly developed leadership skills, realising they should have delegated more tasks to their colleagues. One team leader noted: *"I didn't delegate actions, when I was looking for the doctor"* (team 1, leader).

Interestingly, some students inferred a lack of communication skills, even though the timeline has no explicit communication events: *"two nurses were measuring the blood pressure (pointing to the timeline where those actions appeared) and then, after a while a third came, which means that there was no good communication*" (team 1, RN3). One might hypothesise that if the students had been able to access the sociogram visualisation for that moment (not available at the time of evaluation), this might have added further insights.

5.6.4.3 *Accountability*

Students' discussions revealed mixed preferences in terms of who should be able to see this proxy and for what purpose. A common view among all students was that they would like the teacher to guide the reflection using the timeline to confirm procedures, reinforce knowledge and suggest improvements. One student explained this as follows: *"Information needs to be confirmed by the teacher. She should confirm what we did, what should be done and what can we improve for the next practice"* (team 1, RN3). Students also agreed that all nurses in a team should be permitted to view and reflect on their performance. For example, one student reported that: *"If we are working together then it's good to [explore the timeline] as a group. If each look at it separately each will have their own stories."* (team 3, RN2). However, two divergent perspectives emerged when discussing timeline access to other teams. Students in four teams argued that sharing their timelines with others could leverage peer discussion, by comparing and contrasting their performance. By contrast, two teams raised concerns about how other groups would react i.e. judging, not taking the reflection seriously. One student said *"probably [another group] will laugh at this"* (team 1, RN2). Finally, all students agreed that sharing the information with the whole classroom immediately after the simulation would help them better reflect on their practice. Ethical questions around access to these kinds of analytics are addressed in the discussion (Sec. 8.5.2).

5.6.5 Results from teachers' interviews

Recall (Table 5.1) that while the students only had access to the timeline at the time of the focus groups, in these interviews with teachers, the other visual proxies were available.

5.6.5.1 *Visibility*

All teachers agreed that the visualisations served as summaries or proxies of activity that may not be easy to understand without having a dedicated teacher analysing the simulation (e.g. "*there may be things that the teacher picked up [during the simulation] but there may be things that the teacher didn't pick up. This is the value added by these visualisations",* T1) and to support reflection (e.g. "*[they] would help students reflect on their practice to see how they could improve it. When you are in the scenario it is very hard to think quickly and efficiently",* T3). All teachers also highlighted the need to go beyond these representations, for example, by "*combining the physical and the social proxies"* [...] "*to detect anomalies, such as assessing whether the nurse speaking more with the patient is the one that is further away"* (T2), which is not ideal.

5.6.5.2 *Awareness*

All teachers expressed the view that the timeline of actions (epistemic proxy) would assist them most in provoking reflection. T4 described her reasoning as follows: "*the timeline is the best* *visualisation to use because it would give people an indication of what they were doing at what time in the scenario, and time is critical, at least for this particular scenario".*

5.6.5.3 *Accountability*

Unanimously, all teachers expressed certain concerns about summative assessment of students and instead offered several ideas about how they would incorporate the proxies into guided self-reflection tasks. This was clearly summarised by T3 as follows: "*these data should be used for reflection. I think simulation in general should not be used for [summative] assessment. The whole point of simulation is to help people reflect and improve on their practice".* Teachers T2, T3 and T4 suggested they would like to use the proxies in the classroom, with some strategies suggested to preserve privacy, such as "*avoiding identifying people"* (T2), *"pixelating faces if video was shown"* (T2) or *"aggregating data from all teams"* (T3). Most concerns however were about singling out groups (e.g. students analysing data from other groups). T3 suggested that once the proxies become mainstream, there would be less privacy concerns, as follows: *"if the whole class is participating maybe we don't need written consent because it becomes part of the learning activity for everybody. Maybe we can have verbal consent before starting the reflection activity".* T1 added that showing visualisations "*does not go beyond what's currently done in nursing"* since it is a common practice for students to be recorded during simulations and reflect on pre-recorded scenarios.

5.7 SUMMARY

In order to address *RQ2 (How can low-level multimodal data streams be modelled to serve as proxies for group constructs?),* this chapter described the Multimodal Matrix approach for systematically modelling low-level multimodal data with qualitative insights defined as group constructs. Inspired by Quantitative Ethnography, the modelling approach is composed of four conceptual elements: dimensions of collaboration, multimodal observations, segments and stanzas. While the HCD-approach presented in the previous chapter addressed the definition of group constructs and formalization of MMLA requirements, the Multimodal Matrix approach took a step further by providing a computational model to map sensor data with group constructs.

The Multimodal Matrix was illustrated using multimodal data from two studies in the context of nursing simulation. In addition to designing and prototyping four collaboration proxies, a validation with teachers and students was reported to investigate their perceptions of whether the proxies have the potential to support awareness, accountability and visibility of team behaviours and processes. Findings from the interviews with students and teachers indicated that combining these proxies in one grouped visualisation could support a comprehensive understanding of team's behaviour. However, presenting all the information at once would increase the complexity of data visualisations and its understanding. Building on this promising feedback, Chapters 6 and 7 tackle the challenge of making multiple forms of multimodal data intelligible, by investigating *data* *storytelling* techniques to improve the communication of insights through the design of explanatory visualisations. In addition, a detailed discussion about the Multimodal Matrix implications and future work is presented in Chapter 8.

Chapter 6: Exploring the Potential of Explanatory Visualisations

This chapter motivates and introduces the *learning design-driven data storytelling approach* to the design and implementation of *explanatory visualisations*. Based on Information Visualisation (InfoVis) and Data Storytelling (DS) techniques and principles, this approach offers a way to explain student data by aligning educational visualisations with the intended learning design. This is a step towards the creation of visualisations that communicate *one story at a time*. A preliminary study to exemplify the applicability of data storytelling elements to visualisations is presented using visualisations from a collaborative database design activity context. Findings from this validation with potential users helped to establish the learning design-driven data storytelling approach. A second study illustrated the approach by implementing explanatory visualisations, which were validated with teachers. This validation explored the feasibility and the helpfulness of DS elements to drive visual attention to key aspects of the learning activity. Implications for further implementations are considered, along with a summary of findings[18].

6.1 INTRODUCTION

This chapter describes an approach to increase the *explanatory* effectiveness of the visualisations contained in learning analytics dashboards. Explanatory visualisations are those whose main goal is the presentation and communication of insights (Iliinsky & Steele, 2011). By contrast, exploratory visualisations are commonly targeted at experts in data analysis in search of insights from unfamiliar datasets. While educational researchers and developers are highly motivated, and suitably skilled, to invest time exploring visual analytics, there is a different value proposition to navigate for either students or teachers. Their time is limited, and they typically lack the skills to analyse data, or use more advanced visualisations. Explanatory visualisations may be more helpful, designed carefully to convey specific kinds of insights. Following this argument, the challenge for learning analytics researchers and developers is to discover and communicate insights rather than leaving students and teachers to play the role of data analysts, at the risk of gaining no insight.

[18] Parts of this chapter have been published in OzCHI (Echeverria, Martinez-Maldonado, & Shum, 2017); LAK (Echeverria, Martinez-Maldonado, Granda, et al., 2018) conference proceedings; and in the Journal of Learning Analytics (Echeverria, Martinez-Maldonado, Granda, et al., 2018).

This approach is aimed at providing a way to address some of the issues highlighted above by providing a Learning Design driven approach to bring data storytelling and general visual design principles into the design of visual learning analytics.

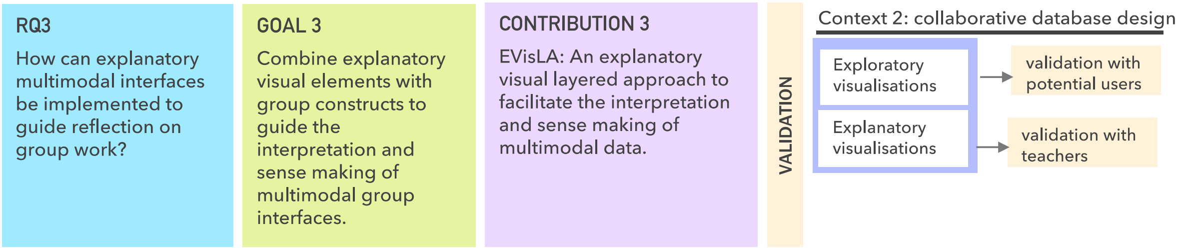


Figure 6.1: Research question, goal, contribution and validation methods addressed in Chapter 6.

Figure 6.1 summarises the research question, goal, contribution and the validation of prototypes addressed in this chapter. This chapter attempts to address the following question: *"How can explanatory multimodal interfaces be implemented to guide reflection on group work?* To this end, this chapter partially contributes to addressing this question by investigating data storytelling (DS) principles and techniques to design *explanatory visualisations* and guide reflection of group visualisations. A preliminary study first enabled the exploration of the potential of DS to align educational visualisations with the learning design. Findings from this study helped clarify the role of different DS elements and their usefulness to support sensemaking. Furthermore, the learning design-driven data storytelling approach was introduced by providing a way to align the learning design and visual elements to design explanatory visualisations that communicate one story at a time. Next, an illustrative study was carefully designed to operationalise the learning design-driven data storytelling approach and explore the potential and helpfulness of DS elements. As a result, a set of improved explanatory visualisations were designed and validated with teachers by triangulating quantitative (eye-tracking) and qualitative (think-a-loud protocol and semi-structured interviews) data. The visualisations used in this chapter corresponded to the ones generated by the DBCollab tool (see Section 3.1.2.2) in the context of a collaborative database design activity.

6.2 PRELIMINARY STUDY: TOWARDS DATA STORYTELLING TO SUPPORT TEACHING AND LEARNING[19]

This section explores how DS can be aligned into educational visualisations to support the sensemaking process. An initial study was designed to explore how DS principles (introduced in Section 2.5.4) could be employed in current visualisations generated by the DBCollab tool (Section 3.1.2.2). Thus, a new set of visualisations were designed endorsing DS principles and validated with potential users.

---

[19] Parts of this section have been published in OzCHI'17 (Echeverria, Martinez-Maldonado, & Shum, 2017).

6.2.1 The data storytelling principles and elements

Ryan (2016) and Knaflic (2015) proposed some practical *golden rules* or principles based on empirical knowledge and visualisation principles ((e.g. Tufte, 2001), to help professionals communicate more effectively with data. Five storytelling principles have distilled that have been reported by both as follows:

R1. Keep a clear goal: defining which is the purpose of the visualisation and who is the audience, could help the users to find insights more effectively.

R2. Eliminate extra-ink: remove elements that do not add informative value.

R3. Using narratives wisely: labels should be used wisely and should be descriptive to explain important points.

R4. Driving attention: push everything to the background and then use pre-attentive attributes to only highlight important aspects of the visualisation.

R5. Call for action: if the visualisation is clear and concise, it should explain the "story" that users should take from it.

Whilst these principles can be applicable in different ways almost using any type of visualisation, a set of visual Data Storytelling design elements (DS elements) that can be related to each rule are proposed for this work. Table 6.1 shows each rule with the corresponding actions that can be performed directly on the visualisations. R1has been omitted because this rule cannot be translated as a specific DS element but rather can be pervasively considered in the way the visualisation is deployed as a whole.

Table 6.1: Data storytelling (DS) principles with their corresponding DS elements.

Data Storytelling principles	**Data Storytelling elements**
R2. Eliminate extra-ink	**(A) Decluttering** (*removing*): Data labels Legends Data markers Tick marks Grids Axis labels
R3. Using narratives wisely	*Adding*: **(B) Narratives** to critical data points **(C) Shaded areas** to cluster information
R4. Driving attention	**(D)** Highlighting with **colours** **(E)** Emphasising key **data points** **(F)** Making **lines** thicker
R5. Call for action	*Adding*: **(G)** Prescriptive **title** delivering a straightforward insight from the data

For the other four rules (R2-5) their corresponding DS elements were defined as follows. Eliminating extra-ink (R2), can be achieved by (A) *decluttering* the visualisation. This means *removing* or *sending to the background* unnecessary data labels (text in each point), data markers, the grid in the background, legends and tick marks. It also involves rotating labels (horizontally) to improve readability. This could be considered a pre-step, before *adding* other DS elements. Narratives (R3) can be used to further explain certain aspects of the visualisation. This can be achieved by adding (B) textual *narratives,* to critical parts of the visualisation; or (C) *shaded areas,* to cluster information. Driving attention (R4) can be supported by first, sending everything to the background (making all visual elements into grayscale as part of the decluttering process) and then, highlighting only the essential parts of the visualisation that help to tell the story. Driving attention could be achieved by (D) *highlighting data with colours*, (E) *emphasising key points* or (F) *making lines thicker*. Finally, to communicate a call for action (R5), a (G) prescriptive *title can deliver a straightforward insight from the data.*

6.2.2 Validation with potential users

This section describes the preliminary study conducted with potential users to gain first insights and impressions of the role of DS in educational visualisations.

6.2.2.1 *Study setup*

A study was designed to expose potential users to a number of explanatory visualisations with data storytelling elements (VDS) and original exploratory visualisations (OV) shown previously to the students (Echeverria, Martinez-Maldonado, Chiluiza, et al., 2017). The aim of this study is to generate an understanding of the potential role that data storytelling elements can play to support interpretations of visual learning analytics and dashboards. This study aimed to obtain first impressions and insights about how potential users reacted to visualisations with and without DS elements by examining two aims:

A1. **DS sensemaking support**. The aim is to investigate if participants could articulate the right story while looking at the visualisations and if DS elements supported the understanding of the visualisation.

A2. **DS helpfulness.** The aim is to explore which data storytelling elements are most helpful to support storytelling through learning visualisations.

6.2.2.2 *Participants*

Five participants (four males, one female, avg. age: 31, std. dev.:5.36) were asked to participate in an exploratory study. All participants had previous experience in teaching and database design. Two had experience on teaching database systems. Four were PhD students and one holds a PhD.

6.2.2.3 *Materials*

The data used for the generation of visualisations was collected from the DBCollab tool that was used to support five teams in a collaborative database design activity (see Section 3.1.2.2.). For this preliminary study, two teams (A and B) were selected because they featured quite distinctive stories

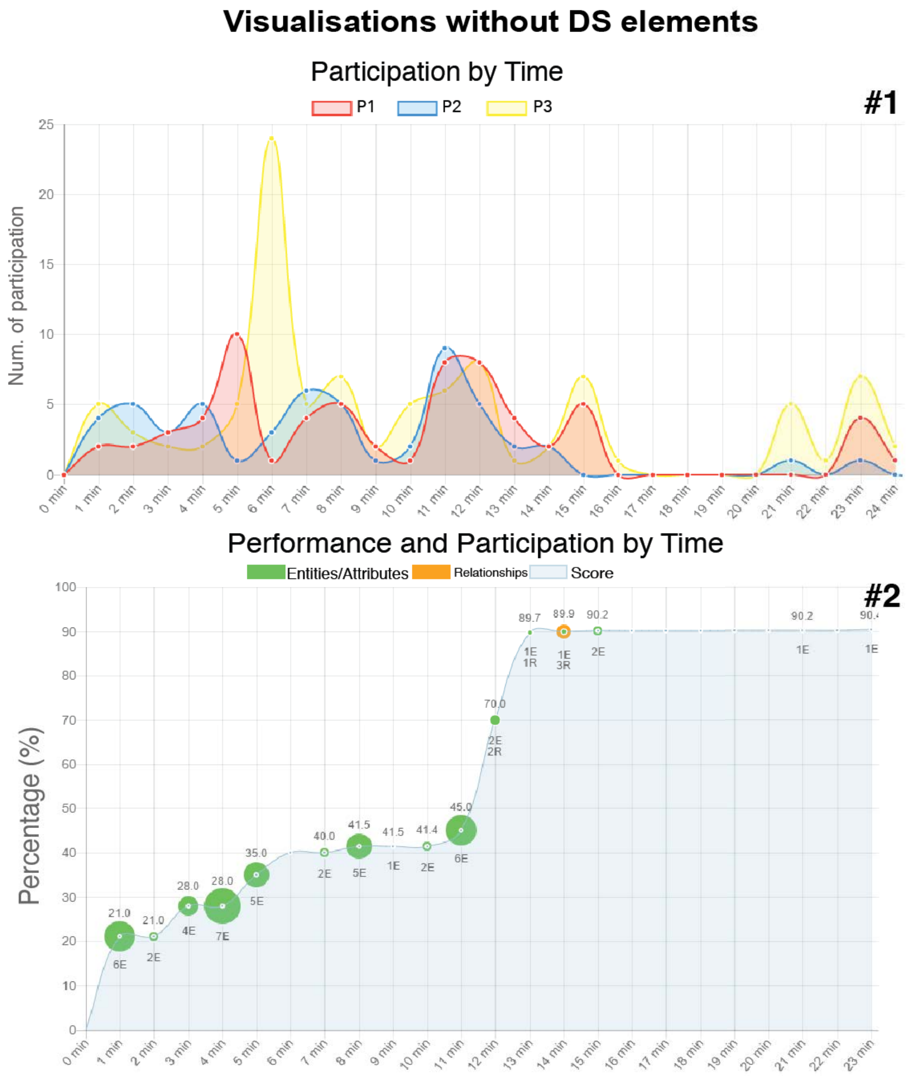


Figure 6.2: Participation (#1) and Performance (#2) visualisations of a group working in a database design activity using the DBCollab tool (Echeverria, Martinez-Maldonado, Chiluiza, & Buckingham Shum, 2017).

in terms of participation and performance, as follows. Participation in Team A was balanced, with one participant leading the activity. Participation in Team B was unbalanced, with each participant dominating the participation at different times. The Performance of Team A was high, as the team was able to determine all key aspects of the task. The Performance of Team B was low, as the team was only able to determine few key aspects of the task. Examples of these are shown in Figure 6.2. Figure 6.2 - #1 shows the individual actions (participation) each student within the team performed during the activity using the collaborative tool. Figure 6.2 - #2 depicts the team performance and individual participation over time.

Having in mind these clear *stories*, a total of ***four*** prototype visualisations with DS elements (Figure 6.3) were crafted using a graphic editor software. Figure 6.3 shows an example of *team participation* and *team performance* visualisations with DS elements. The visualisation about team participation (Figure 6.3 - #3) endorsed A, B, C, D and G DS elements, while the visualisation about team performance endorsed A, B, C, E, F and G DS elements (as listed in Figure 6.3 - #4). The participation visualisation was decluttered by removing unnecessary grid lines and pushing axes and ticks, and team members' data points in the background (A), text narratives (e.g. *"P1 and P3 had coordinated contributions*") were added into critical data points (B), a shaded yellow area was added to indicate a change in the visualisation data points according to a reflection activity planned in the learning design (C), a bright colour (orange) was assigned to P3 to highlight her participation during the activity (D), and a prescriptive title *"P3 lead the team activity*" was added to deliver the

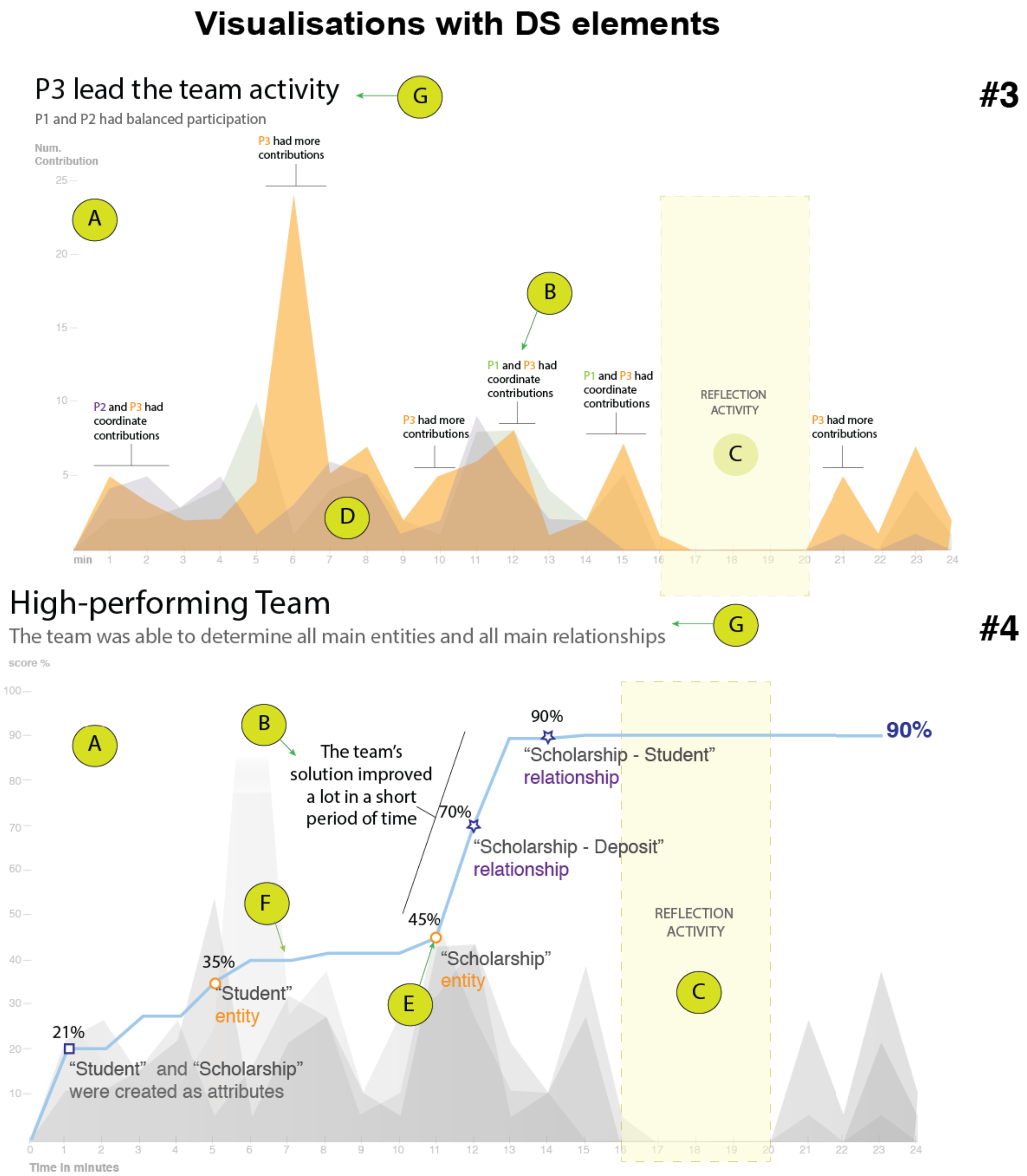


Figure 6.3: Prototypes design of participation (#3) and performance (#4) visualisations with DS elements.

main insight of the visualisation (G).

Similarly, the *performance visualisation* (Figure 6.3 - #4) was decluttered (A), narratives (e.g. *"the team solution improved a lot in a short period of time"*) were added to indicate critical changes in relation to teams' performance (B), critical data points were emphasised with colours and shapes (E), a ticker blue line was added to visualise performance trending clearly (F), and the prescriptive title *"High performing team"* was added to deliver the main insight of the visualisation (G).

In short, each participant would inspect ***eight*** paper prototypes visualisations from two teams, regarding participation and performance, through visualisations with and without data storytelling elements.

6.2.2.4 *Method*

A semi-structured interview was designed to explore the perceptions of participants towards DS elements. The interview consisted of the following tasks:

- *Task 1:* participants were asked to explore the set of paper prototypes visualisations about participation and performance with and without the DS elements, for Teams A and B (8 visualisations in total: 4 without DS elements, and 4 with DS elements) gradually. For example, Figure 6.2 and Figure 6.3 depict all the visualisations corresponding to Team A. Participants were encouraged to externalise their reactions using a think-aloud protocol and try to "tell the story" from each visualisation (**A1. DS sensemaking support**). Although the number of participants in the study is small, the sequential order to which visualisations were presented varied to minimise the learning effect.

- *Task 2:* participants were asked to explain how each DS element was helpful to support the 'story'. Additionally, participants rated each DS element (from 1 to 5 - being 1 the most helpful and 5 the least helpful) illustrated in the visualisations with DS elements (**A2. DS helpfulness**).

Each interview lasted about 30 minutes. All interviews were audio recorded and notes were taken during the interview by the PhD student.

6.2.3 Analysis and results

Audio-recordings were transcribed. In order to understand how DS elements were used to support the interpretation of the story, transcriptions and notes taken from the comments made by each participant were analysed.

To address A1, first, statements were analysed to look if all participants articulated the story behind each type of visualisation, i.e. participation and performance visualisations with and without DS elements. Second, statements that expressed sensemaking support in relation to DS elements were selected and grouped per visualisation type.

Next, to tackle A2, two analyses were carried out: first, ratings from DS elements were gathered and ordered and, second, statements from interviews were analysed thematically. The *themes* corresponded to the DS elements (from A to G). Thus, statements indicating how a DS was helpful to support storytelling were selected and grouped according to the DS element.

The next section presents the results in terms of the effectiveness of the DS elements to support sensemaking (A1) and the perceived contribution of each DS element to support the story (A2).

6.2.3.1 *Sensemaking support*

Considering that each visualisation had its own story, the results from participants' reactions when interpreting the story from each type of visualisation are summarised.

Visualisations about participation. All participants identified the stories from visualisations with and without DS elements. From the former, two participants recognised the *title* as a critical element to understand the 'story' effectively. Nevertheless, a criticism was that the orange *colour* displayed at the first plane dominated the visualisation. They expressed that the colours in the visualisation without DS elements, were better and the opacity helped to observe overlapping participations at a glance. However, one participant indicated that, "*while the information could be easily explored in the visualisation without DS, the [visualisation with DS] helped me to focus only on the specific and critical moments of the team participation*".

Visualisations about performance. It was difficult for participants to infer the story from the visualisations without DS elements. Only one participant commented that the visualisation was about performance. From the visualisation with DS elements, four participants correctly recognised the stories. For instance, participants stated that the information presented in the version with DS elements "*was clearer and consistent*".

Reflecting upon these results, it can be noted that, when a lot of information is displayed in the visualisation, users cannot determine effectively the story behind it due to the extreme cognitive load process for interpreting it. Another cause could be the multiple stories the visualisation tells. The DS elements added to this visualisation attempted to guide the user to focus on *one story at a time*. This proved effective for the visualisations about *performance*, most likely because there was one main story to tell. However, in regard to visualising *participation*, although the DS elements helped participants to focus on only one story at a time, they also wanted to keep other possible 'data stories' available for exploration. This is further addressed in Section 7.4 and discussed in Section 8.4.

6.2.3.2 *Helpfulness of Data storytelling elements*

The next step was to understand how helpful each particular DS element was. Figure 6.4 summarises the ratings from the five participants according to the helpfulness of each DS. Results are presented from the most to the least helpful element for each visualisation.

Visualisations about participation

The DS element *(B) narratives* was the most helpful element to support the story (see Figure 6.4, left). Participants agreed that this element helped them understand the data points from the visualisation by offering brief information of the things that happened. One participant expressed this as following: "*narratives are descriptive. I could understand what happened here [pointing to the narrative and the data]*". The second most helpful element was *(G) title delivering a straightforward insight from the data*. Participants mentioned that *"this [title] explains the main idea"*, and *"describes the things that are shown in the visualisation."* Furthermore, *(D) highlighting with colour* was rated as the third most helpful element. Participants said that this element makes it easier to identify what student participated the most. A participant said: *"this is important if your intention is to emphasise this student"*. However, some participants agreed that as the colour is brighter, it is difficult to see the overlapping participation by other students. For example, one participant expressed that "*it is harder to observe overlapped participations*". Finally, *(A) Decluttering* and adding a *(C) shaded area to cluster information* were rated as the least helpful elements. In brief, participants agreed that the DS elements helped to make the story behind each visualisation clearer to some extent. However, the colours and opacity were judged as detrimental to understanding the participation of the three students. For instance, one participant said that: *"I would rather prefer to use the colours of the original visualisation"*.

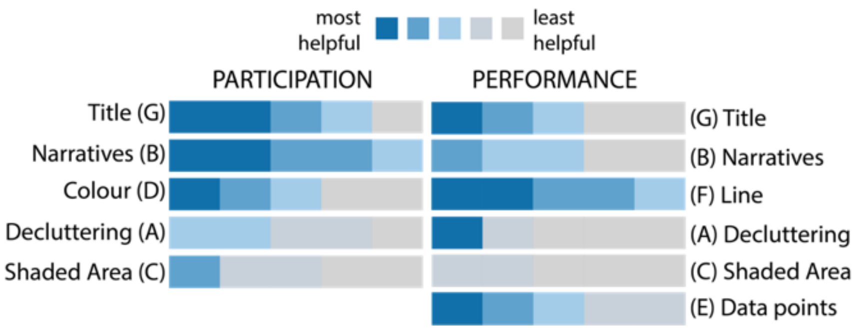


Figure 6.4: Ranking from each DS element for both, participation and performance visualisations.

Visualisations about performance

As depicted in Figure 6.4 (right), *(F) making lines thicker* was catalogued as the most helpful element. Participants expressed that, because of the thicker line, they could understand the main idea of the visualisation. One participant expressed that: *"The line is clear. This [pointing to the line] is more readable"*. Next, *(E) emphasising key data points* was categorised as the second most helpful element. Participants stated that showing this element helped them focus on important aspects of the task that occurred in the activity. One participant expressed this as follows: *"these [key points] summarise the most important things of the task which allows me to focus only on this"*. The third most important element was the *(G) title delivering a straightforward insight from the data*. Some participants expressed that the title contains the main idea of the visualisation. Conversely, other participants expressed that the title *"was not crucial for supporting the message of the visualisation, because the other elements made the visualisation easy to understand [already]"*. Furthermore, *(B) narratives* was rated as somewhat helpful. Some participants said that these annotations *"support the main idea", "describe briefly what happened"* and *"show what the teacher wants to know"*. By contrast, participants who rated this element as less important indicated that annotations *"cause noise to the visualisation"* and *"did not add any extra information to the story"*. Again, *(A) Decluttering* and *(C) shaded area* were rated as the least helpful to supporting the story.

From these results, it can be noted that, as each visualisation has its own story, each DS element can play different roles in the sensemaking process. Participants recognised the pre-attentive attributes (highlight with colour, key points, lines) as the most helpful elements to interpret the message. By contrast, decluttering and the shaded area did not seem to add relevant information to the story but may have helped to make other elements outstand.

6.2.4 Reflections and implications for the next iteration design

This study explored the potential of a set of data storytelling elements to leverage users' sensemaking process. The prototypes of visualisations with DS elements were generated and presented to participants to provoke reflection and record their reactions. Preliminary results indicate that DS elements that permit exploration of each 'data story' in turn could add clarity, especially when there are multiple possible stories in a complex visualisation. Participants also suggested that *additional stories should be available for exploration*. Moreover, visual design choices (e.g. text and line colours) should be carefully considered to avoid deviation of users' attention to less relevant information. Finally, participants mentioned that *interactivity* should be considered for the next design, as this could be helpful for guiding the interpretation of complex visualisations.

It is worth noting that the intention of adding DS elements is not to claim that these visualisations are better than others in all respects. Instead, the aim was to explore if DS elements help to influence the sensemaking process and by guiding users to focus on different aspects of the 'data story'.

Motivated by this initial exploration of DS elements in an educational context, the next sections describe a learning design-driven data storytelling approach to link DS elements with contextual knowledge, by drawing on the initial set of principles and DS elements, as presented above in Section 6.2.1.

6.3 LEARNING DESIGN-DRIVEN DATA STORYTELLING APPROACH[20]

When learners or teachers face a dashboard or visualisation, Verbert et al. (2013) propose that the first stage in the "awareness-sensemaking" process is: (i) **visualising and presenting the data** to the user. Next, since "*data in themselves are not very useful*", the second stage expects that users will (ii) **formulate questions** and assess how data is useful or relevant for addressing those questions. The final two stages are concerned with the sensemaking process by which users (iii) **respond to those questions** (e.g. creation of new insights) in order to (iv) **perform actions** accordingly (e.g. for educators to modify the learning context, or students their behaviours). This process then continues iteratively in what Verbert and colleagues call the "learning analytics process model", summarised in Figure 6.5 (left).

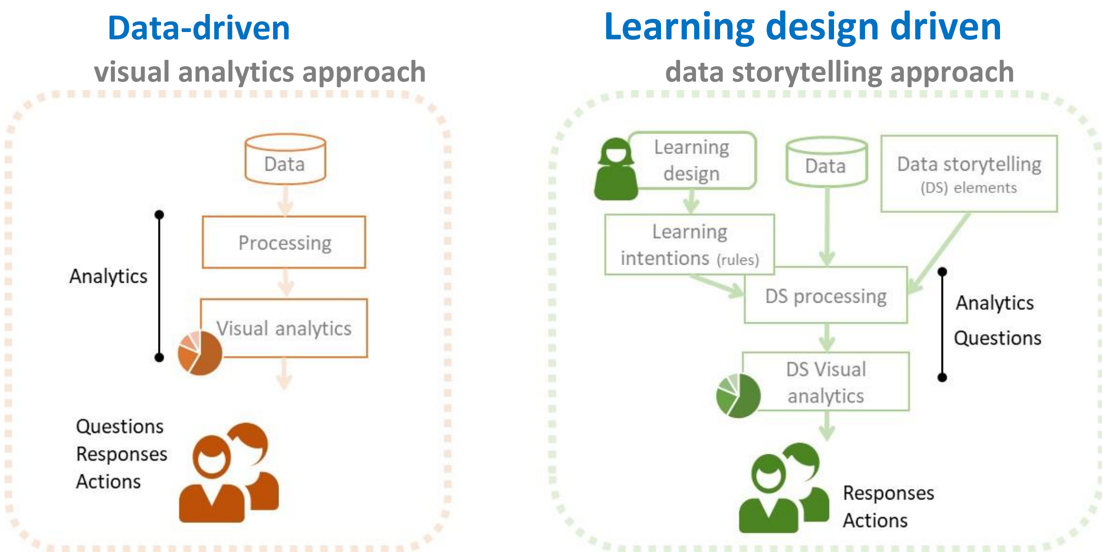


Figure 6.5: ***Left***: **Conventional data-driven visual analytic approach.** ***Right***: **Learning design driven data storytelling approach to support sensemaking.**

Figure 6.5 (left) summarises this common-sense, data-driven approach to visualising datasets. For clarity, in this diagram, there is a distinction between (raw) learner data (presented as data) and the analytics that are needed to present distilled views of these data for human consumption. These analytics include the *Processing* performed on the raw data (e.g. aggregating functions, machine learning, prediction models), and the *Visual analytics* that render the data. In short, the user drives the process of formulating questions, generating insights from the analytics (or 'data' according to Verbert et al., 2013) to respond to those questions, and acting.

In this scenario, the designer of the learning analytics has the challenge of generating interfaces to cater for a number of possible questions that learners or educators could have. This is the conventional data visualisation approach used in most learning analytics dashboards (Bodily &

[20] Parts of this section have been published in LAK'18 (Echeverria, Martinez-Maldonado, Granda, et al., 2018) and the Journal of Learning Analytics (Echeverria, Martinez-Maldonado, Granda, et al., 2018).

Verbert, 2017; Sveen, Klerkx, & Duval, 2014). Teasley (2017) has referred to this approach as "one-size-fits-all" dashboard design. Notable exceptions are visualisations that incorporate mechanisms that allow user input such as customisable dashboards (e.g. see initial explorations in Roberts, Howell, & Seaman, 2017), and personalised visualisations that automatically adapt according to individual's personal traits (Lallé, Conati, & Carenini, 2016). These initial attempts are still in the research lab and have not yet been adopted by the learning analytics community. Teasley (2017) has suggested that personalised visualisations may help address some of the already identified negative effects of dashboards and visual analytics that have been reported.

The learning design-driven data storytelling approach, that this chapter proposes, sits in between "one-size-fits-all" dashboards (simple to produce) and adaptive personalisation (complex to produce and train). Figure 6.5 (right) depicts the approach, which attempts to scaffold the formulation of Verbert et al.'s *user questions* by enabling the dynamic modification of critical features of the visualisations. The aim is to achieve this by translating the learning intentions of the educator into data storytelling elements that highlight particular aspects of the data. Figure 6.5 (right) shows three boxes aligned at the top as the foundations of the learning design-driven data storytelling approach. These are the teacher's learning design, the learner's data, and the data storytelling elements. This approach considers that a teacher's pedagogical intentions, explicitly or implicitly specified in the learning design, can be translated into **rules** which can be read by a data processing system. Once student data has been collected, these rules can be used to reveal interesting behaviours or indicators in student activity to which it is desirable to draw the educator's or learner's attention (e.g. when a student has low performance in an activity because she could not complete a correct step, or when a certain performance threshold is reached).

Variations in data can, to some extent, be detected automatically with no domain knowledge. Progress in this area has recently been made in InfoVis research, to automatically annotate data points of interest in line charts, stacked graphs and alluvial diagrams (Bryan, Ma, & Woodring, 2017); and indoor maps (Metoyer, Zhi, Janczuk, & Scheirer, 2018). However, although specific properties of lines/curves (e.g. peaks and troughs, intersections, gradients), and time-series data (e.g. three consecutive low-performance episodes in a particular session) can be used by the system to highlight a multitude of changes in a graph, from an educational point of view, these may still be irrelevant as feedback to aid learners.

The addition of *learning design* information provides the basis for more powerful filters, since the educator can define precisely what the *meaningful* features are. For example, perhaps the first two of the students' three "low performance episodes" are entirely expected by the educator and part of the learning design, so the dashboard should only highlight the third event as a salient indicator. In terms of Verbert et al.'s model, explicitly directing the user's attention to interesting features at a glance should help them focus on questions that are *relevant* for learning/teaching, since they have been identified and curated in the learning design. The goal is a more productive sensemaking process for users to respond to those questions and perform actions.

In summary, whilst the data storytelling elements can play a critical role for emphasising or de-emphasising visual elements of a dashboard, the learning design (materialised as rules) can provide the contextual information to identify *which* patterns need to be highlighted or sent to the background, driven by the educator's intention. This way, the proposed learning design-driven data storytelling approach aims to keep humans 'in-the-loop', not just as consumers of the learning analytics, but also to shape what is being shown in the interface.

6.4 ILLUSTRATIVE STUDY: EXPLORATORY VS. EXPLANATORY VISUALISATIONS

This section presents an illustrative study to explore the feasibility of the learning design-driven

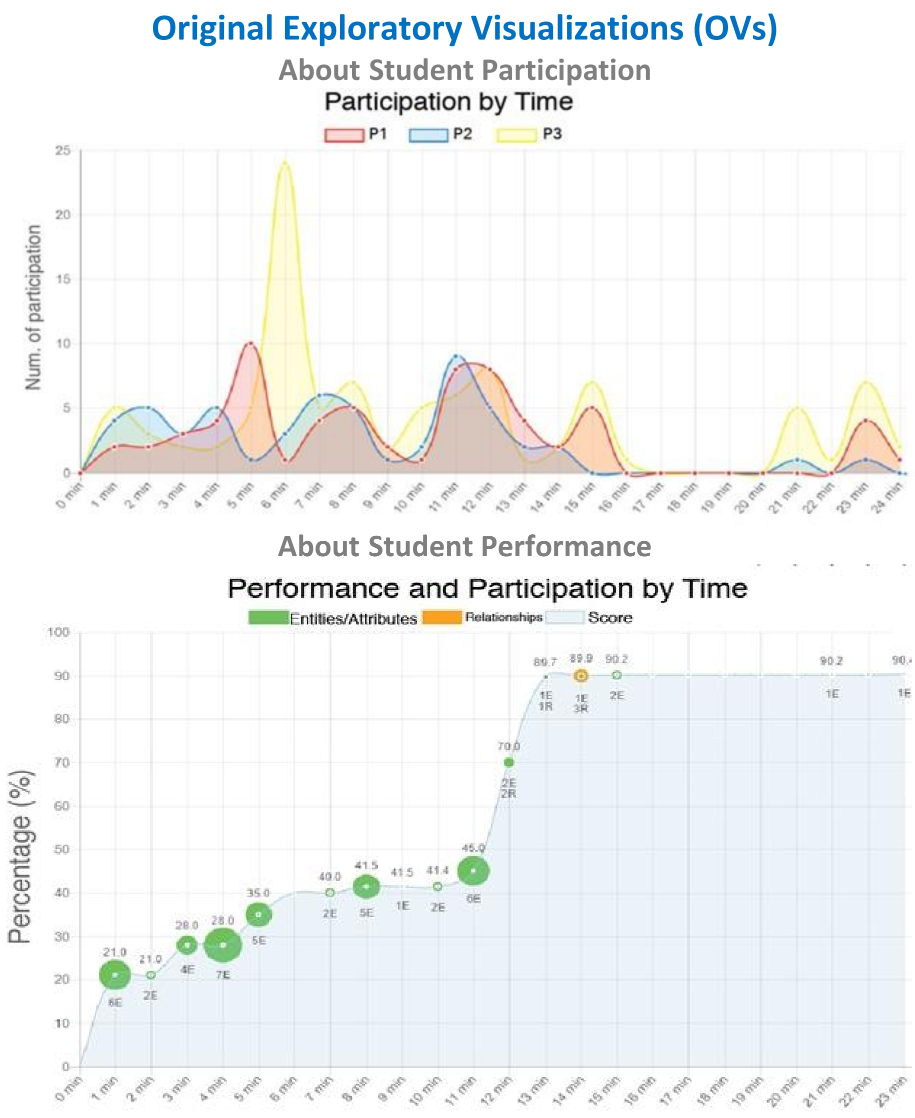


Figure 6.6: Original exploratory visualizations generated by the DBCollab tool and fed back to a student team working collaboratively in a database design activity. *Top*: participation of each team member. *Bottom*: evolution of the team's performance.

data storytelling approach to generate explanatory visualisations. This study was designed and conducted in the context of collaborative database design, as described in Section 3.1.2.2. Similar to the previous preliminary study presented above and as explained in Section 6.2.2, the data collected and visualisations generated by the DBCollab tool from two teams (A and B) were used in this study (referred to from now as *original visualisations*, e.g. Figure 6.6). The section is divided into two sub-sections that present the following:

1. The design rationale of hand-crafted data storytelling-enhanced versions of the original visualisations; and
2. An evaluation study with teachers, to understand the role played by the data storytelling elements to support sensemaking.

6.4.1 Designing explanatory visualisations

This subsection describes the design process to add DS elements to the original visualisations shown to students. Note that although in this study the visualisations were hand-crafted, the purpose is to provide a detailed explanation of how this process can be automated.

Step 1: Understanding the context. Working closely with stakeholders (i.e. teachers and students) is key for contextualising a visualisation. Thus, the context of the visualisation is shaped by the teacher's intentions and the learning design of the activity. The teacher selected a collaborative activity to support collaboration skills. The teacher wanted to provide particular information to students/groups in order to improve their awareness about collaboration skills. Also, she wanted to know the how students' ER diagrams evolved, determining missing concepts and relationships throughout the design process.

Step 2: Mapping teacher's intentions with rules. This second step consisted in extracting the teacher's intentions to be represented through DS elements. For instance, the teacher explicitly stated that she had two critical learning intentions, which were also the drivers of the design of the two visualisations generated for the students:

- *Learning intention 1 (LI1 – Participation):* All students within a group should ideally participate equally in the collaborative database design.
- *Learning intention 2 (LI2 – Performance):* Some specific database elements are critical and should be included early in the database design. This can directly impact the quality of the final design.

These learning intentions enabled the coding of rules in order to automate the generation of new visualisations on-the-fly by translating teacher's learning intentions into DS elements. For instance, a high-level rule **(R1)** can be created for LI1 as follows: *"if a student is dominating the activity in a pre-defined time slot (e.g. for 5 minutes), show [personalised message]"*. For the case of *LI2*, the following rule **(R2)** was generated: *"if [name of a key entity or relationship] is added into the solution, add a marker indicating the name of the entity (or relationship)"*.

Table 6.2: Data storytelling (DS) design elements considered for this study.

DS Elements	Description
A- prescriptive title	a succinct message that is intended to explicitly communicate the main message of the visualisation
B- highlighting specific data series	visual element that can help the audience to focus on the data that is relevant to the current story
C- removing unnecessary data points	removing irrelevant data points and keeping only relevant ones makes explicit which data support the claim in the title
D- decluttering	removing grids, eliminating indirect legends, reducing the number of colours used result in minimising distractions
E- narrative text	text in the form of labels can help to gain a better understanding of or explain interesting changes in the data
F- shaded area	enclose those data points associated with the same insight

Step 3: Translating rules into DS elements. From the list of DS elements presented in Table 6.2, a set of rules was designed to encode the learning design intentions with visual elements into a new set of visualisations that incorporate the foundations of data storytelling. Table 6.3 shows how some rules were used to encode each learning intention, focused on Participation (rules 1 and 2) and Performance (rules 3 and 4) respectively.

> ***Rule 1:*** if the teacher's intention is to know the group's balanced participation, the rule will search for patterns in the graph where this behaviour is explicit: if a student was dominating (i.e. Participation of one student is higher than the other two students) by x minutes (x being defined by the teacher), a visual indicator could be generated on the visualisation itself highlighting this behaviour (see *if-then rule* in Table 6.3, row 1). For example, visual elements can be added to the graph such as a text label with a narrative or triggering a visual alarm.
>
> ***Rule 2:*** If the teacher's intention is to know the overall team Participation, the rule will determine if all students featured a balanced Participation (e.g. using an entropy function). A message could be set to explain this behaviour (see *if-then rule* in Table 6.3, row 2). Then, adding the resulting message as a big title in the graph could be a way to depict this behaviour.
>
> ***Rule 3:*** If the teacher's intention is to identify when a database element that makes the task Performance increase is added (being this element also defined by the teacher), then the rule can insert a key data point when a student performed this action (see *if-then rule* in Table 6.3, row 3). Visual elements that support this rule include shapes and colours to differentiate

different type of elements, data labels to emphasise current score and text labels to describe the name of the entity.

Table 6.3: Cases illustrating how to map intentions, rules and data storytelling elements.

Visualisation	Intentions	Rules	Data storytelling elements	Example

LI1: Participation	Highlight whether the group is contributing equally or not	Rule 1:
		`if dominating (P1,P2,P3) == 'P1' in slot_time(x):`
		`show_message = "P1 dominated the activity during x minutes"`
		`highlight_series (P1)`
			narrative text; highlighting data series	Figure 6.7 (B)

	Summarise the overall group contribution	Rule 2:
		`if balanced (P1,P2,P3) == 1:`
		`show_message = "All participants had a balanced Participation through the activity"`
			prescriptive title; narrative text;	Figure 6.7 (A)

LI2: Performance	Highlight elements that increase the design's quality	Rule 3:
		`if entity == 'Scholarship':`
		`highlight_point (entity)`
		`show_score (line)`
		`message = "Entity Scholarship"`
			narrative text; highlighting data point; shaded area	Figure 6.7 (D)

	Summarise the overall group performance	Rule 4:
		`if overall_performance (team) >= 75:`
		`show_message = "High-performing team"`
			prescriptive title; narrative text	Figure 6.7 (A)

Rule 4: If the teacher's intention is to visualise the overall team Performance and assess if this score corresponds to low or high Performance, a rule can be formulated to determine this depending on the team's final score and a *x threshold* (x being defined by teacher's expertise, see *if-then rule* in Table 6.3, row 4). The resulting message of the rule can be described in the title of the graph or by adding narrative to the graph.

Step 4: Crafting the visualisations. The final step corresponds to generating the visualisations with DS elements by applying the rules that were defined previously. A set of enhanced visualisations was prototyped using a graphic design software (e.g. Figure 6.7, cf. Figure 6.6). In addition to the mapping of intention-rule-elements, three additional steps were completed regarding the design of the visualisations, as follows:

- **Remove elements** that do not add value to the graph. This step is known as decluttering. Therefore, everything in the visualisation was converted into grey scale to then highlight and/or emphasise only important aspects according to the rules. This step can be observed in Figure 6.7- C, where grids, legends, data points were removed. Also, axis labels were rotated, as suggested by Knaflic (Knaflic, 2015) so text could be easily read horizontally.
- **Emphasise lines** to show trending in the Performance (Figure 6.7 - bottom, element E).
- **Add context** through a shaded area (shapes and colour) according to the learning design (Figure 6.7, element F corresponding to the reflection stage).

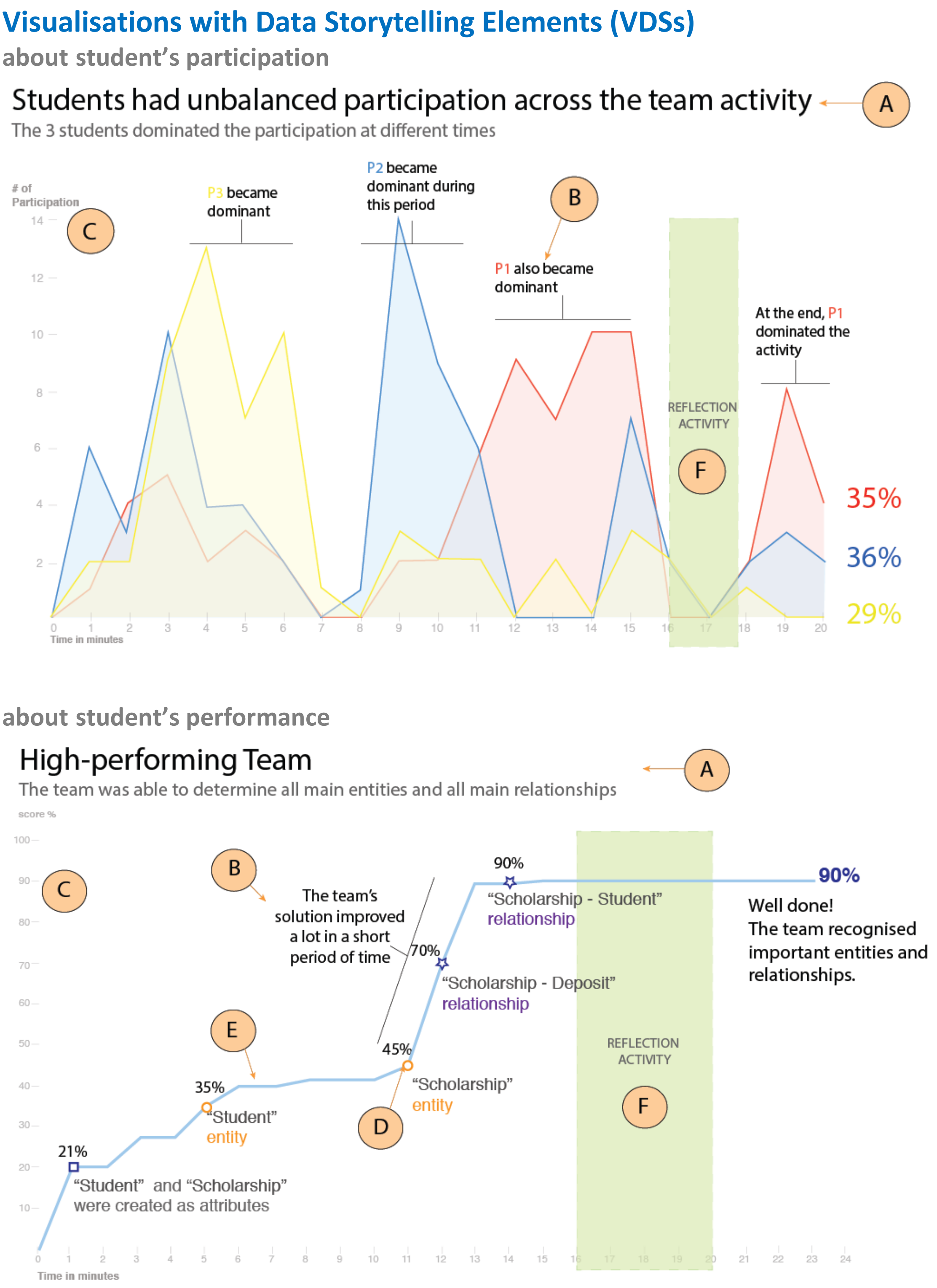


Figure 6.7: Explanatory visualisations after applying data storytelling principles and elements (A-F). *Top*: Participation levels of three students. *Bottom*: Evolution of the team's performance.

6.4.2 Teacher validation study

This section describes the illustrative study conducted with the aim of generating an understanding of the role of the data storytelling elements to support sensemaking. This study illustrates how teachers explored and interpreted exploratory and explanatory visualisations.

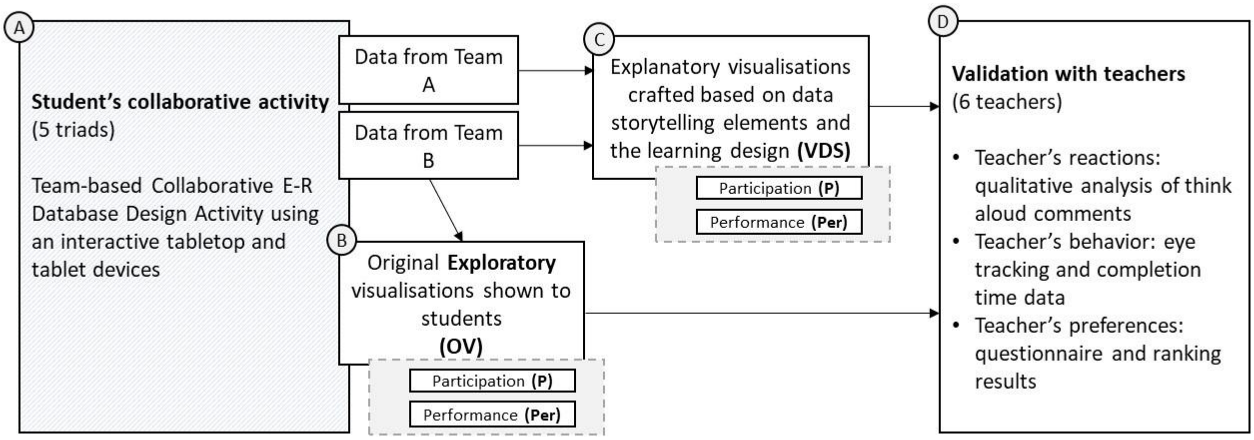


Figure 6.8: Validation process from an authentic student collaborative activity (A) a series of visualisations (OVs) were shown to those students in a dashboard as formative feedback (B). These visualisations (VDSs) were enhanced based on data storytelling elements (C). Both types of visualisations were shown then to teachers to analyse their reactions (D).

6.4.2.1 *Study setup*

A study was carefully designed to expose teachers to a number of explanatory visualisations with visual data storytelling elements (VDS) and original exploratory visualisations (OV) shown previously to the students (Echeverria, Martinez-Maldonado, Chiluiza, et al., 2017). Figure 6.8 presents an overview of the study setup. The intention of this study is not to experimentally test two conditions (e.g. exploratory vs. explanatory visualisations) but rather to focus on generating an understanding of the potential role that data storytelling elements can play to support interpretations of visual learning analytics and dashboards. This study aimed to obtain evidence about how teachers reacted to the different types of visualisations, both exploratory or explanatory, and visual design elements. The analysis is focused on three themes:

A1. **Teacher's reactions**. The aim is to investigate what kind of stories or comments could be formulated by teachers while looking at the visualisations. The hypothesis is that although both exploratory (OV) and explanatory (VDS) visualisations may be useful for teachers to gain insights into learners' activity, the data storytelling elements would play a role in helping teachers aligning those insights with those intentions materialised in the learning design.

A2. **Teacher's behaviour and focus of attention**. The aim is to explore whether explanatory visualisations drive teachers' focus of attention to specific elements (such as text labels, the title and key data points) as intended.

A3. **Visualisations' usefulness and teacher's preferences.** The aim is to examine the usability of visualisations, especially those with data storytelling (VDS) and their power in supporting teachers' monitoring for orchestration of the collaborative activity (Rodríguez-Triana, Martínez-Monés, Asensio-Pérez, & Dimitriadis, 2015). Furthermore, the goal is to explore which data storytelling elements are most helpful to support storytelling through learning visualisations.

6.4.2.2 *Participants*

Six Assistant Professors (T1-6; one female, five males; age range: 26-39 years) from ESPOL University were asked to participate in this study. All teachers had experience in teaching (avg.: 2.33 years), five of them had experience teaching the database systems course, and all of them had experience teaching software design.

6.4.2.3 *Materials*

A total of **eight visualisations** were prepared using data from the study described in Section 3.1.2.2 (see Figure 6.8-A). These visualisations were generated from 2 randomly selected student teams (A and B, from now on). Four original visualisations (OVs) were picked as shown to the students during the reflection task (see Figure 6.8-B). Then, data storytelling principles and elements were applied to each of the OVs, resulting in the second set of four prototypes (see Figure 6.8-C). Figure 6.7 depicts one example generated from team A's participation and performance data respectively. In short, each teacher would inspect data from two teams, regarding participation and performance, through visualisation with (explanatory - VDSs) and without (exploratory - OVs) data storytelling elements.

6.4.2.4 *Method*

The validation with teachers consisted of **four tasks** (Figure 6.8-D). A think-aloud protocol was used during the whole pilot study.

- For the *first task*, each teacher explored both the original (OVs) and the prototyped visualisations (VDSs) on a 21-inch screen connected to a laptop running Windows 10. Teachers were asked to inspect the visualisations and tell a story about what happened with the team, based on what she could interpret from the data **(A1. teacher's reactions)**. From now on, these will be referred to the *inspection episode*, as the task of exploring/verbalising insights from one visualisation.

  In order to control for a familiarity effect, two cases of inspection episodes were set up:

  1. *Case I* included 2 OVs from team A (participation and performance) and 2 VDSs from team B (participation and performance) – four in total; and,
  2. *Case II* included 2 VDSs from team A (participation and performance) and 2 OVs from team B (participation and performance) – four in total.

  Therefore, half of the teachers inspected *Case I* and the other half, *Case II*. In summary, 24 inspection episodes from the six teachers were captured. During the whole first task, data from an eye tracker (Tobii X-60 Pro) was recorded, with the aim of understanding how teachers visually explored the visualisations **(A2. teacher's behaviour and focus of attention)**.

- The *second task* involved the exploration of visualisations regarding their perceived usefulness for supporting monitoring and orchestration of the learning activity. Here, teachers explored OVs and VDSs about participation and performance for team A and B (e.g. Figure 6.6-top representing team A and team B participation, in the same screen). Thus, each teacher had the following sequence of inspection episodes: OVs about participation for team A and B, OVs about performance for team A and B, VDSs about participation for team A and B and VDSs about performance for team A and B (four in total). Teachers were asked to provide commentary on the ease of interpretation and support for orchestration and evaluation **(A3. visualisations' usefulness and teacher's preferences)**.
- In a *third task*, teachers compared both versions: OV and VDS on participation for a random team (A or B) and OV and VDS on performance for a random team (A or B) (e.g. Figure 6.6-top and Figure 6.7-top on the same screen). Teachers were asked to verbalise their thoughts about which of the two versions helped them interpret better the story and insights behind the data. Also, teachers were asked to suggest any change or addition to the VDSs **(A3. visualisations' usefulness and teacher's preferences)**.
- The *fourth task* involved the evaluation of specific data storytelling elements. For this task, teachers were asked to rank each of the data storytelling elements (A-F) in regard to the perceived helpfulness for storytelling. A five-point ranking scale was created to allow teachers to rank the role that each data storytelling played in the VDSs on participation and performance **(A3. visualisations' usefulness and teacher's preferences)**.

Video-recordings, eye-tracking data, screenshots, and responses to the questionnaires and ranking scale were gathered and organised for its subsequent analysis. A post-hoc semi-structured interview was conducted with each teacher after each session for them to explain any additional aspect of their experience.

6.4.3 Analysis and results

A mixed methods analysis approach was followed, aimed at triangulating quantitative evidence collected via the transcripts of the video-recordings of the think-aloud processes and interviews, the ranking instrument, and eye-tracking data with teacher's qualitative reactions. Video recordings were transcribed.

Qualitative data was thematically analysed using NVivo Software. Quotes that expressed teachers' reactions during the exploration of visualisations were selected and grouped by type of visualisation (i.e. exploratory or explanatory – participation and performance visualisations) (A1). Quotes from transcriptions that indicated visualisations' usefulness and teacher's preferences were selected and grouped into two themes: *ease of interpretation and orchestration*, and *teachers' preferences* (A3).

Quantitative data was gathered and processed. The eye-tracking data from each participant was processed by splitting the data according to the type of visualisation (i.e. exploratory or explanatory – participation and performance visualisations). This data was analysed in different ways to tackle A2: 1) time statistical descriptors (i.e. min, max, avg, std. dev) according to each type of visualisation to explore if there are any statistical differences in the exploration time; 2) heatmaps to visualise particular spots that triggered attention; 3) a clustering analysis of gaze data to examine which Areas of Interest (AoI) were critical during the exploration of visualisations, and 4) scanning trajectories to observe teachers' gaze behaviours during the first 10 seconds of exploration. In addition, ratings from DS elements were counted and ordered (A3). This information was triangulated with qualitative data and results are presented above (Section 6.4.3.3).

Next, results are organised and presented by the aim of the study.

6.4.3.1 *Teacher's reactions*

An initial exploration of the transcripts from interviews looked at the kind of reflections that teachers articulated while inspecting each of the visualisations **(A1)**.

For *exploratory visualisations* (OVs), teachers tended to focus on describing the data and then raise questions as they generate hypotheses to explain the data. For example, for the case of OVs about **participation**, teachers recognised that the visualisation was presenting all the information about students' individual participation. This behaviour was affirmed by teacher T1 as follows: *"with this visualisation, I can see an overall behaviour of team's participation*". However, teachers' comments related primarily to counting the number of actions performed by each student during particular periods of time and guessing why sometimes a student performed more (or fewer) actions, instead of gaining specific insights from the data. For instance, this is illustrated by one of the comments expressed by teacher T2 as follows: "*at some point, students in this team stopped the activity, but it is not clear why, maybe they were thinking, or only talking"*. Teachers also tried to explain two or more students' participation, pointing at the equality (or imbalance) in the team's participation. One of the teachers said "*It seems that the participation was imbalanced. The major number of actions were performed by participant 1 and participant 2. The participant 3 was only active during the first seven minutes. I have the impression that only two participants worked actively along the time*" (T3). Finally, some teachers attempted to identify potential leaders by counting the number of actions and guessing turn-takings among participants during the activity. However, as expressed by one of the teachers (T1) this proved not always to be an easy task: *"I can see who has dominated the activity in periods of time easily. However, I cannot tell which participant dominated the participation during the whole activity quantitatively"* Similarly, in regards to the OVs about team's **performance**, comments were in relation to the performance of the final product, how it evolved as students added entities and created relationships and how each action made a (positive or negative) impact on the performance. One of the teachers (T3) said: *"some entities were identified, some were right, others wrong, but the final performance was not any better".*

In short, exploratory (OVs) visualisations prompted teachers to construct stories, but they struggled to know if they were accurate: they inspected the data, described it, and raised questions. Some of their questions could be answered from the data, while others could not be from the limited evidence provided. This behaviour is in line with what Verbert et al. (2013) observed "learning analytics process model": i) data was presented to teachers; ii) they immediately started to formulate questions and iii) assessed how data is useful or relevant for addressing those questions. The critical problem is that teachers were not able to elaborate on statements that could sound like interesting insights about student's participation or performance. In this way, teachers were performing the analysis of the data without being trained to do so. This behaviour was precisely expected to happen when teachers face exploratory visualisations.

For *explanatory visualisations* (VDSs), it was hypothesised that teachers would elaborate on insights gained from the visualisations, rather than merely describing the data. In the case of the VDSs about **participation** (e.g. see Figure 6.7– top), comments were related to a participant's dominance during different periods of time and equality of participation. Moreover, teachers were more confident about dominance. This was stated by teacher T6 as follows: *"it can be clearly seen when a team member is dominating the participation in different time periods".* For the VDSs about team's **performance** (e.g. see Figure 6.7– bottom), teachers articulated their thoughts regarding how key actions made an impact on the final product. Teachers were able to direct their attention to critical moments in a team's process rather than analysing all the data points. Teacher T6 described this behaviour as follows: *"this [VDS] is more understandable when showing what each object means [referring to the text of each entity appearing in the visualisation]. For example, this visualisation shows me that during the first minutes, students identified two objects as attributes, but then, these changed to entities. It was interesting to see how the product evolved, and how the changes made to critical objects affected the performance"* Also, another comment from the teacher T2 reinforced this idea as follows: *"It gives me more detail about the number of entities that were considered for the final score."*

One teacher (T6) explicitly appreciated that the explanatory visualisation reduced the time they needed to spend processing the data: "*this [VDS] visualisation is explicit, … the information is auto-explained"* Another teacher (T5) supported this suggestion: *"this graph [VDS] was more detailed, informative. It does not make me overthink about the objects created. In the other visualisation [OV] I have to start guessing what 6E, 7E was. Then, after a while, I realised that these were the Entities or Attributes."*

As intended, teachers were able to construct more detailed stories about each visualisation. One excerpt of a teacher's story (T2) during the visual inspection of VDS-performance was as follows:

> *"At the beginning, the team created two entities, Students and Scholarship [reads the text of the data points]. The team did not make much progress during several minutes, from minute three to minute seven the team did not have significant activity [reads the text label]. Then, they were adding some objects, but the performance just raised a little bit. It is a low-performance score for fifteen minutes of activity. They have a short reflection time [reads the text in the shaded area], and then, they added one relationship, but it was not enough. In the end, the performance was very low [reads the text]"*

Another teacher (T4) also did use the text labels to explain the story (VDS participation) as follows:

> *"At the beginning of the activity, the three participants did not have any big differences in their participation. However, from minute five to minute seven, for two minutes [reads the text labels], the participant three (P3) have more participation, he dominated the participation [reads the text labels]. Then, P3 also was dominating the participation between minutes fourteen and fifteen. Then, there was a reflection time [reads the text in the shaded area], and no participation was recorded there. Finally, from minute twenty to minute twenty-three P3 became dominant [reads the text label]".*

One question that emerged during this analysis was if teachers used any of the data storytelling elements as a basis for their comments. Thus, each phrase when a teacher grounded their comment on the evidence provided by a particular data storytelling element was coded. Seventy-seven quotes were coded as *text labels*, *axis*, *data points*, *reflection*, or *title*. Summarising the coding results, it was noted that **text** narratives were the most referenced elements with 48%, followed by highlighted **data points** with 22%, the **shaded area** (that depicted a students' reflection period) 16% and the **title** (12%).

To conclude this section, the results from this initial qualitative analysis are encouraging. They suggest that teachers reacted as intended to the data storytelling elements. The exploratory visualisations (OVs) invited teachers to explore all the data but leaving the responsibility on them to discern insights. In contrast, the explanatory visualisations (VDSs) were more directive about the meaning of key points, which the teachers found plausible, and seemed to appreciate.

Self-report is a rich source of evidence, but what did the teachers, in fact, attended visually? The next step in the analysis consisted of an analysis of teachers' visual behaviour when presented with these dashboards, in order to provide answers to questions such as: *Where did teachers look in a given visualisation? What was the role played by each data storytelling element? Did the explanatory visualisation help them explore the data more efficiently?* The next subsection addresses these questions based on a quantitative analysis of gaze behaviour.

6.4.3.2 *Teachers' gaze behaviour*

This section describes how eye-tracking data shed additional light on whether the addition of data storytelling elements appeared to shape teachers' engagement with the visualisations **(A2)**.

Time spent by teachers to inspect the visualisations

Table 6.4: Min, max, average time spent and std. dev. for each inspection episode (n=24)

Purpose of visualisation	Type of visualisation	min (secs)	max (secs)	avg. (secs)	std. dev
Exploratory (OV)	participation	115.11	513.00	220.86	148.71
	performance	91.78	727.07	266.69	223.91
Explanatory (VDS)	participation	57.86	228.69	148.14	57.74
	performance	128.57	266.49	183.09	61.71

A starting point to understand teacher's gaze behaviour was to calculate the time they took for each inspection episode. Table 6.4 summarises the minimum (min), maximum (max) average (avg) and the standard deviation (std. dev.) of the time spent for inspection episode (n=24) of the six teachers. Given the higher standard deviation and the small sample size, significant differences were not found in relation to the average time spent inspecting exploratory and explanatory visualisations ($U=51$, $Z=-1.212$, $p=0.242$). However, it can be observed differences in standard deviations according to the purpose of visualisation, being larger in the explanatory visualisations. This suggests that teachers inspected the explanatory visualisations at similar amounts of times. By contrast, some teachers spent more time inspecting the exploratory visualisations compared to the explanatory visualisations.

Heatmaps of teacher's gaze behaviour

The first step in the analysis of teachers' gaze behaviour consisted of visually inspecting which areas of the visualisations attracted most visual attention. To address this, heatmaps were generated using the gaze data for each of the teachers' inspection episodes. First, heatmaps were generated using all the data points captured by the eye tracker through the inspection of each episode. However, after observing the generated heatmaps, it was noticed that the teacher's gaze covered almost the whole visualisation by the end of the inspection episode, making it difficult to determine which elements in a visualisation appeared to be critical in driving attention. Therefore, in order to enable a normalised visual comparison of the gaze behaviour, another set of heatmaps were generated using the first 60 seconds of each inspection, to try and capture the differences that the design has when first engaged. Twenty-four heatmaps were generated corresponding to the four different teachers' inspection episodes.

Figure 6.9 shows one example for each visualisation type (OVs and VDSs about student's participation and performance each). The heatmaps were scaled so that 1 second's fixation corresponds to a brighter red colour. It can be seen that for the exploratory visualisations (Figure 6.9, top) many hotspots are spread through the data points and legends (located at the top and centre of the OVs), suggesting that the teachers dedicated considerable time looking at each data point in the time series. These examples are representative of what was similarly observed in the other teachers' heatmaps (other examples heatmaps can be found in (https://goo.gl/QQumyz). In the interviews, most teachers explained that they struggled with interpreting the small numbers and letters in each data point and had to consult the legend to remember the meaning of each colour. Although this is a simple design issue quite easily solved during the decluttering of the graph, it is a subtle data storytelling consideration that can have a critical impact while teachers inspect a graph. By contrast, concerning the explanatory visualisations, Figure 6.9 - bottom depicts how the text labels and key data points drew most visual attention. In these two examples, teachers only focused on the narratives and the highlighted data points rather than the axis or the grid. From a visual comparison of the heatmaps, it was observed that, in the first 60 seconds of seeing the display

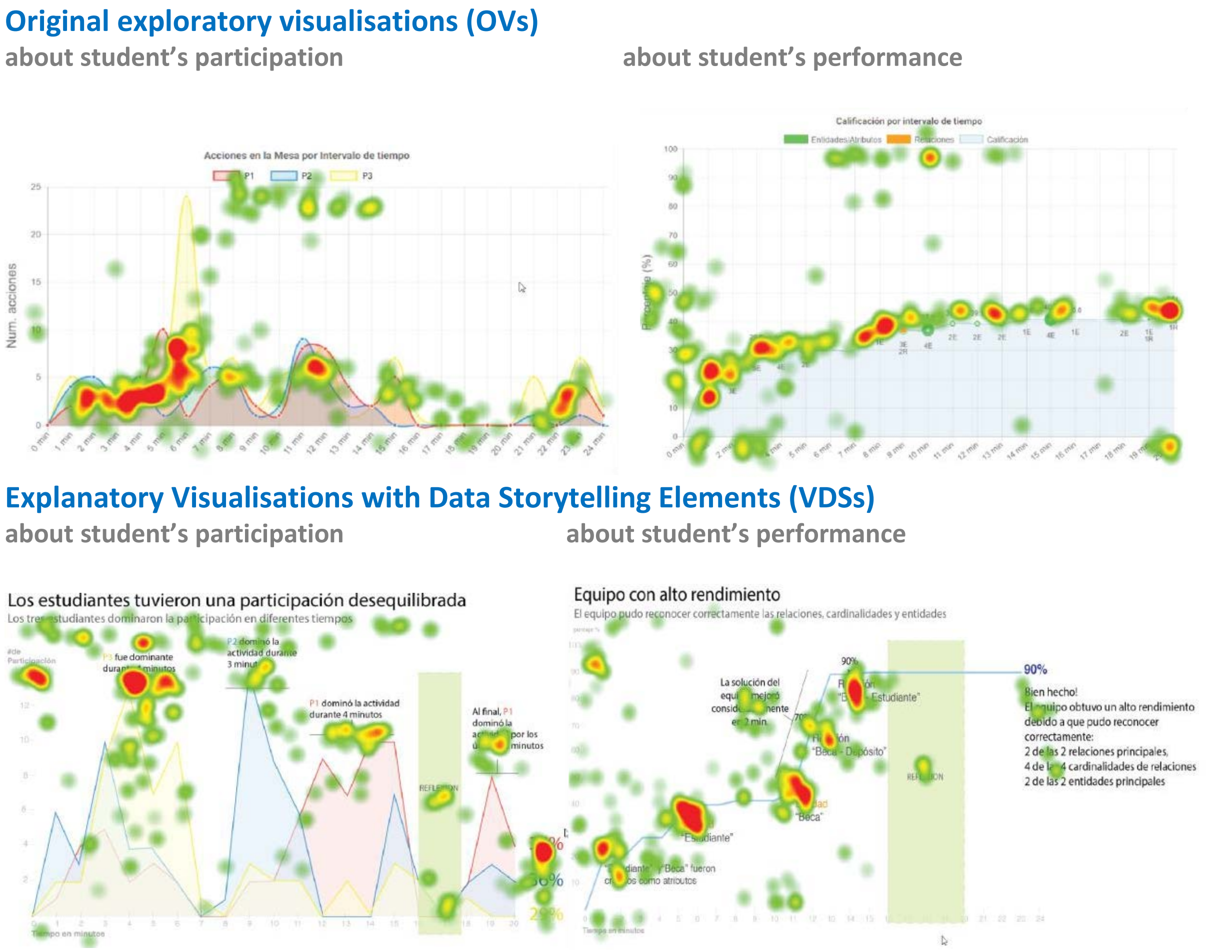


Figure 6.9: Example heatmaps of gaze behaviour of exploratory (top) and an explanatory (bottom) visualisations from four different teachers.

visualisations that featured a clean design, and only highlighted selected data points, and text labels (entities names and percentage of the score) seemed to be effective in reducing visual clutter and helped teachers to focus their attention. While the heatmaps assisted in performing a rapid visual comparison, the next step in the gaze analysis was to test if these differences could be quantified.

Areas of Interest (AOI)

To provide *meaning* to the data points captured by the eye tracking system, the most relevant Areas of Interest (AOI) were identified in each visualisation. This approach is a common practice in eye-tracking analysis methodologies to associate pixels in the screen with higher order areas (Toker, Conati, Steichen, & Carenini, 2013), in this case, data visualisation elements. Table 6.5 lists the AOIs regarding the purpose of the visualisation, either exploratory or explanatory, the latter being related with data storytelling elements. Each visualisation features some shared and some distinct design elements. For example, AOIs which were common across all the visualisations tested in the study included the *axis* (x and y) and the *data points* in the time series. The exploratory visualisations were the only ones containing *legends.* The explanatory visualisations contained three unique AOIs: *text narratives, a shaded area* and a *prescriptive title.* The final AOI was *"not applicable (NA)"* area in a visualisation, either corresponding to a blank space or another irrelevant element of the screen.

Table 6.5: Areas of interest (AOI) related with each visualization purpose.

AOIs in exploratory visualisations	AOIs in explanatory visualisations	Acronym
Axis (x)	Axis (x)	Ax
Axis (y)	Axis (y)	Ay
Data points	Data points	DP
Legends	-	Legd
-	Text narratives	Txt
-	Shaded Area	ShA
-	Title	Title
Non-applicable	Non-applicable	NA

Areas of Interest and Gaze Behaviour: Clustering Analysis

A cluster analysis was performed for each of the two types of visualisation (OVs and VDSs) aimed at giving meaning to the eye tracking data points according to the AOIs introduced above. A total number of 147,045 data points (33184, 43251, for participation and performance OVs, and 35988, 34622 for participation and performance VDSs, respectively) were analysed at this point. The Elbow method was applied (Kodinariya & Makwana, 2013) to identify the appropriate number of clusters in the dataset. This method was performed using all the data points of the inspection episodes of each of the eight different visualisations (2 OVs and 2 VDSs for two teams, A and B, about two types of visualisations, participation and performance). This way, it was found that k=15 would be a reasonable number of clusters that would provide enough information to differentiate AOIs across visualisations. Then, the k-means (k=15) cluster algorithm from the *scikit-learn* Python package (Pedregosa et al., 2011) was executed for each visualisation, grouping the data points of the three teachers' inspections for Case I and the other three teachers' inspections for Case II.

Figure 6.10 and Figure 6.11 depict two examples of the resulting clustering for an OV and a VDS (respectively) on participation for Team A (Figure 6.6-top and Figure 6.7-top show original visualisations without clustering points). The centroid of each cluster is marked with a circle.

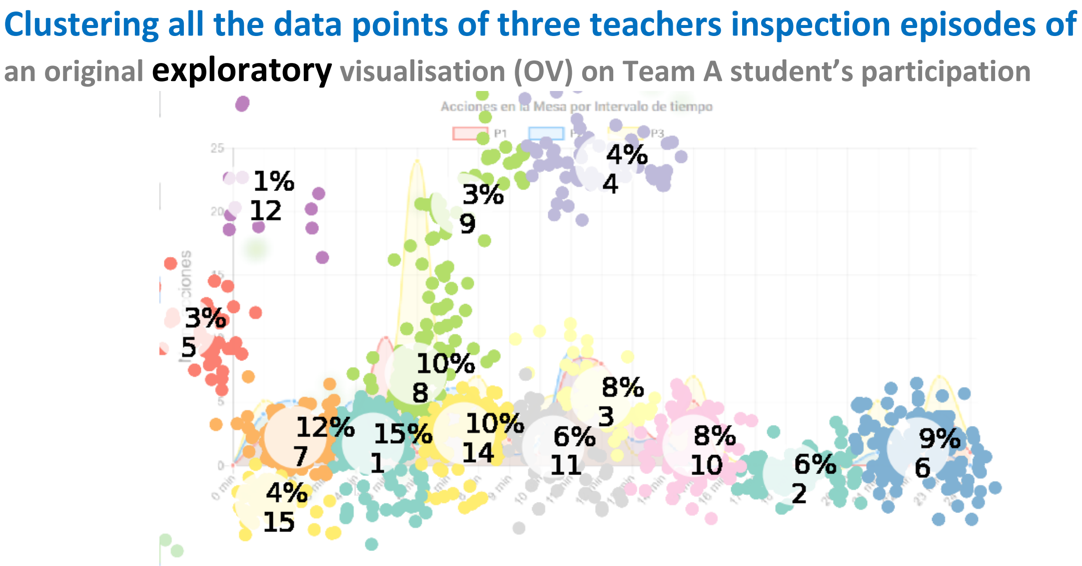


Figure 6.10: Results of clustering the eye tracking data points of the three inspections episodes recorded for one OV about team's participation (k=15). Circles correspond to the centroids of each cluster. Each is associated with the percentage of data points of the cluster (at the top of the centroid) and the id number of the cluster (at the bottom of the centroid).

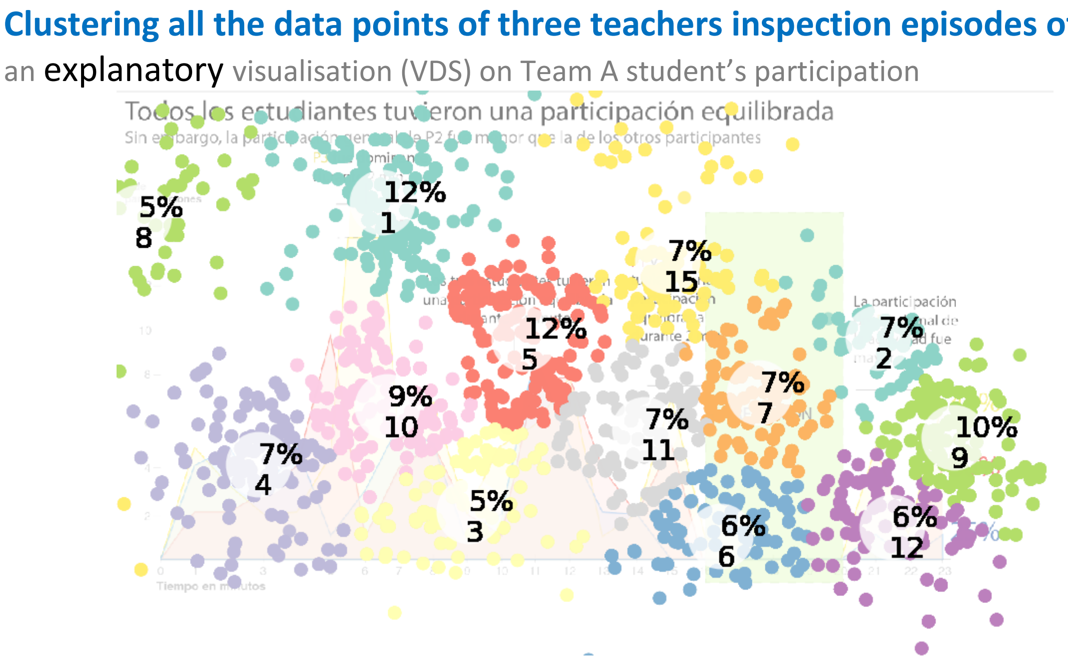


Figure 6.11: Results of clustering the eye tracking data points of the three inspections episodes recorded for one VDS about team's participation (k=15). Circles correspond to the centroids of each cluster. Each is associated with the number of data points of the cluster (at the top of the centroid) and the id number of the cluster (at the bottom of the centroid).

A visual inspection of these clusters gives an idea of the parts of the chart that the eye tracking data points were associated. To associate clusters with meaningful AOIs, one or two AOI (from Table 6.5) was manually assigned to each cluster. When a cluster did not belong to a relevant AOI, it was tagged as a Non-applicable AOI. For example, in Figure 6.10, eye tracking data points in cluster 4 are mostly associated with the *legend* at the top of the chart, cluster 5 with the vertical axis (*y-axis*), cluster 15 with the horizontal axis (*x-axis*), cluster 6 with the horizontal axis and student data of the time series (*x-axis/data points*), etc. Similarly, clusters in Figure 6.11 were associated with AOIs of the visualisations with data storytelling elements. For example, clusters 1, 2, 5 and 15 were associated with text labels (*text*). Table 6.6 summarises the proportion of attention placed on each AOI regarding each visualisation purpose (exploratory, explanatory) and type (participation, performance).

For the exploratory visualisations on participation and performance (columns 1 and 2 respectively) teachers spent more time inspecting the **data points** and the **horizontal axis** – 78.8% and 83.7% respectively from *data points*, *x-axis* and *x-axis/data points* AOIs. The third most inspected AOI was

Table 6.6: Percentage of occurrence of each AOI by visualization type from the six teachers.

Exploratory visualisations (OVs)				Explanatory visualisations (VDSs)			
Participation		Performance		Participation		Performance	
AOI	%	AOI	%	AOI	%	AOI	%
x-axis/ data points	48.6	data points	73	data points	32.4	text	41.4
data points	30.2	x-axis/ data points	10.7	text	26.7	text/data points	29.4
Legends/data points	10.1	Legend	7.9	text/ data points	13.5	Shaded area	4.5
y-axis	4.8	y-axis	5.1	shaded area/ data points	8.4	x-axis	4.3
Non-applicable	2.4	x-axis	2.9	text /title	6.6	Data points	3.8
Legend	2	Non-applicable	0.4	shaded area	3	y-axis/title	3.6
x-axis	1.9			y-axis	2.7	x-axis/data points	3.6
				x-axis/ data points	2.5	y-axis/data points	3.5
				y-axis/title	1.7	x-axis/text	2.9
				title	1	title	1.3
				x-axis	0.92	x-axis/shaded area	1.14
				Non-applicable	0.61	Non-applicable	0.46

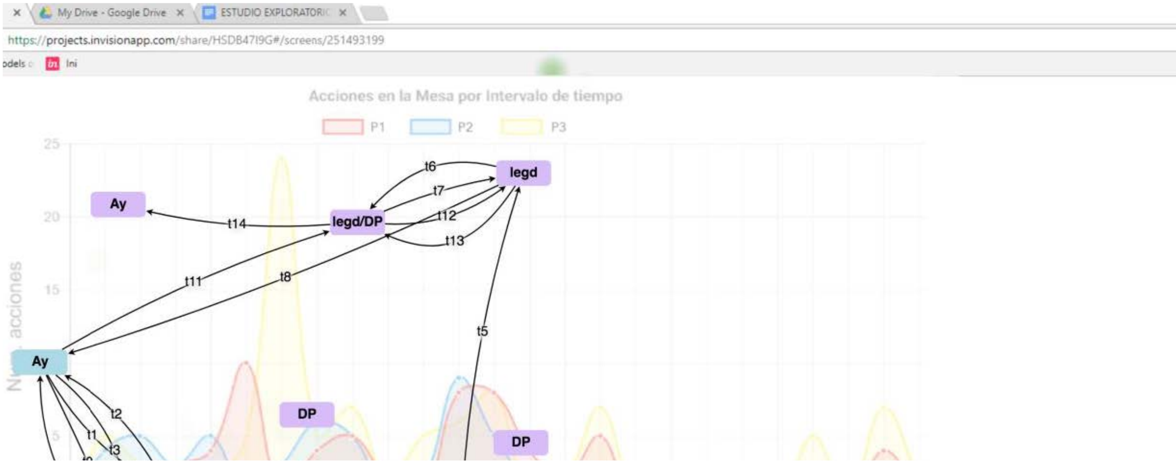
the **legend**. This confirms the qualitative results presented in the previous section. Teachers focused on exploring the data points and making sense of them by looking at the legends and the axis. In contrast, for the explanatory visualisations (columns 3 and 4) teachers read the **text narratives** (text) along with the **data points** - 72.6% and 70.8% for visualisations on participation and performance respectively, from *data points*, *text* and *text/data points* AOIs. This suggests that the addition of brief text narratives attracted the attention of teachers. If these narratives are brief and communicate insights, they can become into effective design elements to be added to learning analytics visualisations. Moreover, in this case, teachers looked at the axis to a lesser extent compared with the exploratory visualisations. This suggests that decluttering the visualisations allowed teachers to focus on key information in the visualisations.

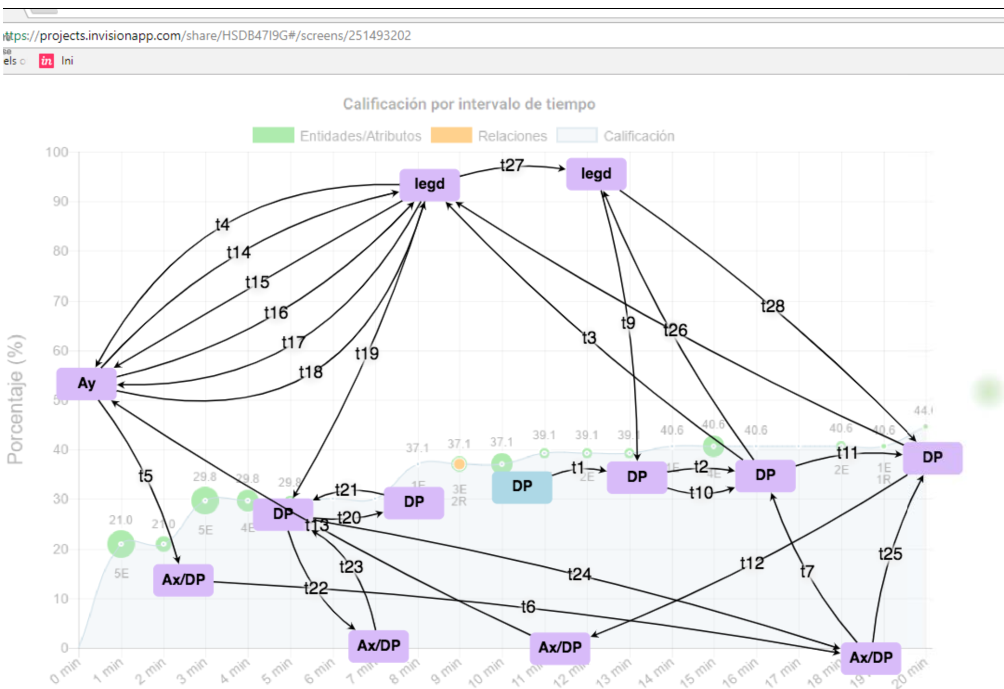
Scanning trajectories

To complete the teachers' gaze behaviour analysis, an additional analysis was performed to know which visualisation elements attracted the most attention in the first seconds of each inspection episode. To achieve this, the analysis was focused on the gaze data during the first 10 seconds of each episode. By associating eye tracking fixation points with the AOIs (through the clustering analysis explained in the previous subsection), visual representations were generated showing sequential trajectories that each teacher followed to make sense of the visualisation when facing it for the first time. Examples of these trajectories are shown in Figure 6.12 and Figure 6.13.

For the trajectories related to **exploratory visualisations** on student's participation, it can be noted in the example in Figure 6.12 - top that this teacher (T2) first explored the vertical axis to then start looking at the data points (see transitions t1-t2 between nodes Ay and DP at the bottom-left of the figure). Then, the teacher looked at the horizontal axis (t4) before inspecting the data points in the middle of the graph (t5). After, the teacher looked up at the region where the legend is and looked at the axis and data points coming back and forth to the legends (t6-14). For the exploratory visualisation about students' performance (see Figure 6.12 - bottom), this second teacher (T3) started inspecting the data points at the middle of the visualisation (see t1 and t2) and then his fixation was focused into the legend (t3).

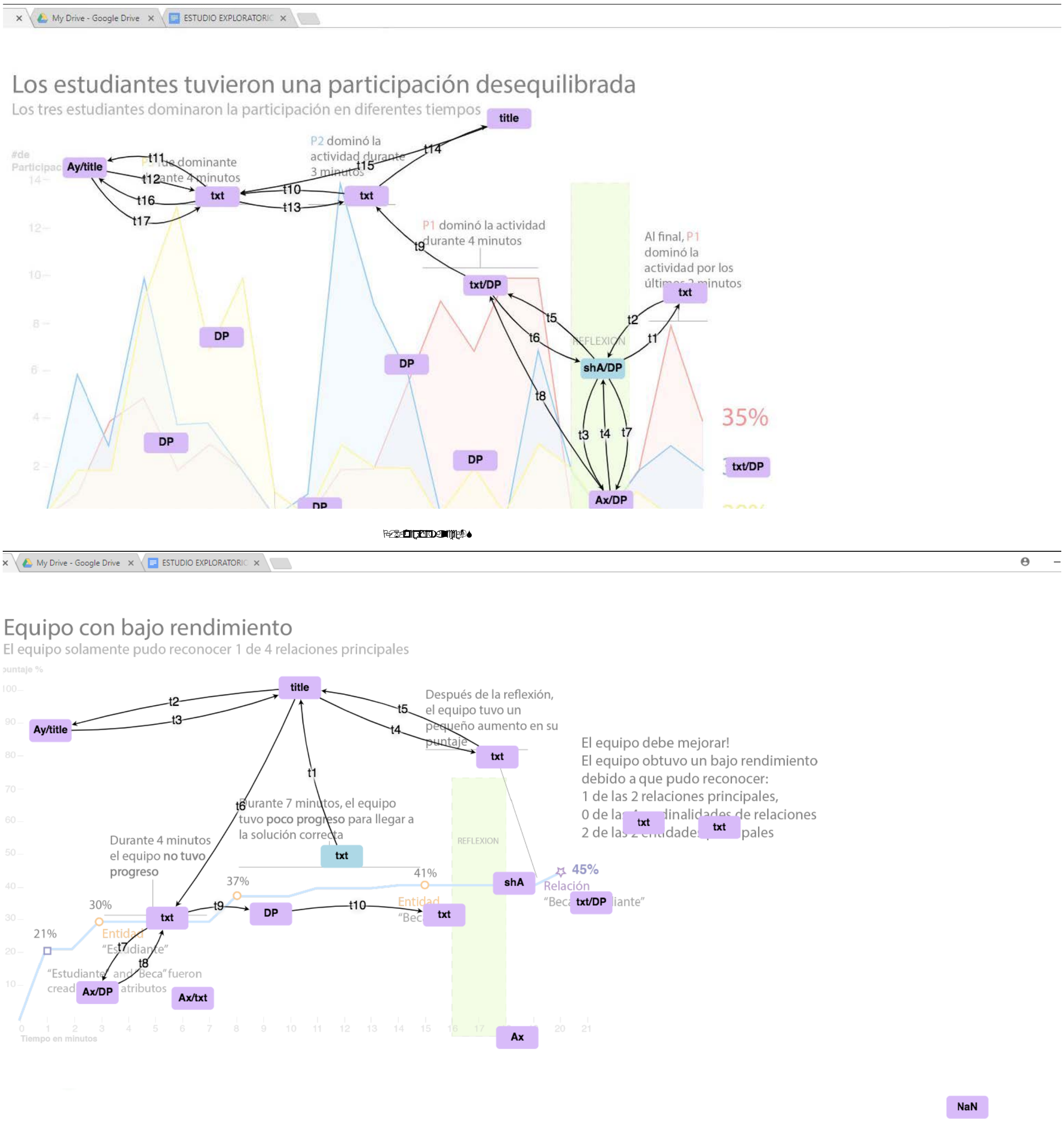
After this, the teacher gaze moved to the vertical axis (t4), the horizontal axis and various data points (t5-t7) to then move back to the legend (t8). The next eye movements scanned more data points, but the teacher had to return to the axis and the legend while inspecting the data points (t9-t28). Interestingly, triangulating the data from these trajectories and the interviews, it was noted that at the beginning of the exploration of these visualisations, teachers commonly verbalised the name of the elements they were looking at and started to ask clarification questions. For example, the teacher (T2) inspecting the first figure said: *"What is the meaning of P1, P2 and P3 in the legend? Are they the participants in the activities, right?"*. The teacher inspecting the second figure (T3) commented: *"[after being quiet for a period] the percentage [the label in the vertical axis] is the score obtained, right? So, the labels related to 5E, 3E? Also, the orange colour in the data points is because students added a relationship object, right?"*. These examples further suggest that teachers needed some time or some training before they could gain any insights from the visualisations at a glance. As illustrated by these exemplars and confirming the summaries of gaze analysis presented in the previous section, **data points, axes and legends are essential to make sense of the information for exploratory visualisations**.





By contrast, in regards of the trajectories related to **explanatory visualisation** about students' participation, it can be seen in the example depicted in Figure 6.13 - top that the teacher (T5) started to inspect the shaded area first (reflection) to then move to the first text narrative next to it (t1) and some data points below that text (see t2). After that, the teacher's gaze moved to the other pieces of text (see t5, t8, t10). Indeed, this teacher did not spend much time looking at the data points in the first seconds but instead paid more attention to the narratives. The teacher (T6) in the second example (Figure 6.13, bottom) started inspecting a text narrative at the centre of the visualisation and then moved up to the title (t1-t3). After, this teacher also looked at a second narrative (t4-t5), and then started to scan the highlighted data points from left to right (see t6 to t10).

These two examples show how **the data storytelling elements, the text narratives, the title and the highlighted data points, drove the attention of these teachers from the beginning**.



These examples are representative of what was similarly observed in the other teachers' heatmaps. This suggests that teachers may have dedicated more time to gain insights from the text rather than expending time in trying to make sense of the graph. From the interviews, one teacher (T5) described this critical difference as follows: "*this is a visualisation about who dominated the participation [referring to an explanatory visualisation] and not as just the number of actions per participants [referring to an exploratory visualisation]*". Besides, data storytelling elements supported the teacher's stories and comments. The teacher (T6) told a story utilising the same information that was illustrated in the text narratives, data points, title and shaded area. Moreover, this teacher followed the same trajectory as the one presented in Figure 6.13 - bottom, when verbalising the story and comments: "*Initially, students added two entities, but as attributes* (t6). *Then, the students added some entities. An object that was added previously as an attribute, Student, now became an entity* (t7). *Then, it* [referring to the visualisation] *depicts a period of inactivity. The team did not make any progress for four minutes* (t8). *Then, the team made progress on a seven percent of their actual performance* (t9). *After this, the team's performance remained the same for seven minutes. They realised that the object Scholarship was an entity, instead of an attribute* (t10)"

In summary, comparing these results with the results from the qualitative analysis, data storytelling elements drove the teacher's attention and support the externalisation of visualisation insights effectively.

6.4.3.3 *Visualisations' usefulness and teachers' preferences*

This section examines the usefulness of visualisations, especially those for explanatory purpose and their power in supporting teachers' monitoring for orchestration of the collaborative activity **(A3)**. Thus, quotes from transcriptions were analysed and grouped into two themes: *ease of interpretation and orchestration*, and *teachers' preferences.*

In addition, this section reports *which data storytelling elements are most helpful* to support sensemaking. Analyses from quotes and ranking scales are summarised below.

Ease of Interpretation and Orchestration

During the interviews, the majority of teachers converged about the ***ease of interpretation*** that explanatory visualisations provided. Most of the teachers also agreed that the VDSs helped them articulate more compelling stories about what happened during students' activity due to the context added through the data storytelling elements. This point of view was echoed by teacher T1 who stated the following: "*these visualisations* [with data storytelling elements] *show me exactly what I need. I do not have to guess what happened in the activity*". Teacher T3 added that *"...this* [with data storytelling elements] *gives a summary of the whole graph. With the original visualisation, I would have had to explore the whole visualisation in order to understand team's performance"*; and teacher

T4 expressed the following: *"the [data storytelling] visualisation shows me what elements students have created very clearly, compared to the other [OV] visualisation"*.

Regarding ***orchestration***, teachers also mentioned that adding contextual information to the explanatory visualisations helped them understand the behaviour of the team. In the visualisations about students' performance, for example, associating text labels to key data points (i.e. the names of entities and relationships added by students) helped teachers understand the causes of potential unintended mistakes, so they could provide students with feedback if this visualisation would be generated in real-time or to guide a data-informed post-hoc reflection. For the visualisation about students' participation, teacher T4 expressed that "*the shaded area in the graph made [her] aware that groups took very different amounts of time to reflect and so, with this information, [she] would encourage students to reflect for a long time before submitting a final version of the task*". Another teacher (T2) mentioned that "*the title showed the right message of the graph regarding the team's participation, which made [him] think that this group needed to collaborate more equally.*"

Teacher's Preferences

During the interviews, teachers were asked to select which visualisation they would prefer and why, in order to explore design issues in detail. When analysing the interviews, it was found that teachers referred to two specific reasons when picking a visualisation: the *functionality* and the *design* of the visualisation. For the visualisation on **participation** data, teachers argued that they would prefer the **VDS visualisation**, because of its *functionality*. Teachers reported that VDS visualisations are informative, and the text is legible and easy to follow. Considering the **performance** visualisation, the majority of teachers also agreed they would prefer **VDS visualisation**. They indicated that colours were helpful to highlight information (e.g. line and data points) and the summary provided facilitated the interpretation of the data very quickly. However, some critiques were made concerning the *design* of the visualisations with data storytelling elements. For the VDS about **participation**, teachers suggested that instead of having straight lines, they would prefer smoother lines (e.g. curves) because they visually look better. Another criticism about the design was related with the balance of additional text labels, as the visualisation could become complex. In fact, two teachers (T1 and T6) mentioned that text labels *"may add noise and graph complexity"* and that the text label information may be on occasions *"redundant with no extra information."*

Data Storytelling Elements Helpfulness

Figure 6.14 shows the final rankings by the six teachers after inspecting VDS. According to teachers' perception, they ranked the prescriptive **title** as a critical element to facilitate the sensemaking of both types of visualisations. Five of them agreed that the title served to indicate the main message from the visualisations before going deeply into inspecting the details from the visualisations. This was stated by two teachers, as follows: (T4) *"[the title] shows an overall description of the graph, which helps to interpret faster the visualisation"* and that (T1) *"[title] was important to summarise and understand the story".*

Next in the ranking of data storytelling elements, is the **shaded area**. Teachers made an explicit comment about how the shaded area helped them to interpret better the sudden drop of activity in the visualisation: (T2) *"[the shaded area] explains what students were doing. Without this, I would think that students were doing nothing related to the activity"*; (T3) *"it adds context to the visualisation, to know what happened during the activity"*; and (T4) *"if I had to review this information after the class, this would help me to remember what happened in that point."* However, this element was ranked as the least helpful in the VDS about performance. Differences in both visualisations results can be explained by the fact that to understand the extent to which students participated, the teacher needs to be more aware of the behaviour of the team (e.g. *why* there was a drop in the activity), whereas, for task performance, teachers did not perceive that it would help to understand the context.

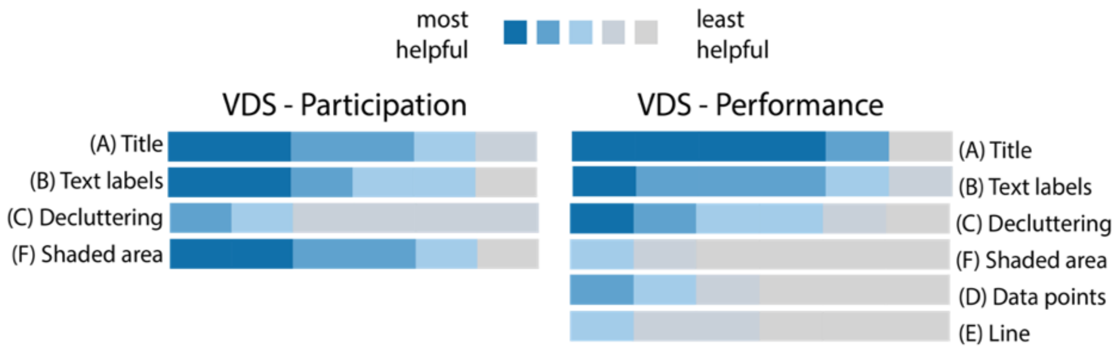


Figure 6.14: Ranking of data storytelling elements to support storytelling in visualisations.

Ranked in second and third place, for VDSs about performance and participation respectively, were **text labels** adding narrative to the visualisations. Teachers pointed out that text labels facilitated the narrative explanation of what occurred in different moments and ease of reading by describing what happened during the activity: (T2) *"[the text labels] helped me to see what happened when the three students were collaborating (for visualisation about participation). The first thing I read in the graph were the labels, which pointed me to explore in detail the data presented".*

Decluttering was ranked in third place for the VDSs on **performance** (Figure 6.14, right). Teachers expressed that the information was presented more clearly than in the OV: *"removing the green background, grids and big circles, made easier to explore the information".* Teachers also indicated that because the original visualisation (OV) was visually overloaded, it was convenient to declutter the graph. As a result, the visualisation was easier to navigate. For example, one teacher (T4) stated that *"it was very useful to declutter the graph. Now it is easier to identify relevant objects that were created over the time"*, and another teacher (T5) mentioned that *"decluttering the graph in this visualisation was very relevant. I can now see exactly where is the main object that increased the* *performance significantly.".* However, the same element was the lowest in the ranking for VDS on **participation (**Figure 6.14, left). From the comments, teachers expressed that the grid was relevant to find data points values. One teacher (T3) suggested that "*it [the OV visualisation] was better with the grid because it is easier to find values during the inspection*". In the same argument, another teacher (T4) indicated that *"you can get confused without the grid when exploring the visualisation"* One explanation of this result is that, due to the higher number data points and lines (one line per participant), backup information is needed to explain the data points values. In the crafted visualisations, the grid was removed and textual information about equity of participation was added. Textual information about the number of actions made by each participant was not added, as the intended message was to prescribe the participation behaviour (balanced or unbalanced).

Notably, other elements which in principle are important for data storytelling (Knaflic, 2015) were lower ranked. These included, for example, the highlighted data points and using a thicker line to highlight activity. For instance, teachers mentioned that **data points** are useful for showing only relevant information and that the thicker **line** makes easier to focus the attention on the information. With this, it cannot be argued that these elements should not be considered for further explorations in learning dashboards. Instead, teachers could perceive other elements as more helpful for them.

When triangulating this information with gaze behavioural data, it was noticed that just highlighting specific data points helped teachers to gain understanding about what occurred with the teams without consulting all the data points available. In addition, heatmaps and trajectories analyses suggested that text labels and data points were the most visited elements. Interestingly, according to teachers' perception, the title, text labels and shaded areas were catalogued as the most helpful ones.

Summing up the analysis and results presented in this illustrative study, the following findings may be relevant for developing tools that invites to communicate insights through explanatory visualisations:

- Data storytelling elements have the potential to help teachers explore visualisations with less effort. All teachers could interpret both the VDSs and OVs, but they explicitly expressed most affinity with VDS.
- Exploratory visualisations (OVs) prompted teachers to explore the data, describe it and raise questions, whereas explanatory visualisation (VDS) provoked more compelling comments about students' activities. This suggests that data storytelling elements may be used as a way of guiding people to gain insights from visualisations without the help of a technician.
- Teachers' behaviours point to the critical role that both narratives and visuals play in explaining the data. Text labels and key data points were the most relevant elements to

draw attention to an explanatory visualisation. This result was confirmed from the eye tracker data and teacher's perceptions. However, as also suggested by teachers, visualisations should be carefully designed in order to avoid clutter with too many text labels.

- Other design implications should be taken into consideration when designing the final visualisations. Colours, (smoother, straight) lines, grids and legends are visual elements that may help users to understand the visualisations better. Whichever is the case, designers should be aware of the visualisation's learning intention.
- A limitation of this study is that, due to the low number of teachers, these results cannot be generalised in the way permitted by an experimental lab study with a statistically meaningful number of subjects. However, this approach helped elicit rapid formative feedback from teachers to start understanding the impact that explanatory visualisation may have on teachers' perception and sensemaking.

6.5 SUMMARY

In order to partially answer *RQ3: How can explanatory multimodal interfaces be implemented to guide reflection on group work?* this chapter aimed to explore the effectiveness of data storytelling elements as a way to guide the reflection of group visualisations. To cope with this, a first exploratory study was carried out with potential users of the tool. The exploratory study served to examine for first time the concept of *educational data storytelling*. Findings from user's perspectives indicated the potential of DS elements to enhance the communicative power of complex learning visualisations that tells *one story at a time*. The preliminary rules and DS elements described in the exploratory study served to develop a more comprehensive approach to link the learning design with DS elements. The *learning design-driven data storytelling* approach brings data storytelling principles to support the interpretation of questions in teachers and students-facing dashboards that can become into explanatory tools. An illustrative study with teachers helped explore teachers' focus of attention and preferences of learning visualisations thoroughly. The results of the study are encouraging and provide the basis to investigate further the means for crafting personalised and context-dependent dashboards.

While the data stories crafted in this chapter addressed *one story per visualisation*, the next challenge in this research is to tell multiple stories from different data sources by merging multiple visualisations or compacting multiple data sources into one visualisation, in order to provide a compelling story of team's constructs and without falling in complex visual representations. This challenge is explored in Chapter 7.

Chapter 7: A Multimodal Layered Approach for Teamwork Reflection

This chapter presents the final outcome of this thesis in the form of an approach to design multimodal learning analytics interfaces that communicate *one insight at a time*. This is done by operationalising data storytelling (DS) principles, the Multimodal Matrix, and findings from the studies presented in previous chapters into a multimodal layered interface. First, this chapter describes the Explanatory Visual Layered (EvisLA) approach to facilitate the interpretation and sensemaking of multimodal data. This builds on elements of HCD-MMLA and the Multimodal Matrix modelling approaches. Then, a *layered model* is introduced to address the complexity of visualising multiple streams of data. Second, this chapter demonstrates how the EvisLA approach was operationalised to generate a high-fidelity prototype of a debriefing tool to guide students' and teachers' reflections in the context of teamwork nursing simulations. Third, the chapter presents a study with eight subject matter experts to explore the orchestration opportunities and data privacy implications of this multi-layered team replay tool, and validate the information presented. The results and discussion of the study are reported, concluding with a summary.

7.1 INTRODUCTION

This chapter is focused on describing the final iteration in the design and development of a solution that facilitates the interpretation and sensemaking of multimodal data. In doing so, this chapter takes the findings and approaches presented in previous chapters to deliver an Explanatory visual Layered (EvisLA) approach. The requirements elicited for MMLA interfaces presented in Section 4.5.2, and the semantic representation of group data through the application of the Multimodal Matrix modelling approach described in Section 5.2, allowed the project to map from learning constructs to low-level data and meaningful data representations. Findings from the validation of collaboration proxies indicated that *combinations* of proxies may provide a more complete view of a team's behaviour and performance. Next, the exploration of the potential of linking data storytelling (DS) principles and elements with the learning design were described in Chapter 6:. This allowed the project to design and deliver explanatory visualisations that communicate one story at a time. Findings from the illustrative study offered a comprehensive understanding of how visual DS elements could drive attention of teachers to focus on learning insights and findings. Adding narratives and highlighting data points to regular learning analytics visualisations did indeed appear to drive visual attention. The final step in this thesis is to bring these two strands together.

Figure 7.1 illustrates the research question, goal, contribution and validation methods addressed in this chapter. This chapter is focused in responding the following question: *How can explanatory multimodal interfaces be implemented to guide reflection on group work?* (RQ3). To this end, the EvisLA approach (Contribution 3) combines the *Multimodal Matrix approach* (Chapter 5) and the *Data Storytelling approach* (Chapter 6) to offer a solution that combines explanatory visual elements with key group constructs to guide the interpretation and sensemaking of multimodal data (Goal 3). A *layered model* is introduced as a solution to address the complexity of visualising multimodal evidence. This approach is illustrated by designing and generating a multimodal layered interface, using the information from the collaboration proxies presented in Chapter 5 in the context of nursing simulations. The multimodal layered interface is validated through a study with eight experts in nursing education to explore the usefulness of these multimodal layers, including *i)* the validation of the rules that connect the DS elements with the learning design, *ii)* orchestration opportunities for the multimodal layered interface, and *iii)* data privacy concerns. The next section describes the motivation of a layered model for visualising multimodal data, which is an essential component of the EvisLA approach.

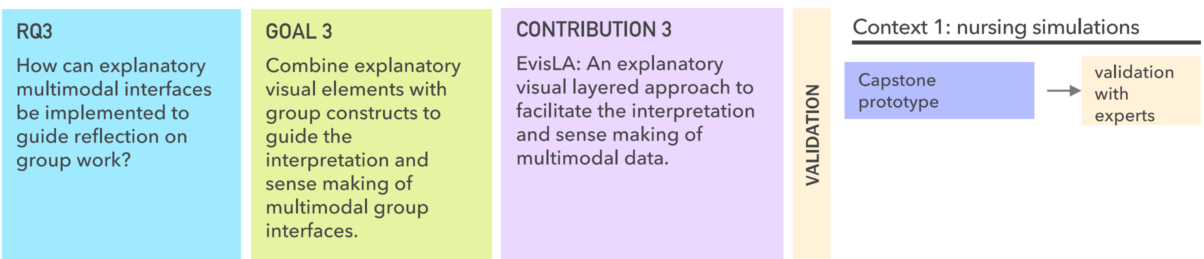


Figure 7.1: Research question, goal, contribution and validation methods of Chapter 7.

7.2 EXPLANATORY MULTIMODAL LEARNING ANALYTICS INTERFACES

While there has been significant progress to integrate and analyse multimodal data, little research has tackled the challenge of providing methods to visualise, understand and support the sensemaking of collaborative data in various situations, to inform teaching and learning practice (see discussion in current literature on Section 2.4.3). For instance, students could benefit from multimodal evidence about their success or failure moments (e.g. Spikol et al., 2017) by reflecting upon their practice and for future activities. As for teachers, multimodal visualisations could provide opportunities to guide a deeper reflection on groups' expertise (e.g. Ochoa, Chiluiza, Mendez, et al., 2013) or competencies (e.g. Cukurova et al., 2018) with evidence such as verbal interactions, emotional traits, gaze direction and performance. However, it is naïve to expect that simply visualising multiple data streams enables a learner or teacher to make sense of them, either individually or in combination, as each data stream can be interpreted and used for a particular purpose, and in combination, further meanings can be inferred which may or may not be helpful for learning and teaching (Ochoa, 2017).

As reviewed in Section 2.4.3, approaches to visualise multimodal data have been implemented for researchers to annotate groups' regulation strategies (Noroozi et al., 2018); to visualise multiple streams of synchronised data (Di Mitri et al., 2019); to inform students about their emotional regulation (Azevedo et al., 2017); and to provide personalised support to teachers on students' disengagement (Aslan et al., 2019). In addition, most current interfaces that support teamwork feedback are limited to mirroring simple information, such as the amount of speech, or have been deployed only in controlled lab settings (see discussion in current literature on Section 2.4.3.2). This work is, therefore, still in its infancy, with little evidence on how to create multimodal analytics interfaces that are meaningful to teachers and students, in contrast to researchers.

In the previous chapter, a learning design-driven data storytelling approach was introduced to link the learning design (i.e. teachers' intentions, expert knowledge) with DS elements aimed at designing and delivering explanatory interfaces. These explanatory interfaces are intended to drive a user's attention to critical aspects of the activity. Findings from the validation study reported in the last chapter demonstrated that adding narratives, such as text labels and prescriptive titles, highlighting key points, and communicating *one story at a time* all have the potential to help teachers understand the visualisation with less effort. However, the low-fidelity prototypes generated in the previous chapter from the DBCollab tool did not exploited multimodal evidence. Therefore, the challenge now is to deliver explanatory visualisations that communicate one insight at a time, but at the same time address the difficulty of interpreting low-level multimodal data by fusing multimodal data into a single interface. The next section describes the EvisLA approach to tackle the aforementioned challenge.

7.3 THE EVISLA APPROACH[21]

It is important to provide mechanisms for learning visualisations to deal with the complexity and integration of learning data (Alhadad, 2018; Schwendimann et al., 2017; Vieira et al., 2018). Firstly, as highlighted by Aguilar (2016) *"students may not need more information to learn more"* (p. 126). Instead, learning visualisations should focus on *"what information is required to provide meaningful feedback to the students"* (Matcha et al., 2019). In addition, Alhadad (2018) suggested that the LA community needs to move towards human-centred data visualisations, where knowledge of humans' limited capacity to process information (Miller, 1956) guides the design of learning

[21] Parts of this section and following sections are in preparation for publication in CHI'20 (Martinez-Maldonado, Echeverria, Fernandez, & Shum, 2020) and ILE journal (Martinez-Maldonado, Echeverria, et al., 2019)

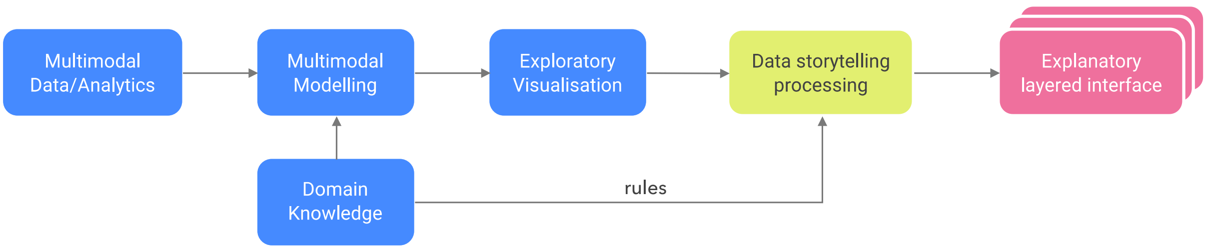
visualisations. The Explanatory Visual Layered (EvisLA) approach explores this perspective by highlighting what is really important for teachers to guide the visual attention to key aspects of the activity and scaffolding the exploration of complex multimodal data. The aim of the EvisLA approach is two-fold:

1) reduce the complexity of multimodal data by categorising the underlaying data into meaningful layers of information, and
2) apply DS principles and elements to drive visual attention to key events of the learning activity.

7.3.1 EvisLA components

This is accomplished firstly, by operationalising the Multimodal Matrix approach to add meaning to multimodal data (detailed in Sec.5.2). Figure 7.2 shows the elements from the Multimodal Matrix approach (blue rectangles), where multimodal low-level data are encoded into a data structure, the *Multimodal Modelling*, that incorporates learning design information. The data structure is used to generate *exploratory* visualisations, that is, providing no explicit guidance on what to attend to, or what the information might mean (see Section 5.2 for details).

Next, the learning design-driven data storytelling approach is applied to the exploratory visualisations. This is depicted in the green rectangle in Figure 7.2. The data storytelling processing encodes DS principles and learning design rules (extracted from the learning design) into DS visual elements (Table 7.1) that can be rendered in an *explanatory* visualisation (as described in Section 6.3). These explanatory visualisations can be shown in a *layered interface* (pink rectangle in Figure 7.2), which aims to place the user in control of how and when to view different types of feedback, as



detailed next.

Figure 7.2: The EvisLA approach to integrating multimodal data/analytics with learning design-inspired storytelling visual elements, to generate layered visualisations.

7.3.2 Explanatory layered interface

Layered models and user interfaces have been widely used in different domains, such as neuroscience and computer systems, as a method to understand complex systems (Gazzaniga, 2015). For example, Geographic information systems (GIS) commonly involve multiple types of data, and we are now familiar from the popularisation of GIS (Google Maps; Open StreetMap; etc) with turning layers on and off to combine lenses or levels of detail. Such controls have been shown to facilitate collaborative geographic sensemaking (e.g. Gaillard et al., 2015; Haslett, Wills, & Unwin, 1990). Any complex system can be simplified by breaking it down into individual layers. Each *layer* can encapsulate a concept, idea or a small part of the complex system. Moreover, according to Munzner (2014), an approach to handle visual complexity is by combining multiple *layers* within a shared frame.

Motivated by Doyle and Csete (2011) and Munzner (2014), a set of principles for crafting layered interfaces using multimodal data is presented:

L1. Each layer should represent a particular view (Munzner, 2014), showing only one data story at a time. This can facilitate the communication of insights while minimising distraction, as suggested from the findings in Section 6.4.3.

L2. A static layer should be used as a shared frame. Layers must share visual landmarks to facilitate user orientation as layers are revealed/hidden (Munzner, 2014). Each data story in a layer should be contextualised to the characteristics of the learning activity, such as duration, participants and critical incidents. This can be achieved by keeping the static layer (faded) in the background and overlaying others on top.

L3. The user should be able to select and combine layers. Each layer has its own purpose to fulfil as a part of the complex system (Doyle & Csete, 2011; Munzner, 2014). A layer presents a data story that communicates a class of insights about team activity. Thus, teachers and students should be able to flexibly select and combine different layers to facilitate sensemaking.

L4. Data points highlighted in different layers should be clearly distinguished. As each layer represents different information, data points can be superimposed within the same coordinate system (Doyle & Csete, 2011). Therefore, different shapes/colours should be used in each layer.

L5. The content of each layer should be defined based on pedagogical intentions. Each data story should be aligned with the learning design (Hernández-Leo, Martinez-Maldonado, Pardo, Muñoz-Cristóbal, & Rodríguez-Triana, 2019) to effectively provoke reflection on aspects that are relevant and that the teacher would normally provide feedback on.

The next section shows how these principles and components of the EvisLA approach were materialised in an authentic study.

7.4 ILLUSTRATIVE STUDY

A high-fidelity prototype was designed and developed based on the design principles presented above. The prototype was tested with eight experts' educators in the context of a teamwork simulation scenario.

7.4.1 Materials and prototype designs

This study switched from the database design task, to nursing simulations (see Section 3.1.1). The particular simulation scenario this study focused on consisted of two phases: in the *first* phase, a simulated patient experiences chest pain which nurses must assess and treat. In the *second* phase, the patient loses consciousness and the nurses must provide basic life support (the same task motivating the collaboration proxies presented in Section 5.5).

A high-fidelity EvisLA prototype was designed using part of the information presented in the physical activity and arousal peaks (Section 5.5, proxy 3) and the team timeline (Section 5.5, proxy 4). These two proxies were selected due to the feasibility of generating (nearly) fully automatic analytics that could be deployed and validated in debriefing. A static layer, a *timeline of actions* (e.g. Figure 7.3) performed by each nurse during the sim (captured by the manikin and the observer), was used as the background of the prototype (in line with principle *L2)*.

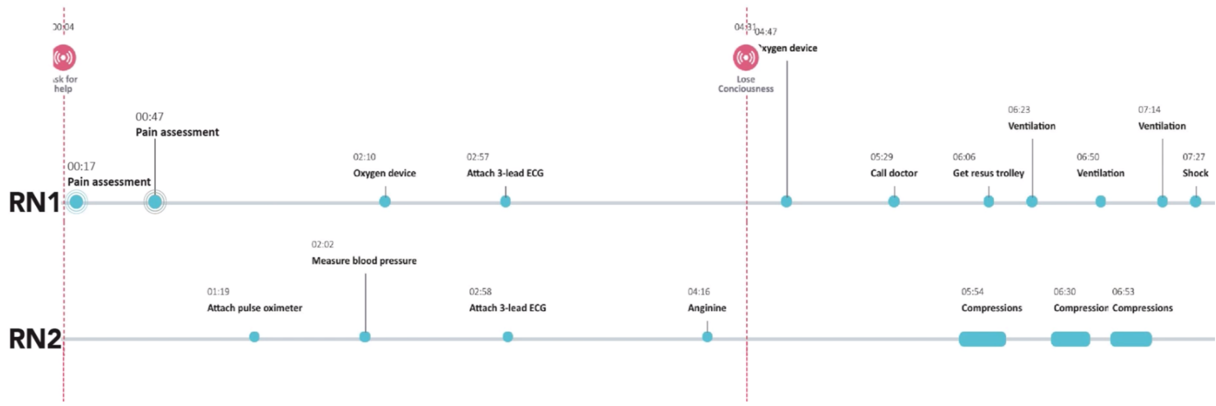


Figure 7.3: Timeline of actions that served as the common reference (background) of the EvisLA prototype: a team of two nursing students working before and after the patient's loss of consciousness.

7.4.1.1 *Layers*

Four layers, each representing a particular view of the multimodal data (*L1*), were generated based on the simulation tasks (*L5*). These focused on: *i*) *time responsiveness* of nurses performing critical tasks; *ii*) *mistakes made,* or actions performed poorly or in the wrong order; *iii*) *arousal (skin conductance) peaks* detected by an increase of 0.03 µs (Braithwaite et al., 2013) using EDA Explorer (Taylor et al., 2015); and *iv*) *location* of nurses in key spaces. The prototype was developed using a JavaScript library, with the functionality for users to select one or more layers in combination (*L3*). The timeline is decluttered, and its opacity is reduced for selected data points to appear emphasised (L4). These four layers are examined in the following sections.

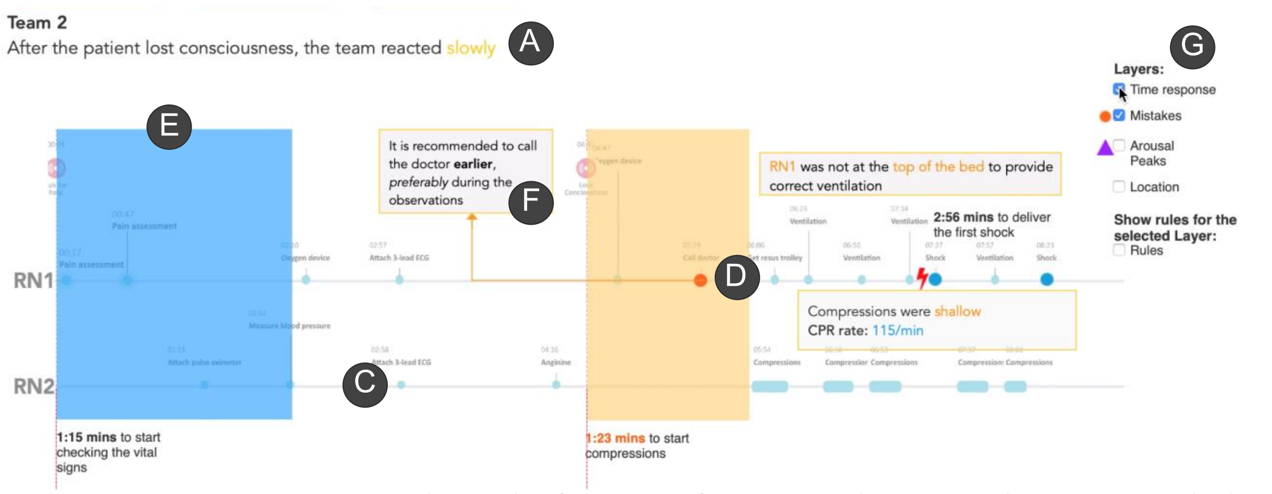


Figure 7.4: Layers *Response time* and *mistakes* for a team of two nurses showing DS elements to encode the different learning intentions of the simulation.

7.4.1.2 *Formalising learning intentions*

Findings from the interviews held with teachers (see Section 4.5. and Section 5.6) provided insights into the learning context, and helped to define the learning intentions (LI) to be displayed in the explanatory layered interface. Teachers proposed that providing evidence of the response time and mistakes would be a great opportunity for students to reflect upon their performance. In addition to this, during the debriefing session it is normal for teachers to ask specific questions about emotional feelings (e.g. *how did it feel participating in the simulation in your role?*). Eleven learning intentions were defined from teachers' inquiry and the learning design script (see Appendix A.1):

- LI1: The team should perform a set of vital signs in a timeframe of 2 minutes, since the patient asks for help.
- LI2: The team should perform CPR in less than 30 seconds, since the patient loses consciousness.
- LI3: Overall, the team should respond timely to both critical incidents.
- LI4: The team should call the doctor during observations and after the patient asks for help.
- LI5: The oxygen device should be attached before the pulse oximeter, during observations and after the patient asks for help.
- LI6: An ECG test should be done during observations and after the patient asks for help.
- LI7: The compression depth should be between 5 and 6 cm.
- LI8: The CPR rate should be between 100 and 120 per minute.
- LI9: The nurse who is providing airway assistance should be located at the top of the bed during compressions.
- LI10: Nurses should show any arousal state during the activity.
- LI11: According to the role enacted nurses should cover different zones of the room.

7.4.1.3 *Mapping learning intentions into high-level rules and DS elements*

High-level rules were created to encode the learning intentions into visual elements. Examples in Figure 7.4 and Figure 7.5 illustrate how the rules encoded DS design elements (Table 7.1) on each layer according to the pedagogical intentions. Figure 7.4 shows how a teacher selected two layers (G- mistakes and arousal peaks). Key insights obtained from the learning intentions list were communicated, such as RN2 not presenting any arousal peak (L10) encoded in the prescriptive title (A), nurses missing to call the doctor during observations (LI4) displayed as a narrative (F) and performing a set of vital signs check during the first two minutes (LI1) as enclosure and narratives (E).

Table 7.1: Data storytelling (DS) design elements considered for this study.

Visual elements	Description
A – Descriptive title	a succinct message that is intended to explicitly communicate the main message of the visualisation
B – symbols	visual element to categorise data points
C - changes in colour/opacity	pushing everything in the background by setting the opacity to low. This allows to focus on insights by adding visual elements with brighter colours
D – highlighting critical data points	visual element that can help the audience to focus on the data that is relevant to the current story
E - enclosure	enclose those data points associated with the same insight
F - narratives	text in the form of labels can help to gain a better understanding of or explain interesting changes in the data

Below, a description of each layer is presented with the corresponding mapping from learning intentions to rules to visual elements.

Response time

Table 7.2 and Table 7.3 lists the learning intentions (LI1 to LI3) that were translated into rules and encoded into visual elements for the layer *response time*.

In ***Rule 1,*** a narrative is added to indicate the time the team took to check vital signs (e.g. Figure 7.4, *"1:15 mins to start checking the vital signs"*).

In ***Rule 2***, a shaded area is used to highlight, either if the response time of the team was in the allowed time (blue) or not (orange) to check vital signs, according to LI1. For instance, if the difference between the time when the action *measure blood pressure* was recorded and the time when the patient *asked for help* is less than 2 minutes: 1) the colour of the time should be blue; and 2) the visualisation should display a *blue shaded area*, from the time when the patient *asked for help*, to the time when the action *measure blood pressure* was recorded (e.g. see Figure 7.4, E). On the contrary, if the difference between the time when the action *measure blood pressure* was recorded and the time when the patient *asked for help* is greater than 2 minutes: 1) the colour of the time should be orange; and 2) the visualisation should display an *orange shaded area*, from the time when the patient *asked for help*, to the time when the action *measure blood pressure* was recorded.

Table 7.2: Translating rules 1-4 into visual elements for the response time layer

Layer	Learning intentions	Rules	Visual elements

Response time	LI1: The team should perform a set of vital signs in a timeframe of 2 minutes.	Rule 1:
		
		Rule 2:
		`if time[measure blood pressure] – time[ask for help] < 2 mins:`
		`highlight_narrative_time("blue")`
		`display(shaded area, "blue", time[ask for help], time[measure blood pressure])`
		`else:`
		`highlight_narrative_time ("orange")`
		`display(shaded area, "orange", time[ask for help], time[measure blood pressure])`
			Narratives:
			Text label
			
			Enclosure:
			2D area with transparency
			
			Colour:
			blue: positive action
			orange: negative action

	LI2: The team should perform CPR in less than 30 seconds.	Rule 3:
		`add_narrative(compressions)`
		
		Rule 4:
		`if time[start compression] – time[lose consciousness] < 30 secs:`
		`highlight_narrative_time("blue")`
		`display(shaded area, "blue", time[lose consciousness], time[start compression])`
		`else:`
		`highlight_narrative_time ("orange")`
		`display(shaded area, "orange", time[lose consciousness], time[start compression])`
			Narratives:
			Text label
			
			Colour:
			blue: positive action
			orange: warning action

In ***Rule 3***, a narrative is displayed to indicate the time the team took to start compressions.

In ***Rule 4***, a shaded area is used to highlight, either if the response time of the team was in the allowed time (blue) or not (orange) to perform CPR, according to LI2. For instance, if the difference between the time when the action *start compression* was recorded and the time when the patient *lost consciousness* is less than 30 seconds: 1) the colour of the time should be blue; and 2) the visualisation should display a *blue shaded area*, from the time when the patient *lost consciousness*, to the time when the action *start compression* was recorded. However, if the difference between the time when the action *start compression* was recorded and the time when the patient *lost consciousness* is greater than 30 seconds: 1) the colour of the time should be orange; and 2) the visualisation should display *an orange shaded area*, from the time when the patient *lost consciousness*, to the time when the action *start compression* was recorded (e.g. see Figure 7.4, D).

In ***Rule 5***, four cases were identified to describe the overall performance of the team considering the reaction time (LI3).

Case 1: If the difference between the time when the action *measure blood pressure* was recorded and the time when the patient *ask for help* is less than 2 minutes and; if the difference between the time when the action *start compression* was recorded and the time when the patient *lost consciousness* is less than 30 seconds, then the *descriptive title* of the visualisation will be displayed as "*The team reacted quickly*" and the colour of the text "*quickly*" should be blue.

Case 2: If the difference between the time when the action *measure blood pressure* was recorded and the time when the patient *asked for help* is greater than 2 minutes and; if the difference between the time when the action *start compression* was recorded and the time when the patient *lost consciousness* is greater than 30 seconds, then the *descriptive title* of the visualisation will be displayed as: *"The team reacted slowly"* and the colour of the text "*slowly*" should be orange.

Case 3: If the difference between the time when the action *measure blood pressure* was recorded and the time when the patient *asked for help* is less than 2 minutes and if the difference between the time when the action *start compression* was recorded and the time when the patient *lost consciousness* is greater than 30 seconds, then the *descriptive title* of the visualisation will be displayed as: *"After the patient lost consciousness, the team reacted slowly"* and the colour of the text "*slowly*" should be orange (e.g. see Figure 7.4, A).

Table 7.3: Translating rule 5 into visual elements for the response time layer

Layer	Learning intentions	Rules	Visual elements

Response time	LI3: Overall, the team should respond timely to critical actions	Rule 5:
		`if time[measure blood pressure] – time[ask for help] < 2 mins and time[start compression] – time[lose consciousness] < 30 secs:`
		`add_title("the team reacted quickly")`
		`highlight_keyword("quickly", "blue")`
		
		`if time[measure blood pressure] – time[ask for help] > 2 mins and time[start compression] – time[lose consciousness] > 30 secs:`
		`add_title("the team reacted slowly")`
		`highlight_keyword("slowly", "orange")`
		
		`if time[measure blood pressure] – time[ask for help] < 2 mins and time[start compression] – time[lose consciousness] > 30 secs:`
		`add_title("After the patient lost consciousness, the team reacted slowly")`
		`highlight_keyword("slowly", "orange")`
		
		`if time[measure blood pressure] – time[ask for help] > 2 mins and time[start compression] – time[lose consciousness] < 30 secs:`
		`add_title("After the patient asked for help, the team reacted slowly")`
		`highlight_keyword("slowly", "orange")`
			Prescriptive title

Case 4: If the difference between the time when the action *measure blood pressure* was recorded and the time when the patient *asked for help* is greater than 2 minutes; and if the difference between the time when the action *start compression* was recorded and the time when the patient *lost consciousness* is less than 30 seconds, then the *descriptive title* of the visualisation will be displayed as: *"After the patient asked for help, the team reacted slowly"* and the colour of the text "*slowly*" should be orange.

Mistakes

Table 7.4 lists the learning intentions (LI4 to LI9) that were translated into rules and encoded into visual elements for the layer *mistakes*. Mistakes were highlighted using a rectangle with a narrative (message according to the mistake done) and changing the colour of the data point related to the mistaken action.

Table 7.4: Translating rules 6-12 into visual elements for the mistakes layer

Layer	Learning intentions	Rules	Visual elements

Mistakes	LI4: The team should call the doctor during observations	Rule 6:
		if time[call doctor] > time[lose consciousness]:
		    add_narrative(call_doctor)
		    highlight_datapoint(call_doctor, "orange")
			Rectangle:
			2D area with transparency
			
			Colour encoding:
			blue: positive action
			orange: negative action
			
			Narratives:
			Text label

	LI5: The oxygen device should be attached before the pulse oximeter	Rule 7:
		if time[attach pulse oximeter] > time[oxygen device]:
		    add_narrative(pulse_oximeter)
		    highlight_datapoint(pulse_oximeter, "orange")	

	LI6: The ECG test should be done during observations.	Rule 8:
		if time[attach 3-lead ECG] between(time[start compression], time[stop compression])
		    add_narrative(3-lead ECG)
		    highlight_datapoint(ECG, "orange")	

	LI7: The compression depth should be between 5 and 6 cm	Rule 9:
		if compression_depth < 5:
		    add_narrative("Compressions were shallow")
		    highlight_keyword("shallow", "orange")	

	LI8: The CPR rate should be between 100 and 120 per minute.	Rule 10:
		add_narrative("CPR rate:", number_rate)
		Rule 11:
		if number_rate < 100 and number_rate > 120:
		    highlight_keyword(number_rate, "orange")
		
		if number_rate > 100 and number_rate < 120:
		    highlight_keyword(number_rate, "blue")	

	LI9: The nurse who is providing airway assistance should be located at the top of the bed.	Rule 12:
		if location_RN != "top of bed" between(time[start compression], time[stop compression]):
		    add_narrative(airway)
		    highlight_keyword(RN, "orange")
		    highlight_keyword("top of the bed", "orange")	

Rule 6 highlights if the team called the doctor during observations (LI4). For instance, if the time when the action *call doctor* was recorded is greater than the time when the patient *lost consciousness*, then a *narrative* is displayed: *"It is recommended to call the doctor earlier, preferably during the observations"* (Figure 7.4, F) and the *data point* rendered in the timeline of the nurse that performed that action is highlighted with an orange colour (Figure 7.4, D).

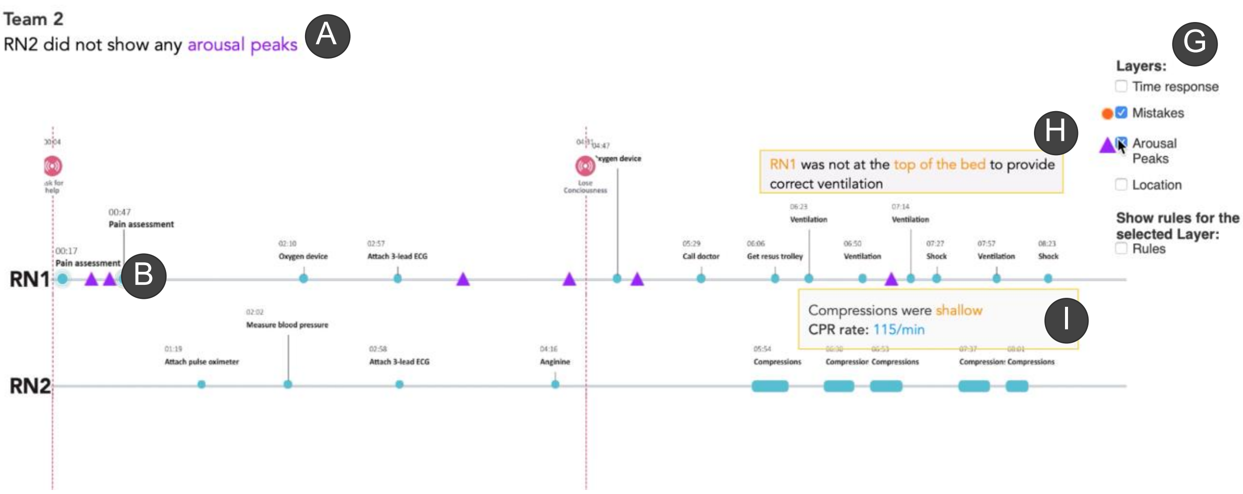
Rule 7 indicates whether the oxygen device was attached before the pulse oximeter or not (LI5). For instance, if the time when the action *attach pulse oximeter* was registered is greater than the time when the action *oxygen device* was registered, then a narrative is displayed: "It is recommended to attach the pulse oximeter before providing oxygen therapy" and the *data point* rendered in the timeline of the nurse is highlighted with an orange colour.

Rule 8 highlights if any team member performed the ECG during patient's observation (LI6). For instance, if the time when the action *attach 3-lead ECG* was registered is between the time when the action *start compression* and the action *stop compression* were registered, then a narrative is displayed: "*It is recommended to monitor the cardiac rhythm with the 3-lead ECG during observations and not during compressions*" and the *data point* rendered in the timeline of the nurse is highlighted with an orange colour.

Rule 9 emphasises if team members did a correct chest compression depth during CPR (LI7). For instance, if the average of compressions depth registered by the manikin is less than 5cm[22], then a narrative is displayed: *"Compressions were shallow"* and the colour of the text "*shallow*" is orange (Figure 7.5, I).

Rule 10 indicates the CPR rate of chest compressions. By adding narrative: *"CPR rate: number_rate/min"* This is to make this value explicit (Figure 7.5, I).

Rule 11 highlights if the average of the CPR rate was between 100 and 120 compressions per



minute (LI8). For instance, if the average of CPR rate is less than 100 and greater than 120 (per min)[23] during the duration of the CPR, then the colour of the text indicating the rate is orange. On the contrary, if the average of CPR rate is greater than 100 and less than 120, then the colour of the text indicating the rate is blue (Figure 7.5, I).

In ***Rule 12,*** the position of the nurse providing airway assistance is highlighted if she was not located in the right spot, which is the top of the patient's bed (LI9). For instance, if the location of the nurse is different from *"top of bed"*, where the location of the nurse should be estimated between the time the action *start compression* and the time *stop compression* were recorded, then a narrative is displayed: *"RN# was not at the top of the bed to provide correct ventilation"* and the colour of the text *"RN#"* and *"top of the bed"* is *orange* (Figure 7.5, H).

Arousal

Table 7.5 lists the learning intention (L10) that was translated into a rule and encoded into visual elements for the layer *arousal.*

Rule 13 was created and linked with DS elements to encode L10. Arousal peaks from nurses were shown as purple rectangles in the timeline (Figure 7.5, B). Also, if any of the nurses did not show any arousal peak, then a *descriptive title* of the visualisation is displayed as: *"RN# did not show any arousal peaks"* and the colour of the text "*arousal peaks*" is purple (Figure 7.5, A).

Table 7.5: Translating rule 13 into visual elements for the arousal layer

Layers	Learning intentions	Rules	Visual elements

Arousal Peaks	LI10: Nurses should show any arousal state during the activity.	Rule 13:
		`if arousal_peaks(RN) = 0:`
		`add_title("RN# did not show any arousal peaks")`
		`highlight_keyword("arousal peaks", "purple")`
			Prescriptive title
			
			Colour:
			purple: highlight peaks
			
			symbols:
			triangles

Location

An additional layer was generated using location data from each nurse to show how each nurse covered different zones of the room (LI11). Location data was visualised as a heatmap, from light to dark shades of blue colours, meaning that areas where nurses were located most of the time were shown in a dark blue colour. Different heatmaps were generated according to the two phases of the simulation to compare if there were any differences or similarities on the zones covered on each phase.

---

[23] This range is recommended for all ages, as stated in the ANZCOR guidelines (Anzcor, 2016)

7.4.2 Participants

Eight experts (1 male, 7 female), most of them educators, with vast experience (7-20 years) in nursing simulations participated in this study. Seven of them were teachers at the Faculty of Health. Two (E4, E8) were graduate students at the same faculty. All experts had worked as Registered Nurses. E1 and E3 were full professors, with extensive experience in the use of information technologies to improve clinical practices and the quality of teaching and learning. E5 was director of Health Simulation, with teaching experience on simulation, patient safety and clinical reasoning related subjects. E2 and E7 were both lecturers, specialised in clinical simulation environments. E6 was a Senior Research Fellow and a Lecturer, who had led multiple teaching and learning projects, focusing on student learning and the development of professional capabilities. E4 was a second-year postgraduate student and lecturer, who had experience in paediatric care, and who is currently investigating the role of errors and mistakes to develop student's critical thinking skills. E8 was a professional skilled nurse who had led research projects to investigate better ways of providing feedback to students about their clinical skills (e.g. CPR techniques). Table 7.6 lists the demographics of the eight participants. Each expert will be referred as E1 to E8.

Table 7.6: Demographics of the eight experts in nursing simulations

Expert	Experience (years)	Gender	Role
E1	30	M	Professor, Medical-Surgical Nursing and Research in Health
E2	11	F	Lecturer, Simulation Manager
E3	26	F	Professor, Nursing Education and Discipline Lead for Nursing
E4	13	F	Registered Nurse, PhD student, Lecturer
E5	7	F	Director of Nursing (Health Simulation), Registered Nurse
E6	7	F	Senior Research Fellow, Registered Nurse, Lecturer
E7	12	F	Lecturer, Subject Coordinator Medical-Surgical Nursing
E8	10	F	Registered Nurse

7.4.3 Method

This study was conducted using LATEP, an elicitation protocol for understanding how non-data experts envisage the use of LA systems (Martinez-Maldonado, Echeverria, et al., 2019). Based on this, this study aimed at investigating: A1) the *added value* of making multimodal data visible through data stories; A2) the anticipated in-classroom *orchestration* of the tools; and A3) the potential impact on students' *accountability*. Inspired by the growing interest on explainable AI (Wang et al., 2019), this study helped to explore A4) the implications of *exposing the algorithms* used for crafting the stories.

Each expert participated in a semi structured interview of approximately 60 minutes, using a think-aloud protocol (Hanington & Martin, 2012). The interview was divided into four tasks:

- For the *first* task, experts were asked to describe the behaviour and performance of the three teams according to the information presented in the timeline of actions (baseline), without DS elements (e.g. Figure 7.3). With this, experts got familiarised with the objectives, outcomes and student's roles of the simulation.

- For the *second* task, experts explored the explanatory layered interface. They were asked to navigate through the different layers of information (e.g. layers presented in Figure 7.4 and Figure 7.5) at their convenience, without a specific order or combination. Also, experts were asked to explain the behaviour and performance of the team as they were inspecting the layers. Questions about *i)* the value of each layer and the *ii)* usefulness for combining layers were asked to trigger discussion with the experts (A1).

- A *third* task involved the completion of a short questionnaire using a 5-point Likert-scale to quantitatively measure their perceived usefulness of each layer. Experts were also encouraged to comment on each answer (A1).

- In a *fourth* task, experts were asked to review the rules that were designed to highlight or emphasise information relevant to the learning intentions (e.g. Figure 7.6). Three questions were asked to the experts to explore: *i)* who should define rules, *ii)* who should see the rules and *iii)* flexibility to change the rules (A4).

- Finally, the last part of the interview consisted in discussing the orchestration opportunities of the explanatory layered interface with regard to: *i)* how experts anticipate the use of the explanatory layered interface; *ii)* how the explanatory layered interface can be implemented in

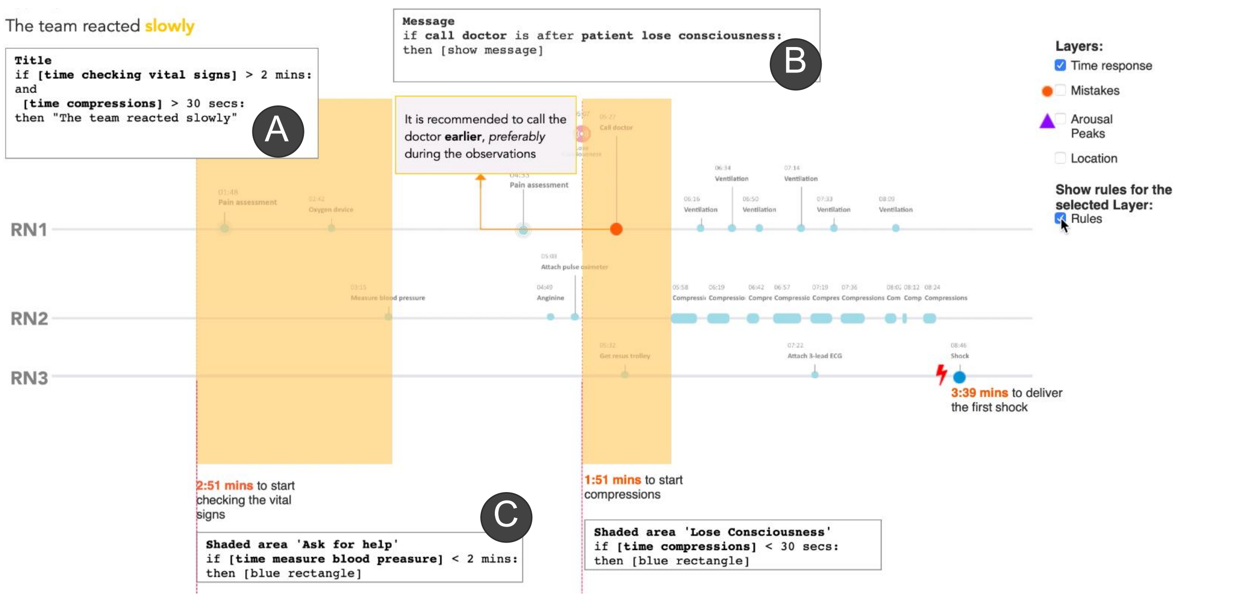


Figure 7.6: Layer that exposes the rule-based algorithm used to add enhancements (e.g. A- a prescriptive title and B- an annotation) to layer time responsiveness of the prototype.

the classroom (A2); and *iii)* who should be able to access the explanatory layered interface for a given team (A3).

The explanatory layered interface was displayed using a Safari Browser on a 21-inch LCD display, connected to a laptop running Mac OS 10.14. Video recordings of interviews and responses to the questionnaires were gathered and organised for subsequent analysis. All video recordings were transcribed using a transcription service.

7.4.4 Analysis

Experts statements and their actions exploring the prototypes were examined. Following best practices of qualitative research in HCI (McDonald, Schoenebeck, & Forte, 2019, p. 13), and given the direct alignment between the study protocol and the analysis themes, statements of interest were jointly coded (Braun & Clarke, 2006) by two researchers according to the pre-set themes of the study protocol (Martinez-Maldonado, Echeverria, et al., 2019) : a) *algorithmic transparency* ; b) *added value of the layers*; c) *anticipated classroom use*; d) *accountability*; and. Then, resulting coded statements were examined by researchers, who had several discussions to select instances that effectively illustrate opportunities and concerns of the approach to create MMLA interfaces that communicate insights.

7.5 RESULTS AND DISCUSSION

This section examines evidence from the study, organised by themes, that reflect the aims of the studies.

7.5.1 Algorithmic transparency

This section describes experts' responses to the Rules Layer (Figure 7.6), and their suggestions about the design of rules to associate the learning intentions with visual elements, as a way to drive attention to specific aspects of the team's performance. As shown in Figure 7.6, users can reveal the rules layer to see pseudocode IF… THEN… conditional rules driving the DS elements.

Experts were asked about a) who should define these rules, b) who should see these rules and; c) the flexibility to change the rules.

According to experts, rules should be defined by the subject coordinator or the tutor who is teaching the simulation. Also, it emerged that the rules could be defined by the learning design and objectives of the simulation. For instance, E3 commented that *"a tutor may want five key things that matter most in a simulation, so they could be rules related to each of those five things"*. This is also consistent with the approach followed to design the rules (described in Section 7.4.1).

Interestingly, all experts agreed that tutors and students should be able to see the rules, as an additional option in the interface. One expert suggested that rules could be *"turned on or off"* during debriefing, allowing flexibility for using them.

Experts considered that visualising these rules during debriefing could guide the discussion with students by comparing team's performance and the learning expectations. One expert pointed out that: "*It might be useful to say during a debrief: here is the time taken for this action and these are the rules that have been set to that. Do you think that's a reasonable performance?"* [E1].

Moreover, experts indicated that if the rules are shown to students, these could be used as a confirmation of students' expectations and exhibited performance. Experts argued that *"students have a timeframe in their mind"* [E7] and that *"students have a clue on how to fulfil a task"* [E2]. However, they acknowledged that most of the time, students do not have the perception of time and the critical situation they are handling. For example, E7 noted that "*students had no idea of the time and how urgent the situation was"*. Thus, these rules could be used as an opportunity for students to recognise the gap in their practice by comparing their performance with a baseline information. As one expert suggested: "*These rules are a good reminder [for students] to compare their steps and the processes they did and what was required in the simulation"* [E8].

Two different perspectives emerged from experts' suggestions. The majority of experts suggested that tutors or subject coordinators should be able to adjust the *parameters* or *thresholds* for rules. T1 expressed that: *"It would be good if there was flexibility so the course coordinator, for example, could adapt them slightly"*. They suggested that rules' thresholds could change according to the level of experience of students. This was expressed by E6 as follows: "*I would expect third-years get their things done perfectly, like, knowing what they are doing, more so than a first-year"*.

On the contrary, two experts agreed that rules should remain the same. E7 indicated that rules *"shouldn't change to accommodate the student"*. Instead, these rules should be seen as a teaching practice to help students to achieve a particular goal. She also explained that rules would construct awareness on students about what is expected from a registered nurse from the very beginning: *"The rules will be like: this is what you need to aim for, this is what needs to be achieved*". This guidance provides a valuable opportunity for students to learn and reflect what they should change to achieve a better performance. Another idea that emerged during the discussion was that rules should be defined by international guidelines on how nurses should react quickly and effectively to save patient's life. E8 commented that this is critical for nurses to notice: "*It's the potential between saving a life and not saving a life. This is what's needed to be done and this is how quickly you need to do it."*

7.5.2 Added value of layers

Overall, all the experts agreed that the layers added significant value to the current timeline version (timeline of actions, Figure 7.3), where only actions were presented. E3 expressed this as follows: *"[The multimodal layers are] absolutely more valuable, because it tells you the story behind the simulation"*. Likewise, E7 made a similar comment on how all the layers allowed her to have *"a good picture"* of team's performance, specifically when exploring the time response: *"These [the multimodal layers] give me the boundaries and the timeframe as well. That's good because then you know this is how long they took. All of these layers work really well together to give you a really good picture"*. Consistent with the social translucence principle reviewed in Section 5.6, E8 stated that *"these layers keep people more accountable"*.

Experts suggested that the combination of layers offers a great opportunity to discuss a team's performance in more dimensions. The freedom to select one or more layers of information at the same time could give teachers a tailored view of each team's performance. This can lead to richer discussions by focusing on specific needs for each team. As stated by E2, *"If the team exhibited a low response time, and did many mistakes, then the discussion would be around the both layers, trying to find an explanation of that behaviour. I wouldn't need the arousal*".

Significantly, it emerged that the explanatory layered interface offers the possibility of supporting one-to-one coaching to teams. E8 suggested that "*All of the components of the interface provide tutors the opportunity to customise their own feedback for coaching teams"*.

Consistent with the findings from teachers' reactions described in Section 5.6.5, it is clear that a single layer could not provide enough information to interpret a team's performance. Showing different layers of information and allowing teachers to use them according to students' needs, clarifies the potential of developing customisable feedback for guiding and supporting reflection on team's performance. During the debrief, the teacher may use the customisable feedback to intervene when the team performance does not match with their expectations or the objectives of the

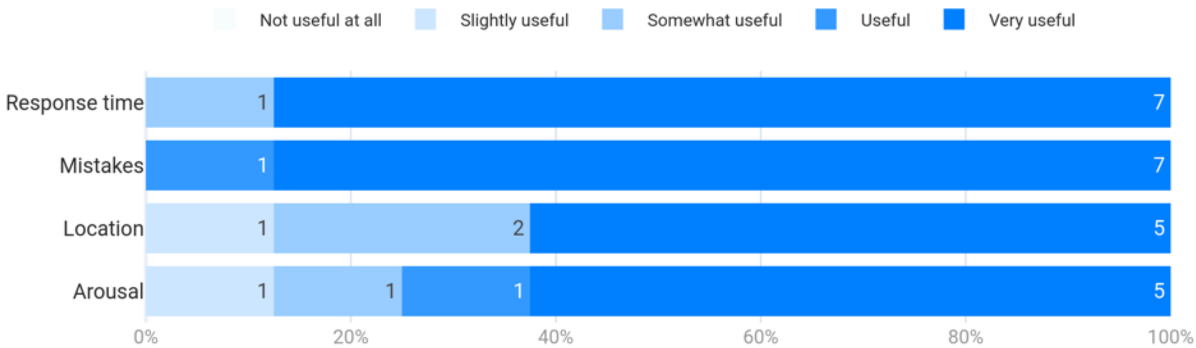


Figure 7.7: Perceived usefulness of each multimodal layer.

simulation.

The results from the questionnaire about the usefulness of each layer indicated that *response time* and *mistakes* layers are very useful to support debriefing and to provide a personalised coaching to teams; and *location* and *arousal* layers are useful at a lesser extent, as portrayed in Figure 7.7. Next, a discussion for each layer and its perceived value is presented, from the most valued to the least valued, according to questionnaire results.

7.5.2.1 *Response time layer*

Most of the experts agreed that the response time layer was very useful (7 experts) and somewhat useful (1 expert; Figure 7.7). During the exploration of this layer, three experts stated that time is "*critical*" and "*crucial*" when managing critical situations. For example, E8 stated that "*Time is a crucial factor within managing cardiac arrest. So, time to detection, time to compression, time to ventilation, time to shock, time is a really important component*". Also, the response time layer offered a **detailed view about delays** occurring during the simulation. One expert noted this as follows: *"So, there's been some pain assessment being done but it has actually taken a bit of time [for them] to actually get things going. And, again, way too long to start compressions. So, this was way too late"* [E1].

Interestingly, some experts discussed how the response time layer helped them to **highlight other behaviours** (e.g. team leadership strategies) by comparing the actions shown in the baseline timeline with the response time displayed in the superimposed layer. For example, one expert first inspected the actions made by each nurse simultaneously with the response time, to then raise concerns about the nurse in the team leader role: *"There was no other delegation in here [pointing to a specific part of the explanatory layered interface of Team 1] because there were no activities"* [E8]. Another expert made a comment in relation to what would be the ideal scenario for the team leader after inspecting the response time layer: *"It just takes the team leader to say, right, I want you to do this, and explain it to the patient as they're going. So, you're actually getting all that that done within the first 60 seconds"* [E7]. Thus, the clinicians were able to 'read' the timeline in quite sophisticated ways, using the presence or absence of events, within a timeframe, to infer the competence with which a particular role was played. This exemplifies how the visualisation could support tutors in leading an evidence-based debrief, or with suitable scaffolding, a personal reflection task assigned to students.

Some experts used the layer to explain **team's responsiveness** by arguing that student's previous knowledge and expertise could impact the quality at which him/her react quickly and positively to unexpected events. One expert [E6] confirmed this as follows: *"In a group of four, for instance, you've got your one or two Type A's [excellent students] that have done most of it. You know, they can do a set of [observations] in like a couple of minutes. For another student, vital signs can take so long"*.

However, it was also discussed that depending on the **simulation and their learning outcome**s, response time layer might not be useful: *"Depending on the simulation I think time response can be a useful tool, like this is obviously a cardiac arrest simulation. But for other simulations may not necessarily be as important."* [E5]. This suggestion offers a reflection point on the approach followed to design and implement the explanatory layered interface. Besides teachers' and students' perspectives, the pedagogical intentions should be aligned with the information provided in the visualisation. The same expert reflected on this by suggesting that *"time response would be purposeful depending on the simulation"* [E5].

Time management and team responsiveness are critical in a life-threatening situation (as reviewed Chapter 4). Yet, in real practice is hard to make these constructs and behaviours visible. Thus, this layer has the potential to provide awareness of the notion time, which can be used to understand team's performance (e.g. delays in performing a critical action).

7.5.2.2 *Mistakes layer*

As depicted in Figure 7.7, experts indicated that this layer was useful (1 expert) and very useful (7 experts). Some experts mentioned that this layer **highlighted mistakes at glance**, in relation to the sequence of correct actions and missing actions. For example, one expert mentioned that "*It is important to identify where [the team] may not have done things in [the right] order*" [E5]. Highlighting mistakes could be helpful for students to reflect on their practice. One expert commented that this layer shows "*how [students] can improve their practice*" [E7] based on the information provided in this layer: "*It's good for them to know that the compressions needed to be deeper, based on what is shown here*" [E7]

Consistent with teachers' and students' needs (as described in Chapter 4) some experts found this layer very helpful to retrospectively **analyse a team's performance** according to the actions that teachers cannot directly observe or provide a proper amount of attention, either for being in another physical location (e.g. *"if you were looking at it [multimodal layer] and you haven't been there [during the simulation], then it's helpful for me to see that"* [E2]); or due to technical challenges to access to the information promptly (e.g. *"for example the depth of compression. I can guess that they're not going deep enough, but I'm not in the manikin."* [E7]).

However, one expert expressed **concerns in prescribing these mistakes**, if the tool is not capable of highlighting all of them (false negatives). E3 noted that this layer could misguide the actual team's performance: *"if the layer did not highlight all the mistakes, it could make students feel that they performed very well when in fact they missed critical things"*. Again, this critique is consistent with what was discussed before, regarding the alignment of the pedagogical intentions and the information. Researchers and designers should always aim to include teachers' and students' requirements to justify the selection of the information presented in the explanatory layered interface, but where this is not possible — as may well be the case when designing analytics for higher order competencies — such tools need to 'design for imperfection' (Kitto, Shum, & Gibson, 2018).

7.5.2.3 *Location layer*

The majority of experts rated this layer very useful (5 experts), whereas two rated it somewhat useful and one slightly useful (Figure 7.7). This layer helped experts to understand students' behaviour in relation to the **engagement with the activity**. For example, E8 noted that a student in Team 3 had a limited engagement due to her position, which was relatively far from the bedside: "*RN3 was not even at the bedside really and that's reflective of their engagement through this scenario"*. Also, E7 noticed the same behaviour from the same RN3 by focusing on the large spot in the location colour map, far from the bedside during the critical emergency (i.e. patient lost consciousness): "*RN3 was just standing there, probably just watching"*. Interestingly, some experts inspected simultaneously the location layer and the actions in the timeline to compare if the actions performed by each student correlated with the space occupied. For example, E4 noted that *"This layer gives you an idea of all the movement and where each of the RNs were in relation to an action"*. Similar, E6 noted that "*you see who is ventilating, who is doing compressions. It actually shows you what activities they're doing"*.

Most of the experts agreed that showing the location during debriefing could provide a great opportunity for teams to reflect on their **correct positioning**. Depending on the role taken, students should stand or move to appropriate places in relation to the action being performed. Therefore, this layer offers the opportunity for teachers to trigger a discussion around correct positioning. One expert commented this as follows: *"From a debrief perspective, I would be saying that the scribe [RN4] needs to be at the foot of the bed so they can see everything that's going on"* [E1]. This means that from this view, the student could monitor all the actions being performed by other students clearly and could anticipate if her help is needed. Also, one expert mentioned that this layer could be useful to identify poor behaviours from students in relation to the action being performed and the location. For example, E5 suggested that: *"You can potentially show to a student that she was trying to provide care to the patient without actually moving into the bed space"*.

However, one expert argued that this layer was not really useful, considering that the simulation had already **prescribed roles** and that some actions were expected from students: *"In this type of simulation, it is not really useful, because it's quite predictable. If it was an unscripted simulation, you'd see people go everywhere and I think it'd have more meaning"* [E3].

While this layer did not prescribe any mistake in relation to the correct positioning, it is worth noting that a more compelling story can be articulated by having this extra information. This suggests that, while video recordings could give teachers more details about positioning of students, a higher level of data abstraction could provide similar insights, which may potentially address concerns raised about data privacy (see Subsection 7.5.3).

7.5.2.4 *Arousal layer*

Most of the experts found this layer very useful (5 experts) and useful (1 expert). However, two of them were less convinced about the usefulness of this layer (Figure 7.7). The interviews with the experts revealed both benefits and concerns.

Some comments around the benefits suggested that this layer might be useful to **reflect on *imperceptible* behaviours**, such as feelings: "*It gives you insights into how the person is feeling"* [E7]. Also, it could trigger questions to explore in depth why a student exhibited stress (e.g. "*was there something at this particular point in time that caused you to feel stressed?* [E2]); or if the arousal state was caused by the situation (e.g. *you were doing a pain assessment and then you were about to put on the oxygen delivery device. What was the trigger there?* [E1]). Finally, it could help teachers to reflect if students perceived the certain actions as a stressful situation: *"educators would be aware of where that stress point was for [students]"* [E3].

However, some concerns were expressed about its usefulness, considering the **accuracy of the device** to capture all arousal states and showing back all this information. E6 and E8 indicated that this layer can be impractical if this is not showing all arousal states, as they were expecting more arousal peaks: *"The arousal doesn't match with what I would expect"* [E6]; and a correlation between actions and arousal peaks: *"I'd expect to see peak before they do anything."* [E8]. In addition, E3 and E6 pointed out that showing these arousal states without certainly knowing the reason why there was a peak could cause misinterpretations about student's behaviour: *"You need to think that [students] are so calm, or you could think that they're completely not interested. There's no gauging on what is going on"* [E6].

In brief, each layer motivated experts to explore different aspects of team's performance. Some benefits and concerns were suggested by experts in terms of each layer of information. The next subsection explores the value of combining layers.

7.5.3 In-classroom orchestration opportunities

The explanatory layered interface could be used as a **post-hoc reflection tool** to support debriefing. The evidence provided by the explanatory layered interface could offer an opportunity to have a better debrief experience. As discussed previously, some layers have the potential to build a *translucence* effect on often imperceptible behaviours such as "time perception", "stress" or "emotional states", which had been previously invisible to teachers. For example, during the inspection of the arousal layer (e.g. Figure 7.5), experts reflected that simulations should provoke a stress response of some sort. Therefore, if students did not show any arousal peak, teachers should investigate the reason for this particular behaviour to give an inform explanation. As stated by E5, "*if [students] have no arousal peaks, I would ask for reasons: What were you actually thinking? Why you were not engaged with that? We, as teachers need to let students understand that that is okay.*" Similarly, E7 commented that "*Teachers could do a more informed debrief, drawing out more information from the students: How [students] were feeling? What they were thinking during the process?*". However, another expert [E5] revealed that, because of the stress that simulations provoke, students react differently. Thus, the Timeline could help students to "*alleviate some of their anxiety by actually showing them that your team was a little bit quicker to get things done*".

The explanatory layered interface offers **objective evidence** to better inform students about their actions and give tailored feedback. As E8 expressed: "*It's in black and white. This is what you did well, this is what you did poor. This is what you can improve on in a real-life situation, you do X, Y and Z*". Thus, the explanatory layered interface could be used to "*make bolder statements that cannot be disputed*". This view is consistent with nursing education, which is focused on developing an evidence-based culture to support reflection on their practice (Ireland, 2008). For instance, the same expert [E8] mentioned that "*this why objective data on clinical scenarios is so important because it is not just someone subjectively giving their opinion. It's providing evidence*". Also, this objective evidence could be used to keep a record on individual's and team performance, allowing students revisiting this information in further simulations. As expressed by E5: *"[Using explanatory layered interface] some students really would remember exactly how they would do it differently next time"*.

With a structured debriefing, the information presented in the explanatory layered interface and critical questions, students may have fa more opportunity for **deep reflection**. As stated by E3, "*if teachers gave [students] structured reflection questions and this [explanatory layered interface] information and asked them to reflect on these aspects of teamwork, it would be a really valuable tool for deep reflection*". This suggestion is aligned with Dewey's (1997) argument, expressing that reflection involves the observation of the experience retrospectively to find explanations for what happened. The majority of experts offered examples of questions that they could ask students to spark reflective conversations and provoke student's reflection based on the evidence provided by explanatory layered interface. These questions were related to a range of variables, including:

- performance (e.g. *"Am I doing a good job? Am I getting things done?"* [E7]);
- arousal (e.g. *"Can you talk me through what you were feeling in this moment. Can you tell me what you were thinking at this point?"* [E2]);
- actions or mistakes performed (e.g. *"So, let's have a look at these compressions, and as I said, you know, they're too shallow. So, tell me a bit about what depth the compressions need to be at, and have you achieved that? How do you know you're achieving that and the rate of*

compression?" [E2]; *"In what ways do we diagnose someone having a cardiac arrest prior to them losing consciousness?"* [E8]);

- response time (e.g. *"Do you think that that's a definitive time frame to do that in? Or looking at the ABCDEFG algorithm, when would you hope to provide your first shock?"* [E8]);
- wrong positioning during the simulation (e.g. *"What were you doing over here? Because you didn't go and get the resus trolley but what took you over here? And then this person stood the most time off to the side and did nothing. And do you think that's a fair way to behave in a team?"* [E8]);
- unsafe practice (e.g. *"Is this clinically safe for your patient? Are you performing in a way that's safe for your patient?"* [E5]).

The explanatory layered interface could be used to **re-design and plan better simulation experiences**. For example, E1, E2 and E5 noticed that Team 1 and 2 put the oxygen before the oximeter. Therefore, the comparison among many other teams' mistakes could provide an opportunity to draw attention into commonalities in their behaviour. As E5 argued: *"The explanatory layered interface will give us a sense about the real practice or highlight if a lot of groups of students are doing the same errors"*. This could lead to highlight and reflect on not only deficits in students' learning, but also in their teaching practice. Ultimately, the information presented in the explanatory layered interface could serve as evidence to improve the simulation, addressing the most common errors exhibited by teams. E1 mentioned this as follows: "*what [teachers, subject coordinators] could do is to re-structure that scenario, having a discussion around the concerns highlighted, and the things that are needed to work with teams"*. This is consistent with Lockyer et al. (2013) suggestions, that analytics could support planning and designing new curricula. Usually, teachers use past experience and what they may recall from the classroom session, which can be seen as an informal practice. Hence, a formal practice is needed to allow teachers corroborate and reflect on evidence from past simulations.

Interestingly, some experts identified both benefits and risks of using explanatory layered interface as a summative **assessment tool**. For example, E8 noted that the explanatory layered interface could be used *"as further evidence to determine competency"*. One of the reasons that led to this point of view is the challenge of providing evidence in high-stakes scenarios about team's performance. As E3 expressed: *"one of the issues with high-stakes assessment, is making sure it is valid and consistent; and objective rather than subjective"*. Therefore, it can be arguable that the explanatory layered interface could support this claim. However, E7 argued that explanatory layered interface should be used to provide formative feedback rather than summative feedback. She expressed that "*Simulation is not a (summative) assessment. It is to give formative feedback, which for help students improve their practice"*.

In addition, one expert indicated that explanatory layered interface could be used as a **briefing tool**, to be used as a case study before the simulation with new students. E1 suggested that using the explanatory layered interface could help students to "*get a learned behaviour much more quickly*" by showing good and bad examples of teams doing the simulation and make students reflect on this.

Another theme that emerged was in relation to the orchestration strategies that could be implemented during debriefing. Depending on the learning design of the session, the debriefing may occur *"in-between scenarios*" [E2], or once all teams have finished the simulation, in a "*whole classroom setting"* [E1]. Thus, the Explanatory layered interface could be used for small team discussion or in a wider class approach enabling (possibly anonymised – see next section) Explanatory layered interfaces to be compared. Interestingly, one expert mentioned that while the Explanatory layered interface could be used during debriefing, it should be used at teachers' discretion, considering that students react differently to feedback. E4 argued that: "*It depends on your cohort. [teachers] are going to know how the class is going to cope with this information"*. This was also noticed by another expert, who indicated that "*it might be better to wait and see how they are all feeling*" [E7]. This suggestion hints that, while the Explanatory layered interface could provide a great opportunity for evidencing team's performance, this may not always be perceived as beneficial for students. Some of the threats for the adoption of the Explanatory layered interface could be that students perceive this information as a summative assessment rather than a formative feedback reflection tool. In this sense, designers and researchers should be careful when designing the rules and visualisations, being more informative rather than evaluative. Finally, one expert suggested that the Explanatory layered interface can be implemented in a large study across semesters to compare students' performance. As stated by E1, "*If [teachers] do these scenarios in the different semesters, what does that look like a second-year student performance compared to semester five?*"

7.5.4 Accountability

A consequence of making evidence about teamwork visible is that people can be made responsible for their actions (Erickson & Kellogg, 2000). This was recognised by educators in this study. E8 explicitly stated that: *"these layers keep people more accountable"*. All experts agreed that the explanatory layered interface should be used by teachers to support debriefing and give personalised feedback to teams (at individual and team-level). However, some concerns were raised by experts when discussing if the explanatory layered interface information should be accessed by other teams or external people. From the discussion, it emerged that sharing the information with the ***whole classroom*** would make students feel "*a bit uncomfortable*" [E7]. One expert [E5] expressed that this could be a good opportunity to discuss teams' performance with the whole classroom. Nonetheless, some experts noticed that this could cause a "*comparison of performance"* [E3 and E5] among teams, especially when a team had a "*lower performance"* [E3]. Two different perspectives arose to address this concern. On one hand, this information should be kept as "*confidential*" [E3], where only each team could access its own information. On the other hand, if this information is de-identified, it can be shared with other students, within the same classroom, to avoid criticism. However, E1 found that it is challenging to completely de-identify information, as it can be guessed by students: *"If students know who was being monitored for simulation, they would know who was in Team 1, Team 2 and Team 3."*, that is, even with de-identified information students could guess who was doing what and it could potentially make students feel uncomfortable. E3 cast doubt on the value to have de-identified information: *"I think if it's de-identified, that would be okay, but I'm not sure how valuable de-identified information would be"*. At the same time, some experts also suggested strategies that can be used to address a whole-classroom reflection. For example, E2 mentioned that the tutor could randomly select a team: "*pick a random team and have a look at [the explanatory layered interface information] and see what went well and what did not go so well, and what [students] can improve on next time"*.

Also, one expert remarked that the information generated from students should be accessed by the students and teachers involved in the simulation, and no other students and academics outside the classroom. As E5 expressed: "*I do not go and talk about [this information] with other tutors [outside the classroom]. I do not talk about [this information] with other students [outside the classroom]. I only talk about [this information] with the students of my class.*"

7.6 EMERGING DESIGN FEATURES

Besides the exploration of the aims of this study, experts also suggested some features for improvement, summarised as follows:

- **Social Interaction:** E3 made a comment about the possibility of *"visualising interactions between members"*, either verbal or physical interaction, in order to understand *"how one person's stress is transmitted to other"* and *"how this impact on performance and outcomes"*. Also, E1 suggested a layered information about verbal interaction, to observe *"who was talking at what point in time or who was quiet"* to relate that with particular actions. Visualisations for some of these aspects were previously developed (see Section 5.5.1). However, this was not part of this validation, as the intention of this design was to generate layers of information that could be fully developed in authentic classrooms.
- **Interaction with the patient:** E1 proposed the tool could visualise information about verbal and non-verbal communication with the patient, recommending a patient-centric interface: *"It would be interesting to see who is interacting with the patient, who is talking with the patient? It doesn't matter to some extent what they are saying, but that sort of verbal and non-verbal interaction*". (again, see Section 5.5.1).

- **Incorporate video:** Another feature that tutors were interested to explore was synchronisation with the video recording, since video provides rich information that helps to make sense of the terser event timelines: *"Sometimes you can tell by the video about facial expressions, are they working collaboratively, are they listening, have they got open body posture? All the key predictors of team dynamics."* [E3], and that critical information from other layers (e.g. mistakes, or arousal states) could serve as indexing points of the video. The teachers envisaged being navigating to the relevant point in the video by clicking on Timeline event markers (cf. related work on semantically indexed meeting replay tools; Page et al., 2005).

- **Add a patient layer:** E5 and E6 indicated that it would be interesting to see the patient's status from the manikin data as an additional layer. E5 expressed that *"it would give an additional context of the scenario"* and E6 said that *"[students] could actually watch the scenario unfold"*.

- **Team-based performance report from the timeline information:** A final suggestion was to generate a team-based report, using the information from all layers to generate a narrative textual summary. As E3 suggested, this could be used as evidence and for assessment: *"Based on these and their timing, the team and the tutor and the course coordinator would have information of the overall summary of the performance. What they did well, what they omitted, what they didn't do well"*.

7.7 SUMMARY

This chapter presented the EvisLA approach to design explanatory layered interfaces. Because of the complexity of understanding multimodal data at a glance, this approach applied a "layered" model to design multimodal interfaces. Each layer is intended to communicate insights derived from one construct or data stream. In addition to this, data storytelling principles were considered on this approach to draw the user's attention to key aspects of the simulation. The approach was illustrated through the design and validation of an explanatory layered interface, a tool that could be used during debriefing. to guide and support reflection. This iteration encapsulated all the design concepts in this research project, and as documented, eight subject matter experts gave very positive feedback as they explored the explanatory layered interface.

The concluding chapter now summarises the contributions that this thesis has made, and discusses the implications for future work.

Chapter 8: Discussion and Conclusions

This chapter concludes the thesis by summarising its main findings, and the contributions it makes to Learning Analytics and related research fields (CSCL, CSCW, HCI). Furthermore, the chapter presents guidelines for researchers and designers of multimodal learning analytics interfaces. Finally, the chapter concludes by noting the challenges this kind of research faces, recognising its limitations, and identifying future technical, ethical and methodological directions for this research.

8.1 SUMMARY OF CONTRIBUTIONS

This thesis advances research in group work by supporting the provision of feedback that can be used by teachers and students. This thesis has investigated how explanatory multimodal interfaces can be used to support individual and group reflection about teamwork. To achieve this, the problem was defined in terms of three research questions:

RQ1. How can stakeholders' requirements be formalised to design for multimodal interfaces?

RQ2. How can low-level multimodal data streams be modelled to serve as proxies for group constructs?

RQ3. How can explanatory multimodal interfaces be implemented to guide reflection on group work?

To address these questions, this thesis developed three conceptual approaches, in relation to *designing, modelling* and *visualising* group feedback, by means of explanatory multimodal interfaces. Figure 8.1 summarises the linkage between the research questions, thesis goals and contributions associated with each chapter.

Let us now consider what was accomplished in relation to each Research Question, before distilling the core contributions.

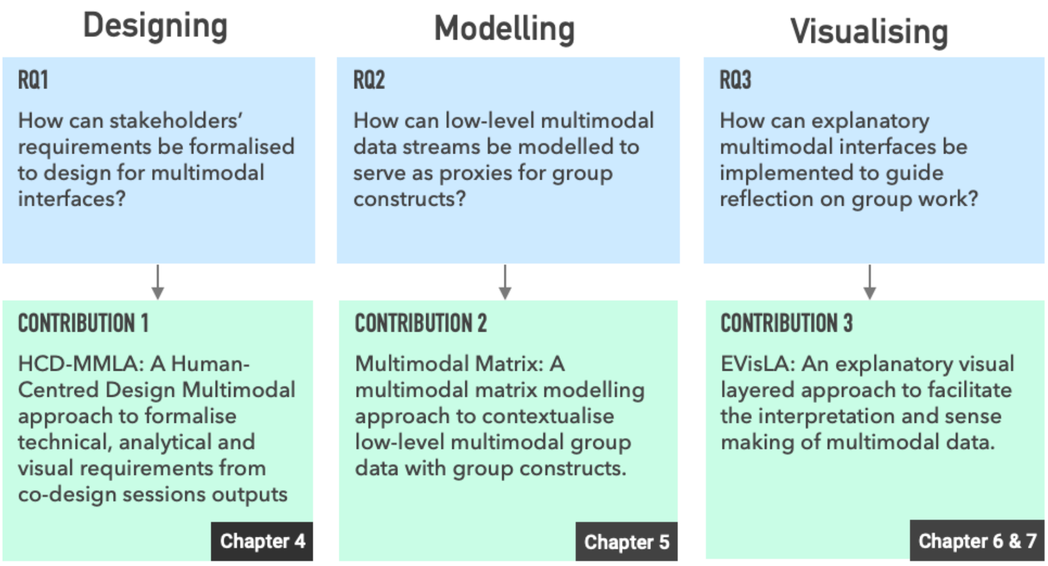


Figure 8.1: Summary of the thesis research questions, goals and contributions (Chapters 4-7).

8.1.1 Main findings for RQ1

How can stakeholder's requirements be formalised to design for multimodal interfaces?

Chapter 4 documented the process of exploiting HCD methods to formalise the stakeholders' requirements for designing a MMLA interface through the conceptualisation of the HCD-MMLA approach. The approach was illustrated as follows:

- Co-design sessions were organised with two subject coordinators and thirteen students from the Bachelor of Nursing program.
- Outputs of these sessions were analysed using the HCD-MMLA approach, first by applying an **inductive content analysis** method for mapping the learning constructs and requirements needed to design a MMLA interface. Second, by applying a **deductive** analysis driven by literature based on pedagogy, evidence and discipline knowledge.
- An outcome of the approach was a **mapping template** to draw explicit relationships among group constructs, subconstructs, behavioural markers, sensors, data/analytics and ways of representing this information.
- Using the mapping template, different prototypes were designed and generated to prove the feasibility of capturing, analysing and visualising multimodal data. The prototypes were implemented using authentic data gathered from two studies, one in a lab setting and the other in an authentic classroom session.

8.1.2 Main findings for RQ2

How can low-level multimodal data streams be modelled to serve as proxies for group constructs?

Chapter 5 reported a systematic conceptual and computational approach to draw on qualitative insights about work practices in order to model low-level multimodal data:

- The conceptualisation of the **Multimodal Matrix** integrated principles from Quantitative Ethnography (QE) methodology and the ACAD framework to systematically encode qualitative insights in a matrix-like data structure.
- The final outcome of the approach was a matrix-like data structure with: (i) columns representing the dimensions of collaboration from the ACAD framework (physical, epistemic, social) and the affective dimension, identified as relevant from the literature; and (ii) rows representing the time (e.g. seconds) the activity unfolded.
- From the matrix-like data structure, four proxies were designed, implemented and validated with four teachers and 23 students. The proxies demonstrated the potential of the Multimodal Matrix to help visualise insights from different dimensions of the group activity.
- An elicitation process was carried out with teachers and students to explore the feasibility of proxies to support awareness, visibility, accountability and orchestration opportunities. Teachers and students valued the proxies for communicating information relevant to a team's performance, rather than all the data captured via sensors and other interactive devices. Proxies were effective in providing insights into the different dimensions of the group activity, separately. Teachers and students reported the need to see multiples modalities in the same visualisation, instead of presenting one modality per visualisation.

8.1.3 Main Findings for RQ3

How can explanatory multimodal interfaces be implemented to guide reflection on group work?

Chapter 6 studied the potential combination of explanatory visual elements with key aspects of the group activity to contextualise learning visualisations. An Educational Data Storytelling approach was conceptualised and illustrated with two studies, as follows:

- The Educational Data Storytelling approach was grounded in current works in Information Visualisation (InfoVis) and Data Storytelling (DS) principles and techniques aimed at communicating insights to different audiences through visual aids.
- A first pilot study to validate the design of explanatory visualisations was carried out using visualisations from the collaborative database design scenario. Five graduate students, with previous experience in collaborative design activities were interviewed to explore the potential of explanatory visual elements. Initial findings indicated that explanatory visualisations added clarity, especially when there were multiple possible stories in a complex visualisation. However, the visualisations should be carefully designed. This led to

more detailed exploration to see if the elements used to explain the visualisation (i.e. data storytelling elements) were correlated with users' visual attention.

- A second illustrative study redesigned the explanatory visualisations from the collaborative database design scenario and investigated the role of each visual element in driving users' visual attention when communicating one story at a time. Quantitative and qualitative data from six teachers were analysed to report their perceptions of the feasibility of DS to support the communicative power of visualisations and recommendations for further designs.
- Evidence from eye tracking data provided a better understanding of the elements that drove teachers' attention during the first minute of seeing a dashboard. Teachers' gaze behaviours when exploring these visualisations were analysed in fine-grained detail in terms of fixations, trajectories and the most viewed areas of interest.
- Eye tracking analysis and teachers' qualitative perceptions confirmed that text labels and key data points were the strongest elements for drawing attention in an explanatory visualisation. Specifically, the DS elements drove teachers' attention on key aspects of the activity rather than having the attention diffused across the whole visualisation.
- This research foresees the potential of learning design-driven data storytelling elements to support sensemaking by guiding the user to "one learning story per visualisation", in order to place the user in control of the rich multimodal data, and prevent disorientation from information overload. It is hoped that both designers and researchers will gain value from this conceptual model as a step towards designing dashboards explicitly contextualised for learning.
- This study showed promising results and motivated the design and motivation of explanatory visualisations for supporting teaching and learning practice in the context of teamwork nursing simulations.

Chapter 7 reported the integration of the previous approaches into one approach. The Explanatory Visual Learning Analytics (EvisLA) approach was conceptualised and illustrated as follows:

- The design of explanatory multimodal visualisations was operationalised by integrating data storytelling and a layered model. In this sense, the EvisLA approach aimed to support teachers and students to focus on "one story per layer". This means that the multimodal data was decomposed in several layers of information, and each layer communicated important insights using DS principles and techniques.

- The EvisLA approach was illustrated in a final capstone study. A high-fidelity prototype was implemented using the multimodal data captured from a teamwork nursing simulation scenario. The prototype was validated with eight subject experts.
- Subject experts appreciated the flexibility to combine different layers of information in a shared view. They mentioned that this feature could support adaptive feedback and guidance during the debriefing session, after the simulation.
- This study showed promising results to operationalise the high-fidelity prototype as a tool to support debriefing, after the simulation has finished. This is in line with the purpose of the tool that was explored in Chapter 4. In addition, subject experts envisaged other potential uses of the prototype, including its use as a tool to reflect on teaching and learning practices in order to design better simulations, and an assessment tool to provide objective evidence to students. They also offered recommendations for further implementations in terms of orchestration opportunities and data privacy.

8.1.4 DBR summary

Considering the Design-Based Research (DBR) process, each conceptual approach and visual representation was illustrated and validated through:

- Authentic empirical studies with experts, teachers and students.
- Use of realistic data collected in the lab and in-the-wild.
- Validation of a technological solution iteratively with experts, teachers and students.
- Various quantitative and qualitative data analysis techniques, such as content analysis, affinity diagram, statistics, clustering, and visualisations.
- A first version of a high-fidelity prototype, a layered interface to communicate one insight per information layer.

The following sections distil the contributions to knowledge that this thesis makes, and the new questions that they open up for future work.

8.2 IMPLICATIONS OF CONTRIBUTION 1: THE HCD-MMLA APPROACH

8.2.1 Discussion and future work

The first contribution of this thesis was the HCD-MMLA approach, a human-centred design approach to formalise group constructs and elicit technical, analytical and visual requirements from co-design session outputs. This contribution can be contextualised in relation to two key challenges now facing the field of LA. Firstly, the growing recognition of LA as needing *human-centred design*

(HCD) methods to engage stakeholders, and secondly, the challenge of translating the *implications of theory* for LA design decisions. These are discussed next.

There is growing interest in the LA research community to adopt HCD processes for the development of LA solutions. For instance, Holstein et al. (2017) reported their reflections on adopting co-design and participatory design techniques to support the design decisions of an Intelligent Tutoring System (ITS). The Eighth edition of the Learning Analytics and Knowledge (LAK18) conference, whose theme was *Towards User-Centred Analytics* research, encouraged researchers to report theoretical, methodological, empirical and technical contributions addressing design challenges of learning analytics tools. Research articles from this conference documented the design of LA tools using co-design and participatory design methods (Dollinger & Lodge, 2018; Fiorini et al., 2018; Holstein, Hong, Tegene, McLaren, & Aleven, 2018) and stressed the role of stakeholder engagement in the design of real-world educational tools (Alhadad, Thompson, Knight, Lewis, & Lodge, 2018; Lang, Macfadyen, Slade, Prinsloo, & Sclater, 2018; Rodríguez-Triana, Prieto, Martínez-Monés, Asensio-Pérez, & Dimitriadis, 2018). Other studies report the use of user-centred design approaches (e.g. the Value Sensitive Design methodology) and techniques (e.g. focus groups, laddering) for designing learning analytics (Chen & Zhu, 2019; de Quincey, Briggs, Kyriacou, & Waller, 2019). The special issue of the *Journal of Learning Analytics* on Human-Centred Learning Analytics (Buckingham Shum, Ferguson, et al., 2019) provides further examples of good research and design methodology, but also confirms that this research area needs to mature.

Turning now to the question of how theory can inform design, Chen (2015) notes that, *"theory has been an important concern since the emergence of the [Learning analytics] field"*. Similarly, Stewart (2017) argued that *"a gap between theory and practice is evident"*. Knight, Shum, and Littleton (2014) argued that no LA system is neutral, such that whether by design or accident, the software and the ways it is used are making epistemological, pedagogical and assessment commitments: *"We have gone beyond "our learning analytics are our pedagogy [...] analytics also embody related epistemological assumptions, thus perpetuating related assessment regimes. [...] learning analytics unavoidably embody these triadic assumptions."*

Ideally then, human-centred approaches should be grounded in epistemological and pedagogical theories and practices for designing learning analytics solutions that are tailored to stakeholders' specific needs. However, this may create a tension, expressed as making a trade-off between how much influence stakeholders' perspectives have on design decisions, versus what 'theory' has to say. Stakeholders are not aware of the potential of learning analytics and data collection methods (Matcha et al., 2019); teachers and students may have some misconceptions about the learning design, aims or outcomes of the activity (DiSalvo, 2016), hindering the development of a robust LA solution. How are LA researchers and designers to reconcile differences between stakeholder perspectives and academically sound theory?

The HCD-MMLA approach developed in this thesis tackles this by combining an *inductive* content analysis process from co-design sessions outputs (elicit and value what is most important to users, and then seek to abstract and formalise as implementable requirements), with a *deductive* theory-driven analysis (seek to align emerging requirements with theory and associated evidence). This dual process goes beyond the challenges in multimodal learning analytics research, by considering critical aspects of the learning environment and stakeholders' specifications (in the inductive step), and pedagogical theory, discipline knowledge and research evidence (in the deductive step). In short, the aim of this dual process within the HCD-MMLA approach is to combine the best of these two realms: human-centred design and theory-driven learning analytics. In this sense, the HCD-MMLA approach can promote the design of learning analytics solutions that enrich stakeholders' (informally expressed, often partial) specifications with research evidence, pedagogical theory and discipline knowledge to address gaps in what stakeholders have expressed, in order to formalise the design requirements. This approach took a theory-driven analysis aimed at: (i) justifying the knowledge, skill or "construct" that can be measured in the learning environment from learning theories or pedagogical practices found in the literature, and (ii) guiding the design decisions that could be implemented in learning analytics solutions. In addition, endorsing stakeholders' specifications with pedagogical theory could lead to the verification or refinement these theories (Chen, 2015).

Since LA is fundamentally a design field, building artefacts for human use, any claim to be grounded in theory must translate in a principled manner into how this shapes analytics and visualisation. It is now a matter of great importance in the LA field to clarify how one establishes a bridge between learning constructs and outcomes (the language of education/learning science) and learners' computationally observable behaviours (the language of data science/analytics). (Wise et al., In Press) advocate for the derivation of construct-aligned metrics to map from meaningful constructs (e.g. based on theory or practice) to human observable behaviours, and from there to digitally captured events, as a way to connect pedagogical theory or practices with analytics. In a more structured view, Shute and Ventura (2013) suggested a theoretical model to support the design of assessment metrics, composed of a *competency model* (e.g. *what collection of knowledge should be assessed?*); an *evidence model* (e.g. *what behaviours or performance should reveal those competencies?*) and a *task model* (e.g. *what tasks should elicit those behaviours that comprise evidence?*). Mislevy, Behrens, Dicerbo, and Levy (2012) describe a similar approach, *Evidence Centered Design*, which again emphasises the importance of designing data for analytics in a pedagogically sound, rigorous manner.

In the context of online learning tools, a substantial body of work is emerging, demonstrating how to map "from clicks to constructs" (Buckingham Shum & Crick, 2016)in environments including MOOCs (Milligan & Griffin, 2016) and high school social media (Martin, Nacu, & Pinkard, 2016).

HCD-MMLA (in combination with the Multimodal Matrix, below) *extends this work into the physical, embodied realm of learning*, providing explicit mapping between the knowledge, skills or behaviours expected from learners (i.e. learning constructs) and machine observable behaviours. Similarly, this template proposed the mapping of constructs to observable behaviours in order to design LA solutions. Thus, this approach can be seen as a formalisation of what this body of research points at, in the context of complex collocated activities: how to assess or model learning constructs from low level data. The HCD-MMLA is a design approach to facilitate this formalisation, which researchers and designers can adopt, with the promise of learning analytics solutions tailored to what stakeholders need. The Multimodal Matrix (next section) is the computational representation of the design decisions.

An important direction for future work includes revalidation of the approach within this setting and extending it to other contexts. For example, it is possible that the template derived from the HCD-MMLA (see Section 4.3.4) may be inadvertently over tuned to the nursing simulation context, and could fail to support other multimodal contexts. In addition, this template can be used in combination with other co-design tools, such as the set of forms developed by Prieto et al. (2018), to structure the design contributions from, and communication between, practitioners, developers and researchers.

8.2.2 Recommendations for researchers and designers

Researchers and designers of learning analytics solutions are advised to use this approach by following these stages:

*Explore the learning situation through **co-design sessions** with teachers and students.* Different HCD techniques, such as ideation, card sorting, interviews, focus groups can help reveal problems, challenges, as well as the context of use for potential solutions promptly. This space provides an opportunity to engage with stakeholders and build a trusted environment. Also, co-design sessions outputs are valuable as they guide the actual design of tools that are tailored to stakeholders' specifications (Prieto-Alvarez, Martinez-Maldonado, & Anderson, 2018; Prieto-Alvarez, Martinez-Maldonado, & Buckingham Shum, 2020).

*Define the learning context and requirements for an MMLA solution using an **inductive content analysis process**.* HCD methods, such as affinity diagrams, and research methods, such as inductive content analysis, allow one to define the learning environment in terms of learning constructs, analytics, sensors and potential visual representations to design and develop learning analytics solutions. As the inquiry process may result in different requirements from stakeholders with different levels of expertise (e.g. tutors, students, subject coordinators), this stage produces a partial mapping of the requirements.

Complete and support the definition of the learning context and requirements for an MMLA solution using a ***deductive theory-driven analysis.*** Reviewing the current literature from different fields offers a better understanding of human perspectives and connects the theory with the practicality of the learning analytics solution. Thus, the formalisation of requirements is completed using a deductive analysis by taking the previous partial mapping and deriving the missing requirements from the literature.

8.3 IMPLICATIONS OF CONTRIBUTION 2: THE MULTIMODAL MATRIX

8.3.1 Discussion and future work

The Multimodal Matrix approach developed in this thesis sought to go beyond the important multimodal data-fusion challenge of integrating information from disparate sensors and devices (Liggins II, Hall, & Llinas, 2017). However, solving the technical data fusion problem is necessary, but not sufficient. Complex as this can be, assuming this can be accomplished, the *interpretive* challenge remains: how can someone make sense of all the multimodal data? (Kipp, Wahlster, Maybury, & Bunt, 2005). The Multimodal Matrix approach facilitated the task of ascribing meaning and understanding of low-level sensory data by grounding the data *theoretically* (e.g. along physical, epistemic, social and affective dimensions; Goodyear & Carvalho, 2014; Roth et al., 2016), and *contextually* (using concepts derived from the learning context, namely the discipline and professional requirements of nursing). Thus, the purpose of this approach was to give meaning to low-level data in an early stage of the analysis.

This work is pertinent to a critical challenge emerging in the LA field, with a growing number of researchers arguing that LA tools should be informed by theory and the use of learning sciences, in order to make a real impact on teaching and learning practice (Gašević, Dawson, & Siemens, 2015; Sedrakyan et al., 2018; Wise & Shaffer, 2015). Wise and Cui (2018), for example, presented seven principles for bridging the gap between learning sciences and learning analytics. It can be seen that the Multimodal Matrix approach and the visualisations generated from this, were aligned with several of these principles as follows:

1. *Grounding analysis in theory:* The Multimodal Matrix provides a way to use insights from theory as constructs (e.g. from the ACAD framework) which define the meaning of data columns, and meaningful combinations of columns. As demonstrated in the case studies, a higher order construct such as "patient-centred care" can be deconstructed into sub-constructs, which eventually map to data streams, which can be modelled in the Multimodal Matrix.
2. *Characterising the context richly:* the structure of the Multimodal Matrix contextualises the multimodal data by defining meaningful representations of the data according to the

learning context (e.g. arousal peaks, in the absence of intense physical activity, may be more useful proxies for stress rather than the raw data from the physiological sensor).

3. *Justify choice of data and/or features:* this is partly done by initially defining the learning constructs that will be part of the analysis. Furthermore, this principle is connected with the purpose of the HCD-MMLA approach (contribution 1), which justifies the learning constructs, data and analysis from stakeholders' specifications and pedagogical theory.
4. *Make sense of high-level patterns using low-level data:* this is the core purpose of the Multimodal Matrix: mapping high-level constructs to low-level data. As explained above (in Section 8.1.2) the qualitative interpretations (e.g. dimensions of collaboration such as social, epistemic, physical and affective) are mapped with quantitative data using QE elements, such as multimodal observations, segments and stanzas. This makes possible the link between high-level constructs with low-level data.
5. *Present analytics results connected to learning processes:* The matrix-like structure derived from the operationalisation of the Multimodal Matrix includes information about the learning context (e.g. critical phases of the learning activity were modelled as stanzas). Thus, the visualisations generated using the matrix-like structure are connected not only to the *task* processes, but also to the *interpretive* processes required to recognise strengths and weaknesses in the team's performance.

This Multimodal Matrix can be framed as a *Quantitative Ethnography* (QE) methodology and support tool to make sense of collocated team activity. *Quantitative Ethnography (QE)* is a methodological approach that respects the insights into specific cultural practices that can be gained from the interpretive disciplines developed in ethnography and other qualitative traditions. QE seeks to apply the power of statistical and other data science techniques to qualitatively coded data, such as observational fieldnotes, interviews, or video analysis (Shaffer, 2017). In the CSCL field, researchers are beginning to demonstrate the value of QE approaches for analysing collaboration processes in online and physical environments, with recent examples including (e.g. in CSCL2019; Eagan, Swiecki, Farrell, & Shaffer, 2019; Gašević et al., 2019; R. Ruis, Pozen, Eagan, & Shaffer, 2019; Sung, Cao, R. Ruis, & Shaffer, 2019; Talafian, Sha, Barany, & Foster, 2019).

Central to QE is the principle of designing ways to model and analyse data that harmonise qualitative and quantitative methodologies, such that all analysis techniques can read, and write to, a *common data representation*. The emphasis on a common data representation distinguishes it from other approaches such as mixed methods which typically have *complementary* representations, and it is this which enables ethnography, and the social sciences more broadly, to contribute in more direct ways to the work of data scientists, and learning analytics specifically.

This thesis adds two principles to QE that reflect the practical design challenge driving the work. A working definition of this QE analysis process and infrastructure is:

> **Applying QE to design timely/real time student-facing analytics requires** *co-design with stakeholders* to model and analyse data that harmonise qualitative and quantitative methodologies, such that all analysis techniques can read, and write to, a common data representation, enabled by an *automated analytics workflow.*

The justification for *co-design* has been provided in detail already in this thesis. What QE highlights is the need to devise a way to model sensor data in ways that both respects, and further enables, culturally meaningful interpretations of human activity (e.g. judgements about the quality of clinical teamwork). Co-design provided a way to understand the experiences and perspectives of students and academics in the specific simulations run in the Health faculty's facilities.

Secondly, *automated analytics workflow* is emphasised, because this is the only way to make sense of data quickly enough to close the feedback loop to educators and students in a timely manner. The Multimodal Matrix is one such common representation, which when connected using scripts for pre-Matrix data fusion and post-Matrix visualisation, provides the automated workflow to generate the Explanatory layered interface for in situ team debriefing.

Finally, QE's emphasis on the importance of the established disciplines for studying cultures and work practices, is a reminder that *the matrix may be populated automatically* by machine, *semi-automatically* by machine plus human coding, or *manually:* in the latter case, new columns or stanzas could be added from analyses performed post hoc by human analysts (but unlike automated analyses, such data would not be available in real time).

Future work should consider the use of the Multimodal Matrix to analyse different aspects of collaboration to those studied here and using different collaboration theories. At this point only a relatively small number of groups have been modelled (see details in Chapter 3) so the consequences of significantly scaling the Matrix are not yet known.

Turning now to the application of machine learning and data mining, such approaches have of course been applied extensively in online contexts (e.g. LMS, MOOCs, adaptive platforms), where all learner interaction is mediated via platforms that can capture clickstreams (Cen, Ruta, Powell, Hirsch, & Ng, 2016; Perera, Kay, Koprinska, Yacef, & Zaïane, 2008). These models have to a lesser degree been operationalised in collocated group work collaboration, using interactive tabletops (Martinez-Maldonado, Kay, et al., 2013; Schneider & Blikstein, 2015). Differential sequence mining and clustering could be used to extract patterns from the Multimodal Matrix which might differentiate team roles (e.g. team leader nurse vs medication nurse). For example, in the case of the team leader role, delegation and monitoring competencies can be modelled by physical presence and speech modalities: if the team leader was delegating and monitoring other students, she should be physically located in a specific spot of the room (e.g. bed footer), observing other students' actions, and directly communicating with the patient and other students. In addition, the analysis of effective steps in relation to the group's performance could be of importance for CSCL and CSCW communities to explore. Thus, applying Human Markov Models (HMM) on multimodal data could produce a representation of learning behaviours for high and low performing groups (McCowan et al., 2005; Zhang, Gatica-Perez, Bengio, McCowan, & Lathoud, 2004). The Multimodal Matrix model can be seen as an initial step towards a deeper analysis of group behaviours and performance. **Future work should investigate the insights that different data mining techniques can provide into collaboration patterns.**

8.3.2 Recommendations for researchers and designers

Other researchers and designers can build on this approach by following the next steps:

- Defining the **higher order constructs** that groups should develop or that can serve to identify 'good' groupwork or teamwork practice for the particular context.
- Selecting a **framework of collaboration** or teamwork to identify the multiple, intertwined *dimensions of collaboration* (i.e. set, social, epistemic, affective) that embrace the complexity of collaborative activity.
- After fusing and synchronising the multimodal data streams, encoding each modality into one or more **meaningful *multimodal observations***. These multimodal observations will define the columns of the matrix.
- The unit of analysis, for example, utterances, a time window or critical incidents, will define the **meaning of each row** (*segments*) of the matrix.
- Defining how segments can be related to each other (grouped into *stanzas*) to facilitate **associating multimodal information** with higher order constructs for a specific part of the activity based on the learning design or by expert knowledge.
- Once the multimodal matrix is filled with encoded, meaningful multimodal data, this information can be **visualised** (e.g. via proxies) or **mined** (e.g. applying rules, sequence pattern mining or other machine learning algorithms).

8.4 IMPLICATIONS OF CONTRIBUTION 3: THE EVISLA APPROACH

8.4.1 Discussion and future work

The final contribution of this thesis is the Explanatory Visual Layered (EvisLA) approach to designing explanatory multimodal visualisations. One of the key concepts of this approach was the adoption of Visual Analytics and Data Storytelling techniques to design explanatory visualisations, whose main goal is the presentation and communication of insights (Iliinsky & Steele, 2011;

Roberts, Ritsos, Jackson, & Headleand, 2017). This approach, driven by the learning design, postulated a middle space of learning visualisations, between undifferentiated 'one-size-fits-all' representations, and personalised learning dashboards. On one hand, undifferentiated visualisations are technically simpler to implement, but can be detrimental if the information represented is not connected with the learning context all (Teasley, 2017). On the other hand, personalised visualisations are more complex to design, but provide flexible adaptation and customisation of the information provided, giving teachers and students the freedom to show or hide certain features that may be or not relevant (Roberts, Howell, et al., 2017). In turn, *explanatory* visualisations are implemented with the purpose of driving attention to key salient features of the learning environment.

In the course of EvisLA's development, other work has emerged that reinforces the rationale for the approach and opens up new ways to extend it. A discussion on the emergence of related ideas in the information visualisation research community is now presented, also considering EvisLA in relation to contemporary debates around *data literacy/design literacy* and *learning analytics/learning design.*

Firstly, the concept of enriching visualisations with tailored messages has been recently explored in the information visualisation community. One example is the work presented by Ren, Brehmer, Lee, Höllerer, and Choe (2017), where a tool called *ChartAccent* was developed to facilitate the annotation of charts, either manually or driven by the data. Another example is *Voder,* a tool developed by Srinivasan, Drucker, Endert, and Stasko (2018) that automatically generates data facts to help users with the interpretation of charts. These two examples provide evidence that similar approaches could be followed to support the generation of explanatory learning visualisations.

Secondly, within LA there have been growing calls for teachers and students to develop stronger *data literacy* skills (Mandinach & Jimerson, 2016; Tsai & Gasevic, 2017; Wolff, Moore, Zdrahal, Hlosta, & Kuzilek, 2016), so that they can benefit from LA tools. This is understandable but could be seen as an analyst-centric response, which risks 'blaming the user' for not understanding visualisations, when at least some of the fault may lie in the poor design skills in the analytics researcher/developer community. An alternative argument from Lonn, Aguilar, and Teasley (2015) is that LA creators need stronger *design literacy* skills. The EvisLA approach, through the operationalisation of educational data storytelling, is a candidate method to support design literacy skills, by enabling researchers and designers to deliver interfaces that more effectively scaffold interpretation. This can be seen as a more natural way to address data interpretation concerns (Greller & Drachsler, 2012; Khan & Pardo, 2016; Teasley, 2017), especially for higher education students and faculty, who have diverse backgrounds and skills.

Finally, there is significant consensus in the LA field on the importance of *aligning learning visualisations with the learning design* (Lockyer et al., 2013; Mangaroska & Giannakos, 2018; Sedrakyan et al., 2018). Providing narratives that are coherently aligned with the learning design to support teachers and students' sensemaking has been recently discussed by researchers (Lim, Dawson, Joksimovic, & Gašević, 2019; Pardo, 2018). Different conceptual models have been proposed to align the learning design with learning analytics tools (Bakharia et al., 2016; Hernández-Leo et al., 2019; Shibani, Knight, & Shum, 2019), but none provides explicit support for *implementation*. The EvisLA approach advances the state of the art by demonstrating a method to contextualise learning visualisations with the learning design, through rules that map the learning (e.g. tasks, objectives, expected outcomes) with visual elements of a visualisation (e.g. title, data points, narratives). In this way, the approach provided to teachers and students with learning visualisations *tailored to pedagogical intentions* (Soller et al., 2005; Wise & Cui, 2018). The rules were designed through consultation with teachers, respecting the principle of giving teachers agency over the design of visualisations (Bakharia et al., 2016; Shibani et al., 2019). EvisLA is also consistent with the recent literature review presented by Jivet et al. (2018), who argued that design decisions shaping a learning dashboard should be aligned with the educational context, otherwise it can result in inadequate solutions for teachers and students.

Future research directions for EvisLA include (i) further evaluation with students, since this thesis has only reported feedback from educators on the layered interface (Chapter 7), and student feedback on an earlier prototype without layers (Chapter 5); **(ii) studying the educational benefits of students critiquing fictional teams using EvisLA 'replays'**; and **(iii) enabling teachers to create their own rules and personalised messages,** potentially using a rule-based platform such as OnTask, which could integrate student activity data from additional learning tools (Pardo, Jovanovic, Dawson, Gašević, & Mirriahi, 2019).[24]

8.4.2 Recommendations for researchers and designers

From the operationalisation of the EvisLA approach developed in this thesis, the following recommendations are proposed for researchers and designers seeking to develop explanatory learning visualisations.

One of the purposes of the EvisLA approach is to make it possible for learning constructs (e.g. tasks, objectives, expected outcomes) to shape the visualisations. Therefore, it is important to **understand the context** in which these visualisations will be rolled out. Several techniques can be used to explore and understand the learning context, from co-design sessions, to field observations.

High-level rules allow the conceptualisation of key aspects of the learning activity. Formalising aspects of the learning design (e.g. a learning outcome: "all members of the group should equally participate") in rules guides the feedback. It is recommended that designers and researchers find a systematic mechanism to document this process. For instance, during the last study reported in

[24]Case studies where OnTask tool has been tested and used can be found in: https://www.ontasklearning.org/case-studies/

Chapter 7, a document template (Appendix B.1) was produced to capture the learning design explicitly, documenting what teachers considered important to visualise in the explanatory visualisations. Once those high-level rules have been defined, the learning analytics team (e.g. researchers, designers, developers) should meet and discuss if all the rules are clear. This should avoid any misinterpretation or misalignment between the learning activity and the design of the visualisation.

Designers of learning analytics solutions should pay attention to **guidelines** from the broader Information Visualisation community to create effective visualisations. For example, Alhadad (2018) presented a set of guidelines for designing data visualisations for attention and cognition, while Sedrakyan, Mannens, and Verbert (2019) provided a compelling list of charts to choose depending on the intended learning goal and feedback dimensions. In Chapter 6 of this thesis, a set of principles were also presented, following Tufte (2001) and Knaflic (2015) guidelines for designing effective visualisations.

Learning data can be complex, especially in the area of multimodal learning analytics where data comes from many sensors, or many features could be extracted from a data source. A layered model, as the one introduced in the EvisLA approach, could help **scaffold** the presentation and interpretation of the data, by attending to the design principles for good layered user interfaces (Sec. 7.3.2).

A methodological recommendation is to **evaluate** the learning visualisations using both quantitative and qualitative data. Self-reported quantitative data from questionnaires may help researchers to evaluate workload (e.g. NASA TLX[25]), technology acceptance (Davis, 1989), usability (Brooke, 1996) of learning visualisations. A more contextualised questionnaire for the evaluation of the quality of learning analytics tools, the EFLA scale (Scheffel, Drachsler, Toisoul, Ternier, & Specht, 2017), could also provide insights about the data, awareness and reflection, and impact of learning visualisations. However, researchers should bear in mind that these tests provide stakeholders' self-reported data, which brings its own subjective biases compared to behavioural data. As illustrated in Chapter 6, using eye-tracking data or other sensors can provide researchers with a better understanding of the design impact, specially pinpointing areas of interest or elements of the learning visualisation. Qualitative data (e.g. think-aloud statements, semi-structured interviews) can be used to explore deeply the weaknesses and strengths of particular features of the learning visualisations.

[25] https://humansystems.arc.nasa.gov/groups/tlx/downloads/TLXScale.pdf

8.5 CHALLENGES, LIMITATIONS AND FUTURE WORK

The research in this thesis has provided well-grounded conceptual contributions that should advance work in the MMLA, LA, CSCL and CSCW communities. Nonetheless, this research has also faced challenges that are inherent to the technical feasibility of capturing multimodal data, and the authenticity of the learning environment in which the work was conducted. The following sections discuss technical, ethical and methodological challenges, limitations and opportunities encountered in conducting this research in-the-wild (in authentic classrooms and with real stakeholders). In doing so, potential opportunities for future work are also highlighted.

8.5.1 Technical aspects

A well-known challenge in MMLA, and in multimodality in general, is the synchronisation of data sources, captured from different sensors, and at different frequencies (Ochoa, Chiluiza, et al., 2018; Reeves et al., 2004). A two-step manual synchronisation was performed with a tool developed for this purpose to store each timestamp in a database (see Chapter 3 for details). **Once the multimodal visualisations reach a maturity point in the design, future work should consider automatic solutions for synchronising multimodal data.** For instance, a promising tool being developed by Schneider, Di Mitri, Limbu, and Drachsler (2018) can be used to simplify tool interoperability. However, this tool should be tested in order to explore the feasibility of using it outside laboratory conditions, in more authentic 'in-the-wild' studies. Another solution to test in the future is the SSI framework (Wagner et al., 2013), an open source project that supports synchronisation and real-time analysis of multi-sensor data. Again, testing of this solution should be done to verify its feasibility, as most of the projects that have used this solution are related to virtual environments (e.g. Anderson et al., 2013; Niewiadomski et al., 2013) or have been tested only in controlled lab studies (e.g. Damian et al., 2015).

The analysis performed over the multimodal data to generate the multimodal visualisations was done semi-automatically. Location, positioning and arousal features were automatically extracted using different tools and scripts (see Section 3.1.1.3 for details). In order to generate the multimodal visualisations, actions were critical to model epistemic episodes of a learning task. However, even though the video recordings of the activities were available, actions were manually tagged by a human observer using a web-based logging application, inspired by tools that followed the same approach (Kharrufa et al., 2017; Qarabash, Heslop, Kharrufa, Balaam, & Devlin, 2019) for collaborative activities. This tool was very helpful to illustrate the conceptual approaches developed in this thesis (as explained in Chapter 7:). Furthermore, other usages can be envisioned for this observer tool to support CSCL research. For instance, another research program is interested to use the tool to generate a post-hoc collaborative visualisation to make error procedures visible (e.g. Palominos, Levett-Jones, Power, & Martinez-Maldonado, 2019). **Future work should focus on the feasibility of inferring actions from the multimodal data, once a substantial human coded**

dataset has been built. For instance, some studies have explored computational methods for recognising students' engagement and attention in lecture classrooms (Raca, Tormey, & Dillenbourg, 2014), or analysing students' strategies using body postures during interaction episodes with a Tangible User Interface (Schneider & Blikstein, 2015), both for post-hoc analysis purposes. A very recent work has started to investigate real-time systems capable of capturing live classroom's interactions (e.g. hands raised, students' speech, instructor's speech) by deploying a suit of computational algorithms (Ahuja et al., 2019). Once MMLA research has dealt with the fully automation of students' actions and interactions, more empirical research on multimodal interfaces could be carried out in live classrooms.

8.5.2 Ethical aspects

Current practices in Learning Analytics and, implicitly in MMLA research, are moving towards **algorithmic transparency** (Pardo & Siemens, 2014). Many have argued that algorithms should be transparent, not an opaque 'black box', considering that algorithms will influence users' decisions. In this sense, data owners should understand how the data is being collected, used, accessed and, how algorithms are applied to abstract the data into the learning visualisations to increase transparency and understanding of the underlying data (2017). This research embraced this principle in two aspects.

The *first* aspect is related to the increase of **transparency** in the design decisions for implementing learning visualisations. Design decisions were corroborated with different stakeholders (e.g. teachers, students, subject experts, subject coordinators) through the whole research project. Chapter 4: reported co-design sessions with stakeholders and initial prototypes considering their perspectives. Also, illustrative studies from Chapter 5:, Chapter 6: and Chapter 7: illustrated the validation of these prototypes and stakeholders' recommendations to improve the design of further prototypes. In addition, the rules that were used to generate the final high-fidelity prototype were validated with experts, in order to make this process transparent and to confirm if the prescriptive information presented in the visualisations were suitable according to what students would expect to receive as feedback (e.g. not criticising, but instead highlighting errors and give recommendations).

The *second* aspect is related to the **use and access** of the information presented in visualisations of groupwork. Considering that the data collected is personal for each student, it is important to identify all the people who will have access to this information. This will ensure that the information is managed and used according to students' consent and request. Chapter 4: and Chapter 5: addressed this aspect by exploring the *accountability* of the visualisations, that is, who should see the information presented in the dashboard and what are the purposes of using the visualisations. Overall, students acknowledged that learning visualisations from group work activities are meant to be inspected by other group members. Since the activity promotes collaboration, this learning also promote a collaborative inspection of the visualisation (Isenberg et al., 2011). Therefore, the information presented can be seen by members of the same group to reflect upon individual and team's performance. However, researchers should be careful about the purpose of the learning visualisation, if they consider sharing this information with external people (e.g. other groups, the whole classroom, other teachers). As pointed out by Teasley (2017), the social comparison effect on performance may lead to feelings of superiority and better goal orientation, but can also induce lower self-confidence or negative feelings (Aguilar, 2016). This is also echoed by Jivet et al. (2018) who found contrasting results (e.g. motivation vs. demotivation) when reviewing the framing references presented in current learning dashboards. More evidence is needed to clarify if and how the learning visualisation can be shared with other groups. In addition, working with affective data in learning environments should be done cautiously, as it may capture emotional states that are triggered by internal and external factors besides the learning environment. **Future work should investigate how physiological information presented to others raises privacy concerns.** Considering that the dashboard invites the exploration of sensitive data in a shared view for team members, a next step would be to explore if students feel comfortable exploring physiological data from their peers.

Capturing students' interactions in authentic classrooms raises a concern in terms of **data privacy** (Picard, 2000) and ownership (Ochoa, 2017; Oviatt, 2018). For example, MMLA data may capture non-related learning actions that conflicts with personal privacy (e.g. when a teacher/student leaves the classroom for an unknown reason). In this research, ethics approval to capture data in the classroom and students' consent were adopted to comply with ethical research. As proposed by Ochoa (2017), ownership of the data should be in hands of the final user (e.g. teachers, students). However, if large scale adoption is expected from multimodal visualisations to support the provision of feedback, how will potentially many thousands of students be given personal control over their data? **Future research should investigate ethical policy frameworks, and the accompanying architectures, curation tools and user interfaces, that could enable considering personal data ownership, storage, and management of multimodal data.**

8.5.3 Methodological aspects

This thesis adopted a Design Based Research (DBR) process in order to exploit the benefits of working in authentic learning environments and engaging with academics and different institutional stakeholders. DBR provided a structure for the research to make more impact by working in authentic environments. However, during the development of this research project several challenges and limitations were faced, that are inherent to DBR.

Engaging with teachers and students is challenging in educational research. As is typical, it was the case that teachers and students lacked understanding about the potential benefits of using learning analytics solutions in the classroom, so building trust with teachers and students took time. This allowed them to become more interested and engaged with the research. A testament to the success of this is that the findings and prototypes from this thesis are being taken forward in ongoing investigations, which we hope will engage more teachers.

DBR depends on the availability of teachers and students. Teachers can be overwhelmed with teaching activities, and students are usually not interested in research activities or have limited availability. The logistics of scheduling time with both threatened at times to impede the progress of the research project, and should be taken into account in future projects.

Similarly, authentic teaching and learning contexts are hard to control. Working with authentic scenarios provides a naturalistic context in which the learning activity unfolds, but also provides challenges and imposes methodological conditions. Contrasting the methodological affordances of DBR with experimental research, external circumstances could hinder the process of research. For example, it is challenging to test specific variables or compare conditions in classrooms, as interventions may need to be negotiated with the teacher responsible of the learning design or subject coordinators. This work tried to engage in authentic scenarios, expecting to generate results that are closer to the reality. Collecting multimodal data is also a challenging and critical activity in conducting research from authentic scenarios. Technological constraints, logistics, ethical implications and teachers' involvement are all variables that can endanger data collection in authentic learning environments. It is of key importance that teachers play an active role in the research and empower their students to participate, as the majority of students do not have a research background.

Finally, in this research, the multimodal data has been used to generate learning visualisations to guide the reflection of groups. During a debriefing process it is expected that students reflect on their behaviour and performance. This opens up an unintended challenge in the use of multimodal data for interpreting behaviour and performance. The problem is to know to what extent people involved in the data capture in authentic environments change their behaviour unintentionally, since they are aware that they are participating in research (Holden, 2001; Parsons, 1974). This challenge needs to be further investigated, since it could lead to erroneous conclusions.

8.6 CONCLUDING REMARKS

Collocated group work is still fundamental in modern work, despite the online revolution. In collocated learning environments, multimodal learning analytics have made it possible to transform teamwork from being an ephemeral activity, to an experience that can be reflected on deeply, as digital traces of collaboration become amenable to computational analysis and visualisation. Providing feedback to collocated groups is generating substantial interest in different research communities, but it is of first order importance that the researchers/designers understand the learning context, stakeholders, and the best way to provide actionable and intelligible feedback.

This thesis has argued, and demonstrated empirically, that human-centred research and design methods can guide the creation of feedback tools that are meaningful to educators and students, even when complex multimodal data is used to capture and model multiple dimensions of group activity. This thesis has demonstrated how stakeholder input, learning design and theory, and knowledge of professional work practices, can powerfully shape the design of data-driven visual dashboards. The contributions of this research open up new possibilities to provide timely feedback, via effective visualisations, tuned to the intended learning outcomes.

Bibliography

Aguilar, S. J. (2016). Perceived motivational affordances: Capturing and measuring students' sense-making around visualizations of their academic achievement information.

Ahuja, K., Kim, D., Xhakaj, F., Varga, V., Xie, A., Zhang, S., . . . Agarwal, Y. (2019). Edusense: Practical classroom sensing at scale. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 3*(3), 1-26.

Al-Qaraghuli, A., Zaman, H. B., Olivier, P., Kharrufa, A., & Ahmad, A. (2011). Analysing tabletop based computer supported collaborative learning data through visualization (M. Universiti Kebangsaan, M. Universiti Putra, M. Universiti Sains, P. University Teknologi, & M. Universiti Teknologi, Trans.). In *2nd International Visual Informatics Conference, IVIC 2011* (Vol. 7066 LNCS, pp. 329-340). Selangor.

Alhadad, S. S. J. (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.

Alhadad, S. S. J., Thompson, K., Knight, S., Lewis, M., & Lodge, J. M. (2018). *Analytics-enabled teaching as design: Reconceptualisation and call for research.* Paper presented at the 8th International Conference on Learning Analytics and Knowledge, LAK'18, Sydney, AUS.

Anderson, K., André, E., Baur, T., Bernardini, S., Chollet, M., Chryssafidou, E., . . . Gebhard, P. (2013). *The tardis framework: Intelligent virtual agents for social coaching in job interviews.* Paper presented at the International Conference on Advances in Computer Entertainment Technology.

Anslow, C. (2014). Reflections on collaborative software visualization in co-located environments. In *2014 ieee international conference on software maintenance and evolution* (pp. 645-650). Los Alamitos: Ieee Computer Soc.

Anzcor. (2016). *Https://resus.Org.Au/guidelines/anzcor-guidelines/*. Retrieved from Australia: https://resus.org.au/guidelines/anzcor-guidelines/

Aslan, S., Alyuz, N., Tanriover, C., Mete, S. E., Okur, E., D'Mello, S. K., & Arslan Esme, A. (2019). *Investigating the impact of a real-time, multimodal student engagement analytics technology in authentic classrooms.* Paper presented at the Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK.

Audhkhasi, K., Kandhway, K., Deshmukh, O. D., & Verma, A. (2009). *Formant-based technique for automatic filled-pause detection in spontaneous spoken english.* Paper presented at the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing.

Azevedo, R., Millar, G. C., Taub, M., Mudrick, N. V., Bradbury, A. E., & Price, M. J. (2017). *Using data visualizations to foster emotion regulation during self-regulated learning with advanced learning technologies: A conceptual framework.* Paper presented at the Proceedings of the Seventh International Learning Analytics & Knowledge Conference, Vancouver, British Columbia, Canada.

Bach, B., Kerracher, N., Hall, K. W., Carpendale, S., Kennedy, J., & Henry Riche, N. (2016). *Telling stories about dynamic networks with graph comics.* Paper presented at the Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.

Bachour, K., Kaplan, F., & 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.

Bakharia, A., Corrin, L., De Barba, P., Kennedy, G., Gašević, D., Mulder, R., . . . Lockyer, L. (2016). *A conceptual framework linking learning design with learning analytics.* Paper presented at the Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, Edinburgh, UK.

Bales, R. F. (1950). A set of categories for the analysis of small group interaction. *American Sociological Review, 15*(2), 257-263.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The journal of the learning sciences, 13*(1), 1-14.

Bardram, J. E. (2005). Activity-based computing: Support for mobility and collaboration in ubiquitous computing. *Personal and Ubiquitous Computing, 9*(5), 312-322.

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences, 12*(3), 307-359.

Beebe, S. A., & Masterson, J. T. (2003). *Communicating in small groups* (Eighth Edition ed.). Boston: Pearson Education, Inc.

Bellanca, J. A. (2011). *21st century skills: Rethinking how students learn*: Solution Tree Press.

Bland, A. J., Topping, A., & Wood, B. (2011). A concept analysis of simulation as a learning strategy in the education of undergraduate nursing students. *Nurse Education Today, 31*(7), 664-670.

Blikstein, P. (2013). *Multimodal learning analytics*. Paper presented at the Proceedings of the Third International Conference on Learning Analytics and Knowledge, Leuven, Belgium.

Blikstein, P., & Worsley, M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. *Journal of Learning Analytics, 3*(2), 220-238.

Bodily, R., Kay, J., Aleven, V., Jivet, I., Davis, D., Xhakaj, F., & Verbert, K. (2018). *Open learner models and learning analytics dashboards: A systematic review*. Paper presented at the Proceedings of the 8th International Conference on Learning Analytics and Knowledge, Sydney, New South Wales, Australia.

Bodily, R., & Verbert, K. (2017). *Trends and issues in student-facing learning analytics reporting systems research*. Paper presented at the International Learning Analytics and Knowledge Conference, LAK'17.

Boud, D., & Molloy, E. (2013a). *Feedback in higher and professional education: Understanding it and doing it well*: Routledge.

Boud, D., & Molloy, E. (2013b). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education, 38*(6), 698-712.

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2013). A guide for analysing electrodermal activity (eda) & skin conductance responses (scrs) for psychological experiments. *Psychophysiology, 49*(1), 1017-1034.

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

Broadhurst, P. L. (1957). Emotionality and the yerkes-dodson law. *Journal of experimental psychology, 54*(5), 345.

Brooke, J. (1996). Sus-a quick and dirty usability scale. *Usability evaluation in industry, 189*(194), 4-7.

Bryan, C., Ma, K.-L., & Woodring, J. (2017). Temporal summary images: An approach to narrative visualization via interactive annotation generation and placement. *IEEE Transactions on Visualization and Computer Graphics, 23*(1), 511-520.

Buckingham Shum, S., & Crick, R. D. (2016). Learning analytics for 21st century competencies. *Journal of Learning Analytics, 3*(2), 6-21.

Buckingham Shum, S., Echeverria, V., & Martinez-Maldonado, R. (2019). *The multimodal matrix as a quantitative ethnography methodology*. Paper presented at the Advances in Quantitative Ethnography, Cham.

Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-centred learning analytics. *Journal of Learning Analytics*, in press.

Bull, S., & Kay, J. (2016). Smili☺: A framework for interfaces to learning data in open learner models, learning analytics and related fields. *International Journal of Artificial Intelligence in Education, 26*(1), 293-331.

California, U. o. (2017). *Learning data privacy principles*. Retrieved from University of California: https://www.ets.berkeley.edu/sites/default/files/general/uc_learning_data_principles_final03.05.2018.pdf

Cen, L., Ruta, D., Powell, L., Hirsch, B., & Ng, J. (2016). Quantitative approach to collaborative learning: Performance prediction, individual assessment, and group composition. *International Journal of Computer-Supported Collaborative Learning, 11*(2), 187-225.

Ceneda, D., Gschwandtner, T., May, T., Miksch, S., Schulz, H., Streit, M., & Tominski, C. (2017). Characterizing guidance in visual analytics. *IEEE Transactions on Visualization and Computer Graphics, 23*(1), 111-120.

Chappel, S. E., Verswijveren, S. J. J. M., Aisbett, B., Considine, J., & Ridgers, N. D. (2017). Nurses' occupational physical activity levels: A systematic review. *International Journal of Nursing Studies, 73*, 52-62.

Charleer, S., Klerkx, J., Duval, E., De Laet, T., & Verbert, K. (2016). Creating effective learning analytics dashboards: Lessons learnt. In K. Verbert, M. Sharples, & T. Klobučar (Eds.), *Adaptive and adaptable learning: 11th european conference on technology enhanced learning, ec-tel 2016, lyon, france, september 13-16, 2016, proceedings* (pp. 42-56). Cham: Springer International Publishing.

Chen, B. (2015). From theory use to theory building in learning analytics: A commentary on" learning analytics to support teachers during synchronous cscl". *Journal of Learning Analytics, 2*(2), 163-168.

Chen, B., & Zhu, H. (2019). *Towards value-sensitive learning analytics design*. Paper presented at the 9th International Conference on Learning Analytics & Knowledge, Tempe, AZ, USA.

Chua, Y. H. V., Dauwels, J., & Tan, S. C. (2019). *Technologies for automated analysis of co-located, real-life, physical learning spaces: Where are we now?* Paper presented at the International Learning Analytics and Knowledge Conference, LAK'19, Tempe, AZ, USA.

Cohen, S. G., & 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.

Corrin, L., & de Barba, P. (2015). *How do students interpret feedback delivered via dashboards?* Paper presented at the Proceedings of the Fifth International Conference on Learning Analytics And Knowledge.
Cotton, W., Lockyer, L., & Brickell, G. (2009). *A journey through a design-based research project.* Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009, Chesapeake, USA.
Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin, 52*(4), 281.
Csanadi, A., Eagan, B., Kollar, I., Shaffer, D. W., & Fischer, F. (2018). When coding-and-counting is not enough: Using epistemic network analysis (ena) to analyze verbal data in cscl research. *International Journal of Computer-Supported Collaborative Learning, 13*(4), 419-438.
Cukurova, M., Luckin, R., Millán, E., & Mavrikis, M. (2018). The nispi framework: Analysing collaborative problem-solving from students' physical interactions. *Computers & Education, 116*, 93-109.
Damian, I., Tan, C. S. S., Baur, T., Schöning, J., Luyten, K., & André, E. (2015). *Augmenting social interactions: Realtime behavioural feedback using social signal processing techniques.* Paper presented at the 33rd annual ACM conference on Human factors in computing systems.
Darrow, C. W. (1964). The rationale for treating the change in galvanic skin response as a change in conductance. *Psychophysiology, 1*(1), 31-38.
Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.
Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In G. Berntson, J. T. Cacioppo, & L. G. Tassinary (Eds.), *Handbook of psychophysiology* (3 ed., pp. 159-181). Cambridge: Cambridge University Press.
de Lima, Y. O., & de Souza, J. M. (2017). *The future of work: Insights for cscw.* Paper presented at the Computer Supported Cooperative Work in Design (CSCWD), 2017 IEEE 21st International Conference on.
De Luca, R. (2014). Review of assessment feedback au - li, jinrui. *Studies in Higher Education, 39*(2), 378-393.
de Quincey, E., Briggs, C., Kyriacou, T., & Waller, R. (2019). *Student centred design of a learning analytics system.* Paper presented at the 9th International Conference on Learning Analytics & Knowledge, Tempe, AZ, US.
Dewey, J. (1997). *How we think.* Chelmsford, USA: Courier Corporation.
Di Mitri, D., Schneider, J., Klemke, R., Specht, M., & Drachsler, H. (2019). *Read between the lines: An annotation tool for multimodal data for learning.* Paper presented at the 9th International Conference on Learning Analytics and Knowledge (LAK19), Tempe, Arizona, USA.
Di Mitri, D., Schneider, J., Specht, M., & Drachsler, H. (2018). From signals to knowledge: A conceptual model for multimodal learning analytics. *Journal of Computer Assisted Learning, 34*(4), 338-349.
Dich, Y., Reilly, J., & Schneider, B. (2018). *Using physiological synchrony as an indicator of collaboration quality, task performance and learning.* Paper presented at the International Conference on Artificial Intelligence in Education, AIED 2018, Cham.
Dillenbourg, P. (1999). What do you mean by collaborative learning. *Collaborative-learning: Cognitive and computational approaches, 1*, 1-15.
DiMicco, J. M., & Bender, W. (2007). Group reactions to visual feedback tools. In Y. DeKort, W. Ijsselsteijn, C. Midden, B. Eggen, & B. J. Fogg (Eds.), *Persuasive technology* (Vol. 4744, pp. 132-+). Berlin: Springer-Verlag Berlin.
DiMicco, J. M., Hollenbach, K. J., Pandolfo, A., & Bender, W. (2007). The impact of increased awareness while face-to-face. *Human–Computer Interaction, 22*(1-2), 47-96.
DiSalvo, B. (2016). *Participatory design through a learning science lens.* Paper presented at the Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, California, USA.
Dollinger, M., & Lodge, J. M. (2018). *Co-creation strategies for learning analytics.* Paper presented at the 8th International Conference on Learning Analytics and Knowledge, LAK'18, Sydney, AUS.
Doyle, J. C., & Csete, M. (2011). Architecture, constraints, and behavior. *Proceedings of the National Academy of Sciences, 108*(Supplement 3), 15624-15630.
Drachsler, H., & Greller, W. (2011). Turning learning into numbers–a learning analytics framework.
Duffy, M. C., Lajoie, S., & Lachapelle, K. (2016). Measuring emotions in medical education: Methodological and technological advances within authentic medical learning environments. In S. Bridges, L. K. Chan, & C. E. Hmelo-Silver (Eds.), *Educational technologies in medical and health sciences education* (pp. 181-213). Cham: Springer.
Dumais, S., Jeffries, R., Russell, D. M., Tang, D., & Teevan, J. (2014). Understanding user behavior through log data and analysis. In S. J. Olson & A. W. Kellogg (Eds.), *Ways of knowing in hci* (pp. 349-372). New York: Springer.
Duval, E. (2011). *Attention please!: Learning analytics for visualization and recommendation.* Paper presented at the Proceedings of the 1st International Conference on Learning Analytics and Knowledge, Banff, Alberta, Canada.

Dykes, B. (2015). Data storytelling: What it is and how it can be used to effectively communicate analysis results. *Applied Marketing Analytics, 1*(4), 299-313.

Eagan, B. R., Swiecki, Z., Farrell, C., & Shaffer, D. W. (2019). *The binary replicate test: Determining the sensitivity of cscl models to coding error.* Paper presented at the 13th International Conference on Computer Supported Collaborative Learning, CSCL 2019, Lyon, France.

Echeverria, V., Avendaño, A., Chiluiza, K., Vásquez, A., & Ochoa, X. (2014). *Presentation skills estimation based on video and kinect data analysis.* Paper presented at the MLA 2014 - Proceedings of the 2014 ACM Multimodal Learning Analytics Workshop and Grand Challenge, Co-located with ICMI 2014, Istanbul, Turkey.

Echeverria, V., Falcones, G., Castells, J., Martinez-Maldonado, R., & Chiluiza, K. (2017). *Exploring on-time automated assessment in a co-located collaborative system.* Paper presented at the 2017 Fourth International Conference on eDemocracy & eGovernment (ICEDEG).

Echeverria, V., Martinez-Maldonado, R., Buckingham Shum, Simon, Chiluiza, K., Granda, R., & Conati, C. (2018). Exploratory versus explanatory visual learning analytics: Driving teachers' attention through educational data storytelling. *Journal of Learning Analytics, 5*(3), 72-97.

Echeverria, V., Martinez-Maldonado, R., Chiluiza, K., & Buckingham Shum, S. (2017). *Dbcollab: Automated feedback for face-to-face group database design.* Paper presented at the 25th International Conference on Computers in Education (ICCE 2017), Christchurch, New Zealand.

Echeverria, V., Martinez-Maldonado, R., Granda, R., Chiluiza, K., Conati, C., & Buckingham Shum, S. (2018). *Driving data storytelling from learning design.* Paper presented at the Eighth International Learning Analytics & Knowledge Conference (LAK'18), Sydney, AUS.

Echeverria, V., Martinez-Maldonado, R., Power, T., Hayes, C., & Shum, S. B. (2018). *Where is the nurse? Towards automatically visualising meaningful team movement in healthcare education.* Paper presented at the International Conference on Artificial Intelligence in Education.

Echeverria, V., Martinez-Maldonado, R., & Shum, S. B. (2017). *Towards data storytelling to support teaching and learning.* Paper presented at the 29th Australian Conference on Computer-Human Interaction, OzCHI'17, Brisbane, Australia.

Echeverria, V., Martinez-Maldonado, R., & Shum, S. B. (2019). *Towards collaboration translucence: Giving meaning to multimodal group data.* Paper presented at the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), Glasgow, UK.

Edelson, D. C. (2002). Design research: What we learn when we engage in design. *The journal of the learning sciences, 11*(1), 105-121.

Egan, G. (2013). *The skilled helper: A problem-management and opportunity-development approach to helping*: Cengage Learning.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing, 62*(1), 107-115.

Eradze, M., Rodríguez-Triana, M. J., & Laanpere, M. (2017). *How to aggregate lesson observation data into learning analytics datasets?* Paper presented at the 6th Multimodal Learning Analytics (MMLA) Workshop and the 2nd Cross-LAK Workshop co-located with 7th International Learning Analytics and Knowledge Conference (LAK 2017).

Erickson, T., Halverson, C., Kellogg, W. A., Laff, M., & Wolf, T. (2002). Social translucence: Designing social infrastructures that make collective activity visible. *Communications of the ACM, 45*(4), 40-44.

Erickson, T., & Kellogg, W. A. (2000). Social translucence: An approach to designing systems that support social processes. *ACM Transactions on Computer-Human Interaction, 7*(1), 59-83.

Erickson, T., Smith, D., Kellogg, W., Laff, M., Richards, J., & Bradner, E. (1999). *Socially translucent systems: Social proxies, persistent conversation, and the design of "babble".* Paper presented at the SIGCHI Conference on Human Factors in Computer Systems, CHI'99, Pittsburgh, Pennsylvania, USA.

Evans, A. C., Wobbrock, J. O., & Davis, K. (2016). *Modeling collaboration patterns on an interactive tabletop in a classroom setting.* Paper presented at the International Conference on Computer-Supported Cooperative Work, CSCW'16, San Francisco, California, USA.

Falcones, G., Wong-Villacres, M., Barzola, V. E., & Garcia, K. C. (2016). *Enhancing quality of argumentation in a co-located collaborative environment through a tabletop system.* Paper presented at the 2016 IEEE Ecuador Technical Chapters Meeting (ETCM).

Farid, Z., Nordin, R., & Ismail, M. (2013). Recent advances in wireless indoor localization techniques and system. *Journal of Computer Networks and Communications, 2013.*

Feese, S., Burscher, M. J., Jonas, K., & Tröster, G. (2014). Sensing spatial and temporal coordination in teams using the smartphone. *Human-centric Computing and Information Sciences, 4*(1), 15.

Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *Int. J. Technol. Enhanc. Learn., 4*(5/6), 304-317.

Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., van Staalduinen, J.-P., & Gašević, D. (2019). *Counting clicks is not enough: Validating a theorized model of engagement in learning analytics.*

Fiorini, S., Sewell, A., Bumbalough, M., Chauhan, P., Shepard, L., Rehrey, G., & Groth, D. (2018). *An application of participatory action research in advising-focused learning analytics.*
Forsyth, D. R. (2009). *Group dynamics* (Third Edition ed.): Cengage Learning.
Gabelica, C., Bossche, P. V. D., Segers, M., & Gijselaers, W. (2012). Feedback, a powerful lever in teams: A review. *Educational Research Review, 7*(2), 123-144.
Gaillard, J., Vienne, A., Baume, R., Pedrinis, F., Peytavie, A., & Gesquière, G. (2015). *Urban data visualisation in a web browser.*
Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64-71.
Gašević, D., Joksimović, S., Eagan, B. R., & Shaffer, D. W. (2019). Sens: Network analytics to combine social and cognitive perspectives of collaborative learning. *Computers in Human Behavior, 92*, 562-577.
Gatica-Perez, D. (2009). Automatic nonverbal analysis of social interaction in small groups: A review. *Image and Vision Computing, 27*(12), 1775-1787.
Gazzaniga, M. S. (2015). *Tales from both sides of the brain: A life in neuroscience*: Ecco/HarperCollins Publishers.
Gergle, D., Kraut, R. E., & Fussell, S. R. (2013). Using visual information for grounding and awareness in collaborative tasks. *Human-Computer Interaction, 28*(1), 1-39.
Giacomin, J. (2014). What is human centred design? *The Design Journal, 17*(4), 606-623.
Gilbert, S. B., Slavina, A., Dorneich, M. C., Sinatra, A. M., Bonner, D., Johnston, J., . . . Winer, E. (2018). Creating a team tutor using gift. *International Journal of Artificial Intelligence in Education, 28*(2), 286-313.
Gillies, R. M., Ashman, A. F., & Terwel, J. (2007). *The teacher's role in implementing cooperative learning in the classroom.* New York: Springer.
Gloor, P. A., Grippa, F., Putzke, J., Lassenius, C., Fuehres, H., Fischbach, K., & Schoder, D. (2012). Measuring social capital in creative teams through sociometric sensors. *International Journal of Organisational Design and Engineering, 2*(4), 380-401.
Gokhale, A. A. (1995). Collaborative learning enhances critical thinking.
Goodyear, P., & Carvalho, L. (2014). Framing the analysis of learning network architectures. In L. Carvalho & P. Goodyear (Eds.), *The architecture of productive learning networks* (pp. 48-70). New York: Routledge.
Green, A., Stawicki, S. P., & Firstenberg, M. S. (2018). Introductory chapter: Medical error and associated harm-the the critical role of team communication and coordination. In S. P. Stawicki & M. S. Firstenberg (Eds.), *Vignettes in patient safety-volume 3* (pp. 1-12). London, UK: IntechOpen.
Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society, 15*(3), 42–57.
Grover, S., Bienkowski, M., Tamrakar, A., Siddiquie, B., Salter, D., & Divakaran, A. (2016). *Multimodal analytics to study collaborative problem solving in pair programming*. Paper presented at the Sixth International Conference on Learning Analytics & Knowledge, Edinburgh, United Kingdom.
Haataja, E., Malmberg, J., & Järvelä, S. (2018). Monitoring in collaborative learning: Co-occurrence of observed behavior and physiological synchrony explored. *Computers in Human Behavior, 87*, 337-347.
Hanington, B., & Martin, B. (2012). *Universal methods of design: 100 ways to research complex problems, develop innovative ideas, and design effective solutions*: Rockport Publishers.
Haslett, J., Wills, G., & Unwin, A. (1990). Spider—an interactive statistical tool for the analysis of spatially distributed data. *International Journal of Geographical Information Systems, 4*(3), 285-296.
Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.
Helms, J. W., Arthur, J. D., Hix, D., & Hartson, H. R. (2006). A field study of the wheel—a usability engineering process model. *Journal of Systems and Software, 79*(6), 841-858.
Hernández-Leo, D., Martinez-Maldonado, R., Pardo, A., Muñoz-Cristóbal, J. A., & Rodríguez-Triana, M. J. (2019). Analytics for learning design: A layered framework and tools. *British Journal of Educational Technology, 50*(1), 139-152.
Herrington, J., Reeves, T. C., & Oliver, R. (2009). *A practical guide to authentic e-learning*: Routledge.
Holden, J. D. (2001). Hawthorne effects and research into professional practice. *Journal of Evaluation in Clinical Practice, 7*(1), 65-70.
Holstein, K., Hong, G., Tegene, M., McLaren, B. M., & Aleven, V. (2018). *The classroom as a dashboard: Co-designing wearable cognitive augmentation for k-12 teachers.*
Holstein, K., McLaren, B. M., & Aleven, V. (2017). *Intelligent tutors as teachers' aides: Exploring teacher needs for real-time analytics in blended classrooms.* Paper presented at the Proceedings of the Seventh International Learning Analytics & Knowledge Conference, Vancouver, British Columbia, Canada.
Hutchins, E. (2010). Cognitive ecology. *Topics in cognitive science, 2*(4), 705-715.
Iliinsky, N., & Steele, J. (2011). *Designing data visualizations: Representing informational relationships*: O'Reilly Media, Inc.

Ireland, M. (2008). Assisting students to use evidence as a part of reflection on practice. *Nursing Education Perspectives, 29*(2), 90-93.

Isella, L., Romano, M., Barrat, A., Cattuto, C., Colizza, V., Van den Broeck, W., . . . Tozzi, A. E. (2011). Close encounters in a pediatric ward: Measuring face-to-face proximity and mixing patterns with wearable sensors. *PLOS ONE, 6*(2), e17144.

Isenberg, P., Elmqvist, N., Scholtz, J., Cernea, D., Ma, K.-L., & Hagen, H. (2011). Collaborative visualization: Definition, challenges, and research agenda. *Information Visualization, 10*(4), 310-326.

Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education, 49*(4), 1037-1065.

Janssen, J., Erkens, G., & Kirschner, P. A. (2011). Group awareness tools: It's what you do with it that matters. *Computers in Human Behavior, 27*(3), 1046-1058.

Järvelä, S., Kirschner, P. A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, J., . . . Järvenoja, H. (2015). Enhancing socially shared regulation in collaborative learning groups: Designing for cscl regulation tools. *Educational Technology Research and Development, 63*(1), 125-142.

Jeong, H., Hmelo-Silver, C. E., & Yu, Y. (2014). An examination of cscl methodological practices and the influence of theoretical frameworks 2005–2009. *International Journal of Computer-Supported Collaborative Learning, 9*(3), 305-334.

Jetter, H.-C., Geyer, F., Reiterer, H., Dachselt, R., Fischer, G., Groh, R., . . . Herrmann, T. (2012). *Designing collaborative interactive spaces.* Paper presented at the Proceedings of the International Working Conference on Advanced Visual Interfaces, Capri Island, Italy.

Jetter, H.-C., Reiterer, H., & Geyer, F. (2014). Blended interaction: Understanding natural human–computer interaction in post-wimp interactive spaces. *Personal and Ubiquitous Computing, 18*(5), 1139-1158.

Jewitt, C. (2006). *Technology, literacy and learning: A multimodal approach*: Routledge.

Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of research in education, 32*(1), 241-267.

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). *Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice.* Paper presented at the European Conference on Technology Enhanced Learning.

Jivet, I., Scheffel, M., Specht, M., & Drachsler, H. (2018). *License to evaluate: Preparing learning analytics dashboards for educational practice.* Paper presented at the International Learning Analytics and Knowledge Conference, LAK'18, Sydney, NSW, Australia.

Johnson, D. M., Sutton, P., & Poon, J. (2000). *Face-to-face vs. Cmc: Student communication in a technologically rich learning environment.* Paper presented at the Annual Conference of the Australian Society for Computers in Learning in Tertiary Education, ASCILITE 2000, Coffs Harbour. http://eprints.qut.edu.au/6682/

Johnson, D. W., & Johnson, F. P. (1991). *Joining together: Group theory and group skills* (Tenth Edition ed.): Prentice-Hall, Inc.

Johnson, R. T., & Johnson, D. W. (1986). Cooperative learning in the science classroom. *Science and children, 24*, 31-32.

Jovanović, J., Gašević, D., Pardo, A., Dawson, S., & Whitelock-Wainwright, A. (2019). *Introducing meaning to clicks: Towards traced-measures of self-efficacy and cognitive load.*

Kannampallil, T., Li, Z., Zhang, M., Cohen, T., Robinson, D. J., Franklin, A., . . . Patel, V. L. (2011). Making sense: Sensor-based investigation of clinician activities in complex critical care environments. *Journal of Biomedical Informatics, 44*(3), 441-454.

Ketrow, S. M. (1999). Nonverbal aspects of group communication. *The handbook of group communication theory and research*, 251-287.

Khan, I., & Pardo, A. (2016). *Data2u: Scalable real time student feedback in active learning environments.* Paper presented at the Proceedings of the 6th international conference on learning analytics & knowledge.

Kharrufa, A., Rix, S., Osadchiy, T., Preston, A., & Olivier, P. (2017). *Group spinner: Recognizing and visualizing learning in the classroom for reflection, communication, and planning*. Paper presented at the Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, Colorado, USA.

Kim, J. W., Sottilare, R. A., Brawner, K., & Flowers, T. (2018). *Integrating sensors and exploiting sensor data with gift for im-proved learning analytics.* Paper presented at the Sixth Annual GIFT Users Symposyum, Florida.

Kim, T., Chang, A., Holland, L., & Pentland, A. S. (2008). *Meeting mediator: Enhancing group collaborationusing sociometric feedback.* Paper presented at the International Conference on Computer Supported Cooperative Work, CSCW'08, San Diego, CA, USA.

Kipp, M., Wahlster, W., Maybury, M., & Bunt, H. (2005). Fusion and coordination for multimodal interactive information presentation. In *Multimodal intelligent information presentation* (pp. 325-339): Springer.

Kitto, K., Shum, S. B., & Gibson, A. (2018). *Embracing imperfection in learning analytics*. Paper presented at the International Conference on Learning Analytics and Knowledge, LAK'18.
Knaflic, C. N. (2015). *Storytelling with data: A data visualization guide for business professionals*: John Wiley & Sons.
Knight, S., Shum, S. B., & Littleton, K. (2014). Epistemology, assessment, pedagogy: Where learning meets analytics in the middle space. *Journal of Learning Analytics, 1*(2), 23-47.
Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of cluster in k-means clustering. *International Journal, 1*(6), 90-95.
Kolbe, M., Burtscher, M. J., & Manser, T. (2013). Co-act—a framework for observing coordination behaviour in acute care teams. *BMJ Qual Saf, 22*(7), 596-605.
Kosara, R., & Mackinlay, J. (2013). Storytelling: The next step for visualization. *Computer, 46*(5), 44-50.
Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest, 7*(3), 77-124.
Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*: Edward Arnold.
Kurtz, S., Draper, J., & Silverman, J. (2016). *Skills for communicating with patients*: CRC Press.
Kusunoki, D. S., & Sarcevic, A. (2015). *Designing for temporal awareness: The role of temporality in time-critical medical teamwork*. Paper presented at the International Conference on Computer Supported Cooperative Work, CSCW'15, Vancouver, BC, Canada.
Lallé, S., Conati, C., & Carenini, G. (2016). Prediction of individual learning curves across information visualizations. *User Modeling and User-Adapted Interaction, 26*(4), 307-345.
Lang, C., Macfadyen, L. P., Slade, S., Prinsloo, P., & Sclater, N. (2018). *The complexities of developing a personal code of ethics for learning analytics practitioners: Implications for institutions and the field*. Paper presented at the 8th international conference on learning analytics and knowledge, LAK'18, Sydney, AUS.
Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction*: Morgan Kaufmann.
Lee, B., Riche, N. H., Isenberg, P., & Carpendale, S. (2015). More than telling a story: Transforming data into visually shared stories. *IEEE Computer Graphics and Applications, 35*(5), 84-90.
Lerner, S., Magrane, D., & Friedman, E. (2009). Teaching teamwork in medical education. *The Mount Sinai Journal Of Medicine, New York, 76*(4), 318-329.
Leshed, G., Hancock, J. T., Cosley, D., McLeod, P. L., & Gay, G. (2007). *Feedback for guiding reflection on teamwork practices*. Paper presented at the 2007 international ACM conference on Supporting group work.
Lewis, R. (1997). An activity theory framework to explore distributed communities. *Journal of Computer Assisted Learning, 13*(4), 210-218.
Liggins II, M., Hall, D., & Llinas, J. (2017). *Handbook of multisensor data fusion: Theory and practice*. Boca Raton, USA: CRC press.
Lim, L., Dawson, S., Joksimovic, S., & Gašević, D. (2019). *Exploring students' sensemaking of learning analytics dashboards: Does frame of reference make a difference?* Paper presented at the Proceedings of the 9th International Conference on Learning Analytics & Knowledge, Tempe, AZ.
Liu, A. L., & Nesbit, J. C. (2020). Dashboards for computer-supported collaborative learning. In M. Virvou, E. Alepis, G. A. Tsihrintzis, & L. C. Jain (Eds.), *Machine learning paradigms: Advances in learning analytics* (pp. 157-182). Cham: Springer International Publishing.
Liu, Z., Wang, Y., Dontcheva, M., Hoffman, M., Walker, S., & Wilson, A. (2017). Patterns and sequences: Interactive exploration of clickstreams to understand common visitor paths. *IEEE Transactions on Visualization and Computer Graphics, 23*(1), 321-330.
Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist, 57*(10), 1439-1459.
London, M., & Sessa, V. I. (2006). Group feedback for continuous learning. *Human Resource Development Review, 5*(3), 303-329.
Lonn, S., Aguilar, S. J., & Teasley, S. D. (2015). Investigating student motivation in the context of a learning analytics intervention during a summer bridge program. *Computers in Human Behavior, 47*, 90-97.
Lotherington, H., & Jenson, J. (2011). Teaching multimodal and digital literacy in l2 settings: New literacies, new basics, new pedagogies. *Annual Review of Applied Linguistics, 31*, 226-246.
Luzardo, G., Guamán, B., Chiluiza, K., Castells, J., & Ochoa, X. (2014). *Estimation of presentations skills based on slides and audio features*. Paper presented at the 2014 ACM International Conference on Multimodal Interaction, Turkey, Istanbul.
Ma, K. L., Liao, I., Frazier, J., Hauser, H., & Kostis, H. N. (2012). Scientific storytelling using visualization. *IEEE Computer Graphics and Applications, 32*(1), 12-19.

Maguire, M. (2001). Methods to support human-centred design. *International Journal of Human-Computer Studies, 55*(4), 587-634.
Malec, J. F., Torsher, L. C., Dunn, W. F., Wiegmann, D. A., Arnold, J. J., Brown, D. A., & Phatak, V. (2007). The mayo high performance teamwork scale: Reliability and validity for evaluating key crew resource management skills. *Simulation in Healthcare, 2*(1), 4-10.
Malmberg, J., Järvelä, S., Holappa, J., Haataja, E., Huang, X., & Siipo, A. (2018). Going beyond what is visible: What multichannel data can reveal about interaction in the context of collaborative learning? *Computers in Human Behavior.*
Mandinach, E. B., & Jimerson, J. B. (2016). Teachers learning how to use data: A synthesis of the issues and what is known. *Teaching and Teacher Education, 60*, 452-457.
Mangaroska, K., & Giannakos, M. N. (2018). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies.*
Manser, T. (2009). Teamwork and patient safety in dynamic domains of healthcare: A review of the literature. *Acta Anaesthesiologica Scandinavica, 53*(2), 143-151.
Martin, C. K., Nacu, D., & Pinkard, N. (2016). An approach to using log data to understand and support 21st century learning activity in k-12 blended learning environments. *Journal of Learning Analytics, 3*(2), 37-87.
Martinez-Maldonado, R. (2014). *Analysing, visualising and supporting collaborative learning using interactive tabletops.* (Doctor of Philosophy). University of Sydney,
Martinez-Maldonado, R., Clayphan, A., Yacef, K., & Kay, J. (2014). Towards providing notifications to enhance teacher's awareness in the classroom. In S. TrausanMatu, K. E. Boyer, M. Crosby, & K. Panourgia (Eds.), *Intelligent tutoring systems, its 2014* (Vol. 8474, pp. 510-515). Berlin: Springer-Verlag Berlin.
Martinez-Maldonado, R., Clayphan, A., Yacef, K., & Kay, J. (2015). Mtfeedback: Providing notifications to enhance teacher awareness of small group work in the classroom. *IEEE Transactions on Learning Technologies, 8*(2), 187-200.
Martinez-Maldonado, R., Collins, A., Kay, J., & Yacef, K. (2011). *Who did what? Who said that?: Collaid: An environment for capturing traces of collaborative learning at the tabletop.* Paper presented at the Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, Kobe, Japan.
Martinez-Maldonado, R., Dimitriadis, Y., Kay, J., Yacef, K., & Edbauer, M.-T. (2013). *Mtclassroom and mtdashboard: Supporting analysis of teacher attention in an orchestrated multi-tabletop classroom.* Paper presented at the 10th International Conference on Computer Supported Collaborative Learning.
Martinez-Maldonado, R., Echeverria, V., Elliott, D., Axisa, C., Power, T., & Buckingham Shum, S. (2019). Designing translucent learning analytics with teachers: An elicitation process. *Interactive Learning Environments*((in press)), 1-30.
Martinez-Maldonado, R., Echeverria, V., Fernandez, G., & Shum, S. B. (2020). *From data to insights: A layered storytelling approach for multimodal learning analytics.* Paper presented at the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), Honolulu, HI, USA.
Martinez-Maldonado, R., Echeverria, V., Santos, O. C., Dos Santos, A. D. P., & Yacef, K. (2018). *Physical learning analytics: A multimodal perspective.* Paper presented at the 8th International Conference on Learning Analytics and Knowledge, LAK 2018.
Martinez-Maldonado, R., Echeverria, V., Yacef, K., Dos Santos, A. D. P., & Pechenizkiy, M. (2017). *How to capitalise on mobility, proximity and motion analytics to support formal and informal education?* Paper presented at the CEUR Workshop Proceedings.
Martinez-Maldonado, R., & Goodyear, P. (2016). *Cocodes: Multi-device support for collocated collaborative learning design.* Paper presented at the Proceedings of the 28th Australian Conference on Computer-Human Interaction, Launceston, Tasmania, Australia.
Martinez-Maldonado, R., Goodyear, P., Kay, J., Thompson, K., & Carvalho, L. (2016). *An actionable approach to understand group experience in complex, multi-surface spaces.* Paper presented at the SIGCHI Conference on Human Factors in Computing Systems, CHI'16, San Jose, California, USA.
Martinez-Maldonado, R., Kay, J., Buckingham Shum, S., & Yacef, K. (2019). Collocated collaboration analytics: Principles and dilemmas for mining multimodal interaction data. *Human–Computer Interaction, 34*(1), 1-50.
Martinez-Maldonado, R., Kay, J., & Yacef, K. (2013). An automatic approach for mining patterns of collaboration around an interactive tabletop. In H. C. Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Artificial intelligence in education: 16th international conference, aied 2013, memphis, tn, USA, july 9-13, 2013. Proceedings* (pp. 101-110). Berlin, Heidelberg: Springer Berlin Heidelberg.
Martinez-Maldonado, R., Pardo, A., Mirriahi, N., Yacef, K., Kay, J., & Clayphan, A. (2015). *The latux workflow: Designing and deploying awareness tools in technology-enabled learning settings.* Paper presented at the Fifth International Conference on Learning Analytics and Knowledge.

Martinez-Maldonado, R., Power, T., Hayes, C., Abdiprano, A., Vo, T., Axisa, C., & Shum, S. B. (2017). *Analytics meet patient manikins: Challenges in an authentic small-group healthcare simulation classroom.* Paper presented at the 7th International Conference on Learning Analytics and Knowledge, LAK 2017.

Martinez-Maldonado, R., Yacef, K., & Kay, J. (2015). Tscl: A conceptual model to inform understanding of collaborative learning processes at interactive tabletops. *International Journal of Human-Computer Studies, 83*, 62-82.

Marzano, R. J., Norford, J. S., Paynter, D. E., Pickering, D. J., & Gaddy, B. B. (2001). *A handbook for classroom instruction that works*: ERIC.

Matcha, W., Gasevic, D., & Pardo, A. (2019). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, in press.

Mazzocchi, F. (2015). Could big data be the end of theory in science?: A few remarks on the epistemology of data-driven science. *EMBO reports, 16*(10), 1250-1255.

McCormack, B., & McCance, T. (2016). *Person-centred practice in nursing and health care: Theory and practice*: John Wiley & Sons.

McCowan, L., Gatica-Perez, D., Bengio, S., Lathoud, G., Barnard, M., & Zhang, D. (2005). Automatic analysis of multimodal group actions in meetings. *IEEE transactions on pattern analysis and machine intelligence, 27*(3), 305-317.

McDonald, N., Schoenebeck, S., & Forte, A. (2019). Reliability and inter-rater reliability in qualitative research: Norms and guidelines for cscw and hci practice. *Proceedings of the ACM: Human-Computer Interaction (PACM: HCI). CSCW*, (in press).

Mead, N., & Bower, P. (2000). Patient-centredness: A conceptual framework and review of the empirical literature. *Social Science & Medicine, 51*(7), 1087-1110.

Mehrabian, A. (2017). *Nonverbal communication*: Routledge.

Metoyer, R., Zhi, Q., Janczuk, B., & Scheirer, W. (2018). *Coupling story to visualization: Using textual analysis as a bridge between data and interpretation.* Paper presented at the 23rd International Conference on Intelligent User Interfaces.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review, 63*(2), 81.

Miller, K., Riley, W., & Davis, S. (2009). Identifying key nursing and team behaviours to achieve high reliability. *Journal of Nursing Management, 17*(2), 247-255.

Miller, M. (2005). Teaching and learning in affective domain. *Emerging perspectives on learning, teaching, and technology. Retrieved March, 6*, 2008.

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

Mislevy, R. J., Behrens, J. T., Dicerbo, K. E., & Levy, R. (2012). Design and discovery in educational assessment: Evidence-centered design, psychometrics, and educational data mining. *Journal of Educational Data Mining, 4*(1), 11-48.

Mohan, A., Sun, H. C., Lederman, O., Full, K., & Pentland, A. S. (2018). *Measurement and feedback of group activity using wearables for face-to-face collaborative learning*. Paper presented at the IEEE International Conference on Advanced Learning Technologies, ICALT'18.

Molenaar, I., & Knoop-van Campen, C. (2017). *Teacher dashboards in practice: Usage and impact.* Paper presented at the European Conference on Technology Enhanced Learning.

Müller, L., Rivera-Pelayo, V., Kunzmann, C., & Schmidt, A. (2011). *From stress awareness to coping strategies of medical staff: Supporting reflection on physiological data.* Paper presented at the International Workshop on Human Behavior Understanding, Berlin, Heidelberg.

Müller, P., Huang, M. X., & Bulling, A. (2018). *Detecting low rapport during natural interactions in small groups from non-verbal behaviour.* Paper presented at the International Conference on Intelligent User Interfaces, IUI'18, Tokyo, Japan.

Munzner, T. (2014). *Visualization analysis and design*: A K Peters/CRC Press.

Nakano, Y. I., Nihonyanagi, S., Takase, Y., Hayashi, Y., & Okada, S. (2015). *Predicting participation styles using co-occurrence patterns of nonverbal behaviors in collaborative learning*. Paper presented at the International Conference on Multimodal Interaction, ICMI'15, Seattle, WA, USA.

Nardi, B. A., & Whittaker, S. (2002). The place of face-to-face communication in distributed work. In P. J. Hinds & S. Kiesler (Eds.), *Distributed work: New research on working across distance using technology* (pp. 83-110). Cambridge: MIT Press.

Neubauer, C., Woolley, J., Khooshabeh, P., & Scherer, S. (2016). *Getting to know you: A multimodal investigation of team behavior and resilience to stress.* Paper presented at the International Conference on Multimodal Interaction, ICMI'16, Tokyo, Japan.

Nicol, D. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. *Assessment & Evaluation in Higher Education, 35*(5), 501-517.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Niemantsverdriet, K., Broekhuijsen, M., van Essen, H., & Eggen, B. (2016). *Designing for multi-user interaction in the home environment: Implementing social translucence.* Paper presented at the Designing Interactive Systems, DIS'16, Brisbane, Australia.

Niewiadomski, R., Hofmann, J., Urbain, J., Platt, T., Wagner, J., Piot, B., . . . Dupont, S. (2013). *Laugh-aware virtual agent and its impact on user amusement.* Paper presented at the International conference on Autonomous agents and multi-agent systems.

Nihei, F., Nakano, Y., & Takase, Y. (2017). *Predicting meeting extracts in group discussions using multimodal convolutional neural networks.* Paper presented at the International Conference on Multimodal Interaction, ICMI'17, Glasgow, UK.

Norman, D. (2013). *The design of everyday things: Revised and expanded edition*: Basic books.

Noroozi, O., Alikhani, I., Järvelä, S., Kirschner, P. A., Juuso, I., & Seppänen, T. (2018). Multimodal data to design visual learning analytics for understanding regulation of learning. *Computers in Human Behavior.*

Ochoa, X. (2017). Multimodal learning analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gaševic (Eds.), *The handbook of learning analytics* (1 ed., pp. 129-141). Alberta, Canada: Society for Learning Analytics Research (SoLAR).

Ochoa, X., Chiluiza, K., Granda, R., Falcones, G., Castells, J., & Guamán, B. (2018). *Multimodal transcript of face-to-face group-work activity around interactive tabletops.* Paper presented at the 3rd International Workshop on Learning Analytics Across Spaces - CrossLAK'18, Sydney, Australia.

Ochoa, X., Chiluiza, K., Mendez, G., Luzardo, G., Guam, B., & Castells, J. (2013). *Expertise estimation based on simple multimodal features.* Paper presented at the Proceedings of the 15th ACM on International conference on multimodal interaction, Sydney, Australia.

Ochoa, X., Chiluiza, K., Méndez, G., Luzardo, G., Guamán, B., & Castells, J. (2013). *Expertise estimation based on simple multimodal features.* Paper presented at the 15th ACM on International conference on multimodal interaction.

Ochoa, X., Domínguez, F., Guamán, B., Maya, R., Falcones, G., & Castells, J. (2018). *The rap system: Automatic feedback of oral presentation skills using multimodal analysis and low-cost sensors.* Paper presented at the 8th International Conference on Learning Analytics and Knowledge, LAK'18, Sydney, AUS.

Ochoa, X., Weibel, N., Worsley, M., & Oviatt, S. (2016). *Multimodal learning analytics data challenges.* Paper presented at the 6th International Conference on Learning Analytics and Knowledge, LAK 2016.

Oh, J., Chee, Y., Lim, T., Cho, Y., & Kim, I. Y. (2014). Chest compression with kneeling posture in hospital cardiopulmonary resuscitation: A randomised crossover simulation study. *Emergency Medicine Australasia, 26*(6), 585-590.

Ohland, M. W., Loughry, M. L., Woehr, D. J., Bullard, L. G., Felder, R. M., Finelli, C. J., . . . Schmucker, D. G. (2012). The comprehensive assessment of team member effectiveness: Development of a behaviorally anchored rating scale for self-and peer evaluation. *Academy of Management Learning & Education, 11*(4), 609-630.

Olguin, D. O., Gloor, P. A., & Pentland, A. (2009a). *Capturing individual and group behavior with wearable sensors.* Paper presented at the Human Behavior Modeling - Papers from the AAAI Spring Symposium, Stanford, CA.

Olguin, D. O., Gloor, P. A., & Pentland, A. (2009b). *Wearable sensors for pervasive healthcare management.* Paper presented at the 2009 3rd International Conference on Pervasive Computing Technologies for Healthcare.

Olson, J. S., Teasley, S., Covi, L., & Olson, G. (2002). The (currently) unique advantages of collocated work. In P. J. Hinds & S. Kiesler (Eds.), *Distributed work: New research on working across distance using technology* (pp. 113-136). Cambridge: MIT Press.

Oviatt, S. (2018). *Ten opportunities and challenges for advancing student-centered multimodal learning analytics.* Paper presented at the International Conference on Multimodal Interaction, ICMI'18.

Oviatt, S., & Cohen, A. (2013). *Written and multimodal representations as predictors of expertise and problem-solving success in mathematics.* Paper presented at the International Conference on Multimodal Interaction, ICMI'13, Sydney, Australia.

Page, K. R., Michaelides, D. T., Buckingham Shum, S. J., Chen-Buger, Y.-H., Dalton, J., Eisenstadt, M., . . . Komzak, J. (2005). Collaboration in the semantic grid: A basis for e-learning. *Applied Artificial Intelligence\, 19*\(9-10\), 881\-904\.

Palominos, E., Levett-Jones, T., Power, T., & Martinez-Maldonado, R. (2019). Healthcare students' perceptions and experiences of making errors in simulation: An integrative review. *Nurse Education Today, 77*, 32-39.

Panitz, T. (1999). Collaborative versus cooperative learning: A comparison of the two concepts which will help us understand the underlying nature of interactive learning.

Pardo, A. (2018). A feedback model for data-rich learning experiences. *Assessment & Evaluation in Higher Education, 43*(3), 428-438.
Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology, 50*(1), 128-138.
Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology, 45*(3), 438-450.
Parsons, H. M. (1974). What happened at hawthorne? *Science, 183*(4128), 922.
Paulus, P. (2000). Groups, teams, and creativity: The creative potential of idea-generating groups. *Applied psychology, 49*(2), 237-262.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Dubourg, V. (2011). Scikit-learn: Machine learning in python. *Journal of machine learning research, 12*(Oct), 2825-2830.
Perera, D., Kay, J., Koprinska, I., Yacef, K., & Zaïane, O. R. (2008). Clustering and sequential pattern mining of online collaborative learning data. *IEEE Transactions on Knowledge and Data Engineering, 21*(6), 759-772.
Petrosoniak, A., Almeida, R., Pozzobon, L. D., Hicks, C., Fan, M., White, K., . . . Trbovich, P. (2018). 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*.
Phielix, C., Prins, F. J., Kirschner, P. A., Erkens, G., & Jaspers, J. (2011). Group awareness of social and cognitive performance in a cscl environment: Effects of a peer feedback and reflection tool. *Computers in Human Behavior, 27*(3), 1087-1102.
Picard, R. W. (2000). *Affective computing*: MIT Press.
Pijeira-Diaz, H. J., Drachsler, H., Järvelä, S., & Kirschner, P. A. (2016). *Investigating collaborative learning success with physiological coupling indices based on electrodermal activity*. Paper presented at the Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, Edinburgh, United Kingdom.
Prieto, L. P., Rodríguez Triana, M. J., Martínez Maldonado, R., Dimitriadis, Y. A., & Gašević, D. (2018). Orchestrating learning analytics (orla): Supporting inter-stakeholder communication about adoption of learning analytics at the classroom level. *Australasian Journal of Educational Technology, 35*(4), 14-33.
Prieto-Alvarez, C. G., Martinez-Maldonado, R., & Anderson, T. (2017). Co-designing in learning analytics: Tools and techniques. In J. C. H. J. M. Lodge, and L. Corrin (Eds.) (Ed.), *Learning analytics in the classroom: Translating learning analytics research for teachers (in press).*
Prieto-Alvarez, C. G., Martinez-Maldonado, R., & Anderson, T. D. (2018). Co-designing learning analytics tools with learners. *Learning Analytics in the Classroom: Translating Learning Analytics for Teachers.*
Prieto-Alvarez, C. G., Martinez-Maldonado, R., & Buckingham Shum, S. (2020). *La-deck: A card-based learning analytics co-design tool.* Paper presented at the 10th International Conference on Learning Analytics and Knowledge (LAK'20), Frankfurt, Germany.
Prieto-Alvarez, C. G., Martinez-Maldonado, R., & Shum, S. B. (2018). *Mapping learner-data journeys: Evolution of a visual co-design tool.* Paper presented at the 30th Australian Conference on Computer-Human Interaction, Melbourne, Australia.
Pugh, L. A., Oldroyd, C. R., Ray, T. S., & Clark, M. L. (1966). Muscular effort and electrodermal responses. *Journal of experimental psychology, 71*(2), 241-248.
Qarabash, H., Heslop, P., Kharrufa, A., Balaam, M., & Devlin, M. (2019). Group tagging: Using video tagging to facilitate reflection on small group activities. *British Journal of Educational Technology, 50*(4), 1913-1928.
R. Ruis, A., Pozen, R., Eagan, B. R., & Shaffer, D. W. (2019). *Finding common ground: A method for measuring recent temporal context in analyses of complex, collaborative thinking*. Paper presented at the 13th International Conference on Computer Supported Collaborative Learning, CSCL 2019, Lyon, France.
Raca, M., Tormey, R., & Dillenbourg, P. (2014). *Sleepers' lag-study on motion and attention.* Paper presented at the Fourth International Conference on Learning Analytics And Knowledge, LAK'14, Indiana, USA.
Reeves, L. M., Lai, J., Larson, J. A., Oviatt, S., Balaji, T. S., Buisine, S., . . . Martin, J. C. (2004). Guidelines for multimodal user interface design. *Communications of the ACM, 47*(1), 57-59.
Reeves, T. C. (2000). Enhancing the worth of instructional technology research through "design experiments" and other development research strategies.
Reimers, G., & Neovesky, A. (2015). Student focused dashboards. *7th International Conferenceon Computer Supported Education (Lisbon).*
Ren, D., Brehmer, M., Lee, B., Höllerer, T., & Choe, E. K. (2017). *Chartaccent: Annotation for data-driven storytelling*. Paper presented at the 2017 IEEE Pacific Visualization Symposium (PacificVis).
Riebe, L., Girardi, A., & Whitsed, C. (2016). A systematic literature review of teamwork pedagogy in higher education. *Small Group Research, 47*(6), 619-664.

Roberts, J. C., Ritsos, P. D., Jackson, J. R., & Headleand, C. (2017). The explanatory visualization framework: An active learning framework for teaching creative computing using explanatory visualizations. *IEEE Transactions on Visualization and Computer Graphics, 24*(1), 791-801.

Roberts, L. D., Howell, J. A., & Seaman, K. (2017). Give me a customizable dashboard: Personalized learning analytics dashboards in higher education. *Technology, Knowledge and Learning, 22*(3), 317-333.

Rodríguez-Triana, M. J., Martínez-Monés, A., Asensio-Pérez, J. I., & Dimitriadis, Y. (2015). Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrating cscl situations. *British Journal of Educational Technology, 46*(2), 330-343.

Rodríguez-Triana, M. J., Prieto, L. P., Martínez-Monés, A., Asensio-Pérez, J. I., & Dimitriadis, Y. (2018). *The teacher in the loop: Customizing multimodal learning analytics for blended learning.* Paper presented at the 8th international conference on learning analytics and knowledge, Sydney, AUS.

Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics, 2*(1), 7-12.

Rosen, M. A., Salas, E., Wilson, K. A., King, H. B., Salisbury, M., Augenstein, J. S., . . . Birnbach, D. J. (2008). Measuring team performance in simulation-based training: Adopting best practices for healthcare. *Simulation in Healthcare, 3*(1), 33-41.

Roth, W.-M., Tenenberg, J., & Socha, D. (2016). Discourse/s in/of cscw. *Computer Supported Cooperative Work, 25*(4-5), 385-407.

Rousseau, V., Aubé, C., & Savoie, A. (2006). Teamwork behaviors. *Small Group Research, 37*(5), 540-570.

Ruiz-Robledillo, N., & Moya-Albiol, L. (2015). Lower electrodermal activity to acute stress in caregivers of people with autism spectrum disorder: An adaptive habituation to stress. *Journal of Autism and Developmental Disorders, 45*(2), 576-588.

Rummel, N., Spada, H., & Hauser, S. (2009). Learning to collaborate while being scripted or by observing a model. *International Journal of Computer-Supported Collaborative Learning, 4*(1), 69-92.

Ryan, L. (2016). *The visual imperative : Creating a visual culture of data discovery*: Elsevier Science.

Salas, E., Burke, C. S., & Cannon-Bowers, J. A. (2000). Teamwork: Emerging principles. *International Journal of Management Reviews, 2*(4), 339-356.

Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On teams, teamwork, and team performance: Discoveries and developments. *Human Factors, 50*(3), 540-547.

Salas, E., Shuffler, M. L., Thayer, A. L., Bedwell, W. L., & Lazzara, E. H. (2015). Understanding and improving teamwork in organizations: A scientifically based practical guide. *Human Resource Management, 54*(4), 599-622.

Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a "big five" in teamwork? *Small Group Research, 36*(5), 555-599.

Schecter, A., Pilny, A., Leung, A., Poole, M. S., & Contractor, N. (2018). Step by step: Capturing the dynamics of work team process through relational event sequences. *Journal of Organizational Behavior, 39*(9), 1163-1181.

Scheffel, M., Drachsler, H., Toisoul, C., Ternier, S., & Specht, M. (2017). *The proof of the pudding: Examining validity and reliability of the evaluation framework for learning analytics.* Paper presented at the European Conference on Technology Enhanced Learning, EC-TEL 2017.

Schmidt, K., & Bannon, L. (2013). Constructing cscw: The first quarter century. *Computer Supported Cooperative Work (CSCW), 22*(4), 345-372.

Schneider, B., & Blikstein, P. (2015). Unraveling students' interaction around a tangible interface using multimodal learning analytics. *Journal of Educational Data Mining, 7*(3), 89-116.

Schneider, B., & Pea, R. (2017). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. In *Educational media and technology yearbook* (pp. 99-125): Springer.

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (2016). *Detecting collaborative dynamics using mobile eye-trackers.* Paper presented at the Transforming Learning, Empowering Learners: The International Conference of the Learning Sciences (ICLS) 2016, Singapore.

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (2018). Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. *International Journal of Computer-Supported Collaborative Learning.*

Schneider, J., Di Mitri, D., Limbu, B., & Drachsler, H. (2018). *Multimodal learning hub: A tool for capturing customizable multimodal learning experiences.* Paper presented at the Lifelong Technology-Enhanced Learning, Cham.

Schulz, H.-J., Streit, M., May, T., & Tominski, C. (2013). *Towards a characterization of guidance in visualization.* Paper presented at the IEEE Conference on Information Visualization (InfoVis). Atlanta, Georgia, USA.

Schumacher, C., & Ifenthaler, D. (2017). Features students really expect from learning analytics. *Computers in Human Behavior.*

Schwendimann, B. A., Rodríguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., . . . Dillenbourg, P. (2017). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies, 10*(1), 30-41.

Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2018). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior.*

Sedrakyan, G., Mannens, E., & Verbert, K. (2019). Guiding the choice of learning dashboard visualizations: Linking dashboard design and data visualization concepts. *Journal of Computer Languages, 50*, 19-38.

Segel, E., & Heer, J. (2010). Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics, 16*(6), 1139-1148.

Shaffer, D. W. (2017). *Quantitative ethnography.* Madison, WI: WI:Cathcart Press.

Sharma, K., Papamitsiou, Z., & Giannakos, M. (2019). Building pipelines for educational data using ai and multimodal analytics: A "grey-box" approach. *British Journal of Educational Technology, 50*(6), 3004-3031.

Shibani, A., Knight, S., & Shum, S. B. (2019). *Contextualizable learning analytics design: A generic model and writing analytics evaluations.* Paper presented at the Proceedings of the 9th International Conference on Learning Analytics & Knowledge, Tempe, AZ.

Shute, V. J., & Ventura, M. (2013). *Stealth assessment: Measuring and supporting learning in video games*: MIT Press.

Siemens, G., & Baker, R. S. J. d. (2012). *Learning analytics and educational data mining: Towards communication and collaboration.* Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada.

Smith, B. L. (1992). *What is collaborative learning.* Paper presented at the Collaborative Learning: A Sourcebook for Higher Education, Pennsylvania State University.

Smith, S. L., & Mosier, J. N. (1986). *Guidelines for designing user interface software.* Retrieved from

Soller, A., Martínez, A., Jermann, P., & Muehlenbrock, M. (2005). From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *International Journal of Artificial Intelligence in Education (IOS Press), 15*(4), 261-290.

Sommerville, I. (2011). Software engineering 9th edition.

Spikol, D., Ruffaldi, E., Dabisias, G., & Cukurova, M. (2018). Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. *Journal of Computer Assisted Learning, 34*(4), 366-377.

Spikol, D., Ruffaldi, E., Landolfi, L., & Cukurova, M. (2017). *Estimation of success in collaborative learning based on multimodal learning analytics features.* Paper presented at the 2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT).

Srinivasan, A., Drucker, S. M., Endert, A., & Stasko, J. (2018). Augmenting visualizations with interactive data facts to facilitate interpretation and communication. *IEEE Transactions on Visualization and Computer Graphics, 25*(1), 672-681.

Stahl, G. (2006). Group cognition: Computer support for building collaborative knowledge. Acting with technology series. In: MIT Press.

Standardization, I. O. f. (2010). *Ergonomics of human-system interaction: Part 210: Human-centred design for interactive systems.* Retrieved from

Stein-Parbury, J. (2013). *Patient and person: Interpersonal skills in nursing*: Elsevier Health Sciences.

Stewart, C. (2017). Learning analytics: Shifting from theory to practice. *Journal on Empowering Teaching Excellence, 1*(1), 10.

Strijbos, J.-W. (2010). Assessment of (computer-supported) collaborative learning. *IEEE Transactions on Learning Technologies, 4*(1), 59-73.

Strijbos, J.-W., & Sluijsmans, D. (2010). Unravelling peer assessment: Methodological, functional, and conceptual developments. In: Elsevier.

Sturm, J., Herwijnen, O. H. V., Eyck, A., & Terken, J. (2007). *Influencing social dynamics in meetings through a peripheral display.* Paper presented at the International Conference on Multimodal Interfaces, ICMI'07, Nagoya. Conference Paper retrieved from

Sung, H., Cao, S., R. Ruis, A., & Shaffer, D. W. (2019). *Reading for breadth, reading for depth: Understanding the relationship between reading and complex thinking using epistemic network analysis.* Paper presented at the 13th International Conference on Computer Supported Collaborative Learning, CSCL 2019, Lyon, France.

Sveen, C., Klerkx, J., & Duval, E. (2014). Learning analytics dashboards. *Journal of Learning Analytics.*

Talafian, H., Sha, M., Barany, A., & Foster, A. (2019). *Promoting and tracing high school students' identity change in an augmented virtual learning environment.* Paper presented at the 13th International Conference on Computer Supported Collaborative Learning, CSCL 2019, Lyon,France.

Taras, M. (2005). Assessment–summative and formative–some theoretical reflections. *British journal of educational studies, 53*(4), 466-478.
Tausch, S., Ta, S., & Hussmann, H. (2016). *A comparison of cooperative and competitive visualizations for co-located collaboration.* Paper presented at the SIGCHI Conference on Human Factors in Computing Systems, CHI'16, Santa Clara, California, USA.
Taylor, S., Jaques, N., Chen, W., Fedor, S., Sano, A., & Picard, R. (2015). *Automatic identification of artifacts in electrodermal activity data.* Paper presented at the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).
Teasley, S. D. (2017). Student facing dashboards: One size fits all? *Technology, Knowledge and Learning, 22*(3), 377-384.
Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation, 27*(2), 237-246.
Thomas, J. J., & Cook, K. A. (2006). A visual analytics agenda. *IEEE Computer Graphics and Applications*(1), 10-13.
Toker, D., Conati, C., Steichen, B., & Carenini, G. (2013). *Individual user characteristics and information visualization: Connecting the dots through eye tracking*. Paper presented at the proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
Tsai, Y.-S., & Gasevic, D. (2017). *Learning analytics in higher education---challenges and policies: A review of eight learning analytics policies.* Paper presented at the Seventh International Learning Analytics & Knowledge Conference, LAK'17.
Tufte, E. R. (2001). *The visual display of quantitative information* (Vol. 2): Graphics press Cheshire, CT.
Tufte, E. R., & Schmieg, G. M. (1985). *The visual display of quantitative information* (Vol. 53): American Association of Physics Teachers.
van Leeuwen, A., Rummel, N., & van Gog, T. (2019). What information should cscl teacher dashboards provide to help teachers interpret cscl situations? *International Journal of Computer-Supported Collaborative Learning.*
Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist, 57*(10), 1500-1509.
Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education, 122*, 119-135.
Viswanathan, S. A., & VanLehn, K. (2017a). *High accuracy detection of collaboration from log data and superficial speech features.* Paper presented at the Making a Difference: Prioritizing Equity and Access in CSCL, 12th International Conference on Computer Supported Collaborative Learning (CSCL) 2017, Philadelphia, PA.
Viswanathan, S. A., & VanLehn, K. (2017b). Using the tablet gestures and speech of pairs of students to classify their collaboration. *IEEE Transactions on Learning Technologies, 11*(2).
Vygotski, L. S. (1987). *The collected works of ls vygotsky. Volume 1: Problems of general psychology, including the volume thinking and speech* (Vol. 1). New York: Springer
Wagner, J., Lingenfelser, F., Baur, T., Damian, I., Kistler, F., & André, E. (2013). *The social signal interpretation (ssi) framework: Multimodal signal processing and recognition in real-time.* Paper presented at the 21st ACM international conference on Multimedia.
Wang, D., Yang, Q., Abdul, A., & Lim, B. Y. (2019). *Designing theory-driven user-centric explainable ai.* Paper presented at the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK.
Wang, G., Zhang, X., Tang, S., Zheng, H., & Zhao, B. Y. (2016). *Unsupervised clickstream clustering for user behavior analysis.* Paper presented at the SIGCHI Conference on Human Factors in Computing Systems, CHI'16.
Ward, J. A., Pirkl, G., Hevesi, P., & Lukowicz, P. (2017). *Detecting physical collaborations in a group task using body-worn microphones and accelerometers.* Paper presented at the 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops).
Webb, N. M. (2009). The teacher's role in promoting collaborative dialogue in the classroom. *British Journal of Educational Psychology, 79*(1), 1-28.
Weston, W. W., & Brown, J. B. (2003). Developing a patient-centered curriculum. In M. Stewart (Ed.), *Patient-centered medicine: Transforming the clinical method* (pp. 229-241): Radcliffe Publishing.
Wilson, J. M., Straus, S. G., & 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*(1), 16-33.
Wise, A., & Cui, Y. (2018). *Envisioning a learning analytics for the learning sciences.* Paper presented at the 13th International Conference of the Learning Sciences.
Wise, A. F., & Hsiao, Y.-T. (2019). Self-regulation in online discussions: Aligning data streams to investigate relationships between speaking, listening, and task conditions. *Computers in Human Behavior, 96*, 273-284.

Wise, A. F., Knight, S., & Buckingham Shum, S. (In Press). Collaborative learning analytics. In U. Cress, C. Rosé, A. Wise, & J. Oshima (Eds.), *International handbook of computer-supported collaborative learning*: Springer.

Wise, A. F., & Shaffer, D. W. (2015). Why theory matters more than ever in the age of big data. *Journal of Learning Analytics, 2*(2), 5-13.

Wolff, A., Moore, J., Zdrahal, Z., Hlosta, M., & Kuzilek, J. (2016). *Data literacy for learning analytics.*

Wong-Villacres, M., Chiluiza, K., Ortiz, M., & Echeverria, V. (2015). *A tabletop system to promote argumentation in computer science students.* Paper presented at the Proceedings of the 2015 ACM International Conference on Interactive Tabletops and Surfaces, ITS 2015.

Worsley, M. (2018). *Multimodal learning analytics past, present, and, potential futures.* Paper presented at the Second Multimodal Learning Analytics Across (Physical and Digital) Spaces (CrossMMLA).

Worsley, M., Abrahamson, D., Blikstein, P., Grover, S., Schneider, B., & Tissenbaum, M. (2016). *Situating multimodal learning analytics.* Paper presented at the International Conference on Learning Sciences.

Worsley, M., & Blikstein, P. (2014). *Deciphering the practices and affordances of different reasoning strategies through multimodal learning analytics.* Paper presented at the Proceedings of the 2014 ACM workshop on Multimodal Learning Analytics Workshop and Grand Challenge, Istanbul, Turkey.

Worsley, M., & Blikstein, P. (2018). A multimodal analysis of making. *International Journal of Artificial Intelligence in Education, 28*(3), 385-419.

Zhang, D., Gatica-Perez, D., Bengio, S., McCowan, I., & Lathoud, G. (2004). *Modeling individual and group actions in meetings: A two-layer hmm framework.* Paper presented at the 2004 Conference on Computer Vision and Pattern Recognition Workshop.

Zhang, L., Stoffel, A., Behrisch, M., Mittelstadt, S., Schreck, T., Pompl, R., . . . Keim, D. (2012). *Visual analytics for the big data era—a comparative review of state-of-the-art commercial systems.* Paper presented at the IEEE Conference on Visual Analytics Science and Technology (VAST), Seattle, Washington, USA.

Zhang, Z., & Sarcevic, A. (2015). *Constructing awareness through speech, gesture, gaze and movement during a time-critical medical task.* Paper presented at the ECSCW 2015: Proceedings of the 14th European Conference on Computer Supported Cooperative Work, 19-23 September 2015, Oslo, Norway, Cham.

Zoltan, R., & Vancea, R. (2015). Organizational work groups and work teams–approaches and differences. *Ecoforum Journal, 4*(1), 13.

Zolyomi, A., Ross, A. S., Bhattacharya, A., Milne, L., & Munson, S. A. (2018). *Values, identity, and social translucence: Neurodiverse student teams in higher education.* Paper presented at the SIGCHI Conference on Human Factors in Computing Systems, CHI'18, Montreal QC, Canada.

Appendix A

A.1 SIMULATION SCENARIO 1: STUDENT GUIDE

SIMULATION SCENARIO – 2ND YEAR BN: AUTUMN SEMESTER

UTS: Faculty of Health

STUDENT GUIDE

Title: Basic Life Support: Patient with chest pain and lethal cardiac rhythm

Background: This simulation will occur in two phases.

In phase one: a group of four students will be assessing and treating Mr Lars for chest pain. These RNs in different roles will communicate with Mr Lars, apply oxygen, assess his pain using PQRST, perform vital sign observations, administer Anginine according to the six rights, connect him to a three lead, identify his cardiac rhythm, document appropriately and call for a clinical review. Following the completion of phase one, a second team will undertake phase two.

In phase two: a second group of four students will take over Mr Lars's care at which point he will lose consciousness due to a fatal cardiac rhythm. The team will need to perform basic life support using the DRSABCD algorithm. Following the resuscitation attempt, the rest of the class in groups of four will rotate through the two phases. Student's not participating in the active simulation will be assessing each other on BLS.

This simulation is focused on team work and communication in the context of caring for a deteriorating patient eventually requiring basic life support.

Rationale for simulation:	Allows for exposure to reality based situations prior to attending clinical placement.
Focus: Consolidation of technical skills:	Management of an acutely deteriorating patient Taking and assessing vital signs Oxygen therapy Oral medication administration 3 lead ECG Rhythm interpretation Basic life support Correct documentation on all associated charts
Target audience:	Bachelor of Nursing – 2nd year students
Subject/ Course:	92322 Medical Surgical Nursing
Aim:	To understand the specific elements in caring for a patient with chest pain who then requires basic life support.

Learning Objectives:
At the end of this simulation, participants will have learnt:

- To assess and respond to the specific needs of a deteriorating patient with chest pain or a lethal rhythm requiring basic life support.

INTERNAL USE Page 1

How did it feel participating in the simulation in your role?

How did you think the team worked together?

How has the simulated experience helped with your preparation for clinical practice?

What will you each take away as something you have learnt from this simulation?

Phase 2: Simulation Running Sheet (Fatal Rhythm)

Scenario:
You are coming onto morning shift. The night team were admitting Mr Lars to the ward when he began to deteriorate. As you reach his bedside he loses consciousness.

Care already provided:

- Mr Lars has had his chest pain assessed.
- He has been administered one dose of Anginine with no relief.
- He is scoring his pain at 11/10. He describes the pain as crushing and it is radiating. He says it is unlike his normal angina pain.
- The clinical review has been upgraded to a rapid response call

Care to be provided:

- Mr Lars's cardiac rhythm will change and he will lose consciousness requiring basic life support according to the DRSABCD algorithm.

Participant Role	Actions / communication to be demonstrated
RN1	• Discover emergency • Check for danger • Check for response (verbal then Trapezius squeeze) • Hit the emergency buzzer and note the time • Clear the airway then perform head tilt/jaw thrust • Verbalise spinal precautions and compression/breath rate • Bag patient using Air-viva noting the rise and fall of the chest after every set of compressions
RN2	• Commence compressions • Verbalise depth and rate of compressions • Perform compressions/count aloud to ensure bagging
RN 3	• Complete MET Form
RN 4	• Get resus trolley • Attach AED/follow prompts • Maintain team safety/clear team before shock

On completion of this scenario, discuss with your tutor:
How did it feel participating in the simulation in your role?

How did you think the team worked together?

How has the simulated experience helped with your preparation for clinical practice?

What will you each take away as something you have learnt from this simulation?

A.2 PARTICIPANT INFORMATION SHEET

Roberto Martinez-Maldonado
Connected Intelligence Centre
Office of the DVC (Education & Students)
University of Technology Sydney

Rm CB02.07.088, PO Box 123, Broadway
NSW 2007, Sydney
T: +61 2 9514 4524 M: +61 407 862578
Email: Roberto.Martinez-Maldonado@uts.edu.au
Web: utscic.edu.au

INFORMATION SHEET

ETH18-2278- Learning Analytics in clinical simulation

WHO IS DOING THE RESEARCH?

This research is being undertaken as a joint venture between the Faculty of Health and the Connected Intelligence Centre. The researchers are Professor Simon Buckingham Shum, Dr Roberto Martinez Maldonado, Vanessa Echeverria, Dr Tamara Power, Mrs Carolyn Hayes, Mr, Carlos Prieto and Ms Jan Forber.

WHAT IS THIS RESEARCH ABOUT?

This project aims to support student learning through gathering data about simulation experiences.

As a result, this research has four main goals:
1) to collect data about students' interactions with the manikins and their peers;
2) to generate understanding of group collaboration during simulations;
3) explore ways to provide feedback to students about their activities in the clinical laboratories;
4) look at patterns of group collaboration that are associated with better outcomes.

WHY HAVE I BEEN ASKED?

Because your input can provide researchers with key information to address the goals of the project.

IF I SAY YES, WHAT WILL IT INVOLVE?

If you agree to participate you may be invited to participate in one or two sessions to perform a simulation task. Data will be collected from available technologies you interact with such as manikins, sim pads and iPads in the lab. We would seek your permission to use this data and audio-visual material in conference presentations and publications.

ARE THERE ANY RISKS/INCONVENIENCE?

There are very few if any risks because the research has been carefully designed. If you agree to participate in the sessions, you will be inconvenienced by giving up two hours of your time. To reduce inconvenience, we would try to schedule this on a day where you were on campus anyway.

DO I HAVE TO SAY YES?

You don't have to say yes. Participation is entirely voluntary and participation or non-participation will have no effect on your grades.

WHAT WILL HAPPEN IF I SAY NO?

If you decide not to participate, it will not affect your relationship with the researchers or the University of Technology Sydney. If you wish to withdraw from the study once it has started, you can do so at any time without having to give a reason. If you withdraw from the study, all data collected from the session will be destroyed.

CONFIDENTIALITY

By signing the consent form you consent to the research team collecting and using personal information about you for the research project. All this information will be treated confidentially. Your information will only be used for the purpose of this research project and it will only be disclosed with your permission, except as required by law. In any publication, information will be provided in such a way that you cannot be identified.

WHAT IF I HAVE CONCERNS OR A COMPLAINT?

If you have concerns about the research that you think I can help you with, please feel free to contact us on Simon.BuckinghamShum@uts.edu.au or Roberto.Martinez-Maldonado@uts.edu.au.

You will be given a copy of this form to keep.

NOTE:

This study has been approved by the University of Technology Sydney Human Research Ethics Committee [ETH18-2278]. If you have any concerns or complaints about any aspect of the conduct of this research, please contact the Ethics Secretariat on ph.: +61 2 9514 2478 or email: Research.Ethics@uts.edu.au], and quote the UTS HREC ETH18-2278. Any matter raised will be treated confidentially, investigated and you will be informed of the outcome.

A.3 SIMULATION SCENARIO 2: HANDOVER, MANIKIN SET-UP AND PATIENT PROMPTS

Case 1 part two (evening shift)

Date: 30th July 2018 **Time:** 1400 hrs

You are a new graduate nurse within a medical surgical unit. You and your team are to continue to plan and provide care for Mr Jim Cooper who was admitted this morning from the emergency department.

Bedside handover is as follows: (read by Lab session leader)

Identification: Jim Cooper, 68 year old male

Situation:
Jim presented early this morning with a 1 week history of productive cough and SOB more than usual. He has been to the GP for Oral AB. Reports having to sit in a chair all night the past few nights and is finding it more difficult to care for himself at home. He lives alone

Past medical history:
COPD diagnosed 5 years ago.
GORD
Hypercholesterolaemia.

Assessment:
Vital signs at 1000 hrs on admission to the respiratory medical unit

HR	112 beats/min
RR	30 breaths/min
SaO_2	89% RA – 93% 2L O_2 NP
BP	142/88 mmHg
Temp	37.8˚C oral

Auscultation: Scattered wheezes throughout both lung fields.

Recommendation:
Treatment with oxygen, IV fluids, IVAB, Prednisolone and physiotherapy.

Case 1 Part Two (evening shift) - Manikin set-up and patient prompts

Deterioration scenario

Jim develops a reaction to the IVAB shortly after it is commenced. Do not tell the team this is happening. You need to provide responses for the team in the role as Jim and the manikin settings will need to be altered.

Begin by informing the team you are beginning to feel dizzy and unwell – nausea.

A little while later, or if questioned, tell the team your chest feels tight and you are developing discomfort – DO NOT SAY THE WORD EPIGASTRIC as Jim would not use this word, just point at your or the manikins epigastric area.

Do not provide clinical information below until the team have actually performed the assessment or asked for the information.
Where the manikin generates the vital signs, the team should get the result from the manikin e.g. HR, RR, breath sounds, BP etc.

	Look	Listen	Feel
Airway	• Patent • Pursed lip breathing	Only talking in one word responses	
Breathing	• 30 breaths per minute • 88% RA • 90% when O2 insitu • uses accessory muscles of the neck		
Circulation	• HR 130 beats/min • BP 150/95 mmHg • cap refill 3 sec		• Temp 38˚C Oral • Warm peripheries
Disability	• Alert – eyes open • Anxious looking	Orientated to person, place and time	
Exposure	• Pale • skin is thin and bruised		
Fluids	• Pale dry mucous membranes		Poor skin turgor
Glucose	• BGL – 6.8 mmol/L		

Appendix B

B.1 TEMPLATE TO SUMMARISE THE LEARNING DESIGN AND CRITICAL INCIDENTS

1. Simulation name:
2. List the actions that nursing students will perform:
   - Action A
   - Action B
3. List the actions or responses that the patient may present during the simulation (e.g. heart rate change):
   - Patient Action A
   - Patient Action B
4. List at least two critical moments that would be interesting to explore in this scenario:
   - E.g. Patient with cardiac arrest
   - Critical moment B

5. How many students will be part of the team?
6. List the name of the roles (if necessary):
   - RN1: Team leader
   - RN2
   - RN3
7. Rule definition for time response:

- Is there any action that need to be performed timely? If so, please fill in the following information:

Action	Expected time
e.g. Vital signs	first 3 mins from the beginning of the simulation

8. Rule definition for errors:

- Is there any particular order of actions to be performed by the team? If so, please fill in the following information:

Action	Order
e.g. Blood pressure	2
e.g. ECG	1

- Is there any error that could be flagged in terms of team's positioning? (e.g. the team leader should be always near the patient). If so, please fill in the following information:

Role	Positioning
e.g. Team leader	Near the patient's bed

9. Additional information: