Towards Collaboration Translucence: Giving Meaning to Multimodal Group Data

Vanessa Echeverria[1,2], Roberto Martinez-Maldonado[1] and Simon Buckingham Shum[1]
[1]University of Technology Sydney, Ultimo, NSW, 2007, Australia.
[2]Escuela Superior Politécnica del Litoral, ESPOL, Campus Gustavo Galindo Km 30.5 Vía Perimetral, Guayaquil, Ecuador.
Vanessa.EcheverriaBarzola, Roberto.Martinez-Maldonado, Simon.BuckinghamShum}@uts.edu.au

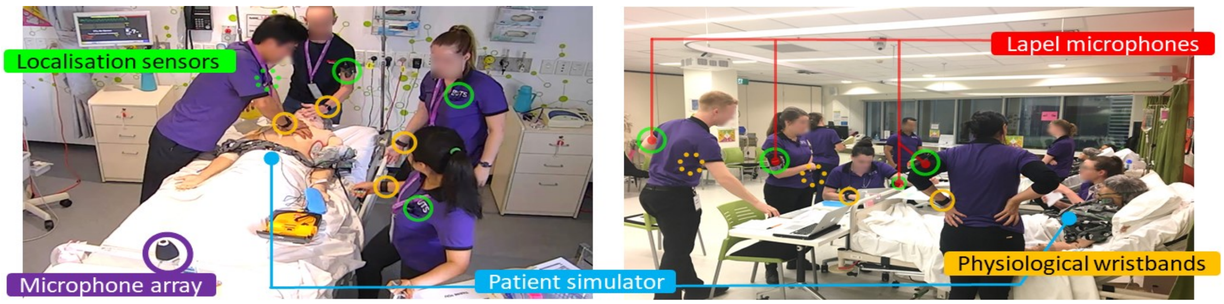

Figure 1: Multimodal analytics in healthcare scenarios: in a simulation-controlled room (left) and in the classroom (right).

ABSTRACT

Collocated, face-to-face teamwork remains a pervasive mode of working, which is hard to replicate online. Team members' embodied, multimodal interaction with each other and artefacts has been studied by researchers, but due to its complexity, has remained opaque to automated analysis. However, the ready availability of sensors makes it increasingly affordable to instrument work spaces to study teamwork and groupwork. The possibility of visualising key aspects of a collaboration has huge potential for both academic and professional learning, but a frontline challenge is the enrichment of quantitative data streams with the qualitative insights needed to make sense of them. In response, we introduce the concept of collaboration translucence, an approach to make visible selected features of group activity. This is grounded both theoretically (in the physical, epistemic, social and affective dimensions of group activity), and contextually (using domain-specific concepts). We illustrate the approach from the automated analysis of healthcare simulations to train nurses, generating four visual proxies that fuse multimodal data into higher order patterns.

ACM ISBN 978-1-4503-5970-2/19/05...$15.00.
https://doi.org/10.1145/3290605.3300269

CCS CONCEPTS

• Human-centered computing → Collaborative interaction

KEYWORDS

CSCW; collaboration; learning analytics; pervasive computing

ACM Reference format:

Vanessa Echeverria, Roberto Martinez-Maldonado, and Simon Buckingham Shum.. 2019. Towards Collaboration Translucence: Giving Meaning to Multimodal Group Data. In *Proceedings of ACM CHI conference (CHI'19).* ACM, New York, NY, USA, Paper 39, 16 pages. https://doi.org/10.1145/3290605.3300269

1 INTRODUCTION

The ability to communicate, be an effective team/group member and collaborate face-to-face (f2f) are key skills for employability in the 21st century workplace [10]. Collaborating in collocated (f2f) settings provides unique benefits that are not easy to achieve in digitally mediated forms of group work [36, 65]. Literature suggests that the rich, multimodal communication channels in f2f interaction can promote social bonding [60], increase creativity [29, 60] and productivity [65]. However, *learning to collaborate* effectively requires practice, awareness of group dynamics and reflection upon past activities [43]. Importantly, it often needs close coaching by an expert facilitator to foster beneficial collaborative interaction [28]. However, although having a coach or teacher closely supervising each group would be ideal, it may be unrealistic in practice.

A promising way to address the above would be to capture behavioural traces from group interactions. Significant effort has been invested in automatically mining digital traces of *online* experiences, where logs can be easily captured, to make group interactions visible, infer patterns of behaviour and broaden our understanding of group activity in various contexts [18]. By contrast, much more needs to be done to help researchers, designers, and users in understanding the complexity of collaboration and in finding ways to support collaboration or learn to collaborate more effectively in collocated settings.

The broad banner of *Educational Data Science* has been used to refer to areas of data-intensive research (e.g. intelligent tutoring systems - ITS, educational data mining, and learning analytics - LA) aimed at providing automated analysis and feedback based on learners' interaction logs and other sources of evidence. However, it has been recently reported that there is a disconnection between logged data and higher order educational constructs [55, 75]. More problematic, decisions can be made and actions can be taken based on misleading information or not acknowledging that logged data is, by definition, incomplete [39]. One of the main reasons for such disconnection is, not surprisingly, the intrinsic lack of meaning in data, and their representations, that are often used as proxies of activity mediated by different interfaces (e.g. in dashboards, recommender systems, feedback tools) [1, 26] or to conclude generalisations about activity without considering the context where it unfolds [80]. Moreover, LA and other data-intensive innovations may work well for well-defined, specific scenarios [24]. Yet, this approach fails when trying to support people in complex scenarios that involve higher order thinking, non-computer mediated interactions, or ill-defined, open tasks, such as in f2f team/group work situations.

Attempts to address the complexity of f2f collaboration through automated means exist but they have tended to reduce or over simplify the interaction space by slicing the activity, only looking at certain quantitative social aspects (e.g. [56]) or very specific parts of the task (e.g. micro formations [46], speech [56], gaze [73]). Sensors are becoming inexpensive and more readily available. However, the more sensors are used to log f2f activity the more complex the meaning making process to make these data into information becomes [12].

In this paper, we propose an approach to give meaning to multimodal group activity data. Inspired by the metaphor of social translucence [21], we propose a vision to make evidence of collaboration translucent. In doing so, we emphasise that collaboration also involves epistemic, physical (the use of tools, devices and the space) and affective aspects besides the social realm. We build on the notion of quantitative ethnography [74] to give meaning to sensor data based on knowledge that can help explain such data (e.g. domain knowledge, theory) in a specific context. We illustrate our approach in the context of simulation in nursing. Simulated scenarios are representative of situations in which it is critical for team members to reflect upon evidence on different aspects of their activity.

We use extracts from our research on group work in two authentic settings: an immersive simulation room (Figure 1, left) and a simulated hospital ward-classroom (right). In these, we capture rich multimodal data, including nurses' localisation, movement, physiological signals, actions, voice and video. We illustrate how ideas from quantitative ethnography can help reveal important aspects of collaboration, from low level logs, to meaningful higher order constructs. We present four prototypes as exemplar 'proxies' of collaboration. We analyse how three of these can play a key role as proxies in terms of social dynamics; embodied strategies; and emotional arousal. We also present a fourth proxy shown to actual students after authentic in-the-wild classroom sessions.

The rest of the paper is organised as follows. Next section provides an overview of the state-of-the-art in the area of collaboration analytics. Section 3 presents the principles underpinning our approach. In Section 4, we describe the studies that illustrate how our approach can be put into action for giving meaning to multimodal group data. In Section 5 we present exemplars of proxies that provide insights into the activity of nine teams of nursing students. We conclude with a discussion of the application of our approach and ideas for future work in Section 6.

2 RELATED WORK

2.1 State-of-the-art: Collaboration Analytics

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 [18]. For example, substantial effort has been posed in extracting quantitative *non-verbal speech* features from dialogue in contexts such as brainstorming [38, 78], problem-solving [56, 66, 81], and group meetings [2, 58, 59, 63, 77].

Other works have looked at a wider range of aspects of f2f collaboration. Some studies have analysed the relationship between traces of *physical activity* (e.g. proximity [26], motion [7] and posture [38]) 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 [77]; the author of meaningful utterances [63], or low rapport [58]. *Face* and *hand tracking* features, extracted from videos, have been used to analyse affective states [61], physical formations [76] and synchronicity [58]. Other studies have found that *physiological* features (e.g. electrodermal activity, hypoxic ventilatory response) may serve as indicators of affective states at individual [32, 61] and group levels [14, 32]. *Gaze* features have been used to identify participation styles [59] and assess the quality of group work [73]. Multitouch screens and input devices can also serve to log group activity [50]. For example, *touch events* have been be used to predict group performance [18, 48, 81] and *stroke* features from digital pens have been used to identify expertise [66] and to assess participation [59].

In terms of interfaces that communicate insights, prior research has mostly been limited to *mirroring* data for group members to decide what steps to take next. Most of these systems have offered simple representations of non-verbal *speech* features, to enhance awareness and accountability of verbal participation [2]; support self-reflection [78]; or promote social regulation [38]. Not many mirroring systems have included other modalities of data. Some exceptions include interfaces aimed at influencing group work by showing *gaze* data [73, 77]; *touch activity* [23, 81]; and quality scores of the group product [18].

Overall, the studies presented above show how multimodal data may serve to complement the analysis of f2f activity. 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 (e.g. social and task-related [18, 47, 66, 72], social and affect [32, 61], or social and physical aspects [7, 26, 76]). Most previous works have been conducted under controlled conditions (e.g. [7, 14, 32, 61, 76]) 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 [23, 56, 66] and the workplace [58, 63]. In sum, although some prior work has started to look at more than one aspect of collaboration, there is a timely need for a holistic approach that can help researchers and designers to associate multimodal group data with meaningful higher order constructs.

2.2 Multimodal Data and Meaning

A significant body of HCI work makes use of behavioural traces of human activity captured through sensors or input devices [17, 41]. As shown above, 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 [17]. Logs can be mined, for instance, to cluster user behaviours [44], identify archetypal users [83], and visualise common paths [44]. However, while logs can illuminate *what* users do, they often say much less about *why* [17]. 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 [74, 75, 84]. For example, far from the 'big data revolution' signalling the 'death of theory' [52], Wise and Shaffer [84] 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* [74].

Theoretically-motivated analytics can be designed in a principled manner for higher order constructs such as students' *"conscientiousness"* [75], or *"crowd-sourced learning"* capacity to learn in a MOOC [55]. 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 [13]); ii) automatically coding multimodal data according to an epistemic frame [85]; and

iii) examining co-occurrence of events or associating certain indicators of activity with collaboration, task performance and learning outcomes [14, 32, 76]. Shaffer [74] argues that "*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).

To summarise, we must go beyond solving the technical data fusion problem [42]; this is necessary, but not sufficient. The next challenge is to give meaning to multimodal data in terms of what people are actually doing.

3 APPROACH

In this section we present our approach to give meaning to multimodal group data.

3.1 Collaboration Translucence: Foundations

Erickson et al. [20, 21] developed the design approach of *social translucence*. This refers to computer-mediated systems that provide social cues that compensate for the loss of *visibility* (of socially significant information), *awareness* (of others' presence or actions) and *accountability* (of people's own visible actions) as a result of moving away from interaction in physical spaces into the digital realm. The term *translucence* foregrounds the intention of making *selective* aspects of activity visible. Translucent systems include the notion of a *social proxy*, a minimalist form of visualisation of people or their activities [21]. In the original literature [22], there were at least two types of social proxies: *real time proxies* (e.g. abstract representations of presence and participation in a chat room); and *summaries of information* over time that allowed reflection on past events to inform future actions (e.g. a timeline that showed past activity in a forum).

In the last two decades, the concept of social translucence has been adopted by several researchers in HCI and CSCW [53, 62], particularly in works focused on online networks (e.g. [27]) and small group work (e.g. [87]). Moreover, social translucence has been embraced as an analysis framework for eliciting design requirements (e.g. [87]). Bilandzic and Forth [4] and Niemantsverdriet et al. [62] argued that social translucence should also be applied to f2f situations, because these are becoming complex multi-user, hybrid spaces enriched with a number of digital devices and sensors. A critical question remains: what needs to be made visible, that already is not, in a collocated situation? Although awareness can be easily diminished in online systems [34], various aspects of collaboration can also get dimmed in the physical world [32]. For example, in classrooms, it is challenging for teachers to closely monitor multiple groups at once [37] and for students to evidence skills development. In other settings, keeping fluid awareness may also be challenging if the same members of the team are not present in all the opportunities of physical contact [71] such as in emergency rooms. Awareness can also wear-off as time passes, especially in fast-paced group settings [40] or in meetings [2].

Most work based on social translucence has been limited to only considering quantitative social aspects (both in online and in the few f2f settings [2, 15, 38]). This is despite the fact that Erickson et al. [20] originally suggested that future work should look at modelling more complex traits of group behaviour (e.g. by looking at the content of conversations) and also tracking and visualising social behaviours over time. In next section, we discuss theoretical perspectives that serve to provide a definition of collocated collaboration translucence.

3.2 Dimensions of collaboration

The social dimension is clearly critical in group work, but other dimensions are equally important. There are various theoretically-inspired, practical frameworks in HCI that decompose collaborative activity into multiple dimensions. For example, the activity-based computing (ABC) framework [3] decomposes group activity into *tasks*, *materials*, *time*, and *users*. The Blended Interaction framework [35] 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 [49] provided a three-dimensional view of group activity: (1) *set*, 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.

A fourth dimension not included in the frameworks above, but that has been identified in numerous multimodal group analytics studies, is 4) *the affective* dimension. Affective aspects have been identified in foundational theoretical work [82] as critical to understand CSCW work, even though it can often remain *invisible* [68]. As discussed above, it is possible to shed light on affective aspects using video recordings or physiological sensors.

To summarise, a preliminary definition of *collaboration translucence* should include not only social aspects, but also making evidence of collaboration translucent according to the multiple, intertwined dimensions of group activity: 1) physical, 2) social, 3) epistemic, and 4) affective). Before diving into possible data representations that can serve as collaboration proxies, we need to perform a first methodological step which would be to understand how to give meaning to complex, multimodal data.

3.3 The Multimodal Matrix

Quantitative ethnography [74] 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 [25]. Inspired by this concept, we introduce our approach to grounding quantitative data in the semantics derived from a qualitative interpretation of the context from which it arises. We propose a conceptual data representation termed the *multimodal matrix* (Figure 2), comprising the following conceptual elements: *dimensions of collaboration, multimodal observations, segments,* and *stanzas.*

Dimensions of collaboration. These explain the complexity of group activity (groups of columns in Figure 2). As an illustration, we selected the dimensions from the ACAD framework (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 key aspects of group activity can help researchers and designers to justify the emphasis of some kinds of connections over others [33] 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* that we will call multimodal observations (columns in Figure 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 one's perspective. Critically, each column should only contain one kind of information and only one term can be used in each (*ontological and terminological consistency* [74] 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 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* [74], segments are the smallest units of meaning considered for analysis (rows in Figure 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 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* [74] p. 167).

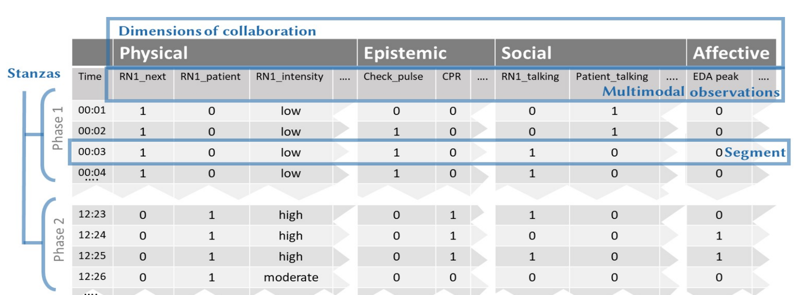

Figure 2: Multimodal matrix

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 Figure 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 [70]).

We next illustrate the application of our approach to give meaning to multimodal group data in two f2f settings.

4 ILLUSTRATIVE STUDIES

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), 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 implementing them in real life. Simulations are sometimes recorded and played back to students so that strengths and areas for improvement can be observed in facilitated debriefing sessions [31]. In practice, achieving this is challenging due to time and logistic constraints, hence the need for summary overview representations that can support reflection upon practice.

In this section we describe two studies in authentic nursing settings. Study 1 was conducted in an immersive simulation room, and Study 2 in a simulated hospital ward-classroom. Both were conducted within the framework of the *university's* curriculum which emphasises *patient-centred care* (PCC) and *teamwork* — higher order constructs that we target. This means that students must learn not only technical skills, but also develop communication skills to enable them to deliver professional care.

4.1 Study 1

This study was conducted as part of an optional program for nursing students to gain further experience in academic research and healthcare practice. As part of this program, a scenario was designed by a teacher in the context of caring for a patient requiring basic life support. Nine undergraduate nursing students (6 female and 3 male), aged from 20 to 53 years (avg= 34, std=10), participated in the study. According to their own time availability, they were randomly organised into three teams (A, B and C), of four students (2 females and 2 males), three (2 females and 1 male) and two (females) students each.

Learning design. The manikin was programmed by the teacher to deteriorate over time, dividing the task into two phases. Phase 1 involved the assessment of a patient's chest pain, including four sub-tasks: i) give oxygen therapy; ii) assess chest pain; iii) give medication; and iv) connect an electrocardiography (ECG) device. In Phase 2 students were expected to perform cardiopulmonary resuscitation (CPR) on the patient suffering a cardiac arrest. Each student was randomly asked to enact one of four roles (RN1-4) with an associated set of subtasks. RN1 was team leader, with RNs 2-4 responsible for subtasks ii, iii and iv, respectively. Depending on the number of students, the subtasks were distributed among the roles. Each simulation 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).

Apparatus. Sessions were conducted in an immersive simulation room. Students were the only people in the room (Figure 1, left). The room is 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. Students' localisation around the manikin was captured automatically through ultra-wide band wearable badges (Pozyx.io). These provide an error rate of 10 cm, recording x and y position of each student at 2Hz. We applied a Kalman filter to improve further calculations. Some student actions were automatically logged by the high-fidelity manikin (Laerdal Simman 3G), including: 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. A microphone array (Microcone) was used to detect nurses' conversations. Physiological wristbands (Empatica e4) were worn by students. These include a photoplethysmography sensor, an electrodermal activity sensor, a 3-axis accelerometer, and an optical thermometer. For this study, we automatically recorded EDA (at 4Hz) and acceleration (at 32 Hz) streams of data captured by the first two sensors. All the sessions were video-recorded. We synchronised the data streams at a 1 Hz, down sampling data streams from sensors that had a higher frequency.

Data gathering and analysis. Two researchers and a teacher were present in each session. Besides the data outlined above, other data gathering included observation notes and recordings of the group debriefing. These were

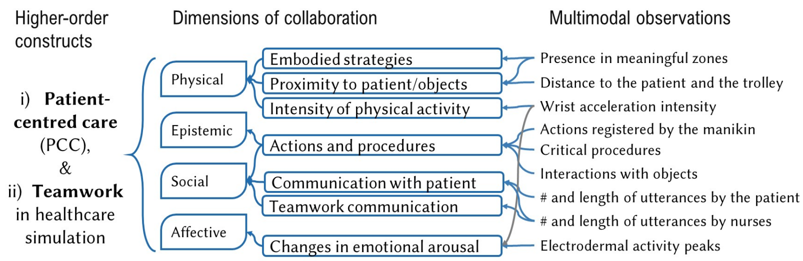

Figure 3: From multimodal logs to constructs in healthcare

transcribed for analysis. Data analysis involved two researchers independently screening the video recordings of the sessions looking for moments of interest that could serve to illustrate the potential use of the collaboration proxies (1, 2 and 3) and to describe individual nurses' behaviours.

4.2 Study 2

The second study was conducted in an authentic setting as part of the regular classes of the third-year undergraduate unit, *Integrated Nursing Practice*. Each class typically has 20-30 students working in teams of 4-6. Classrooms are equipped with 5-6 patient beds like the one used in Study 1 (Figure 1, right). The study focused on six 3-hour classes conducted in week 3 of semester 2, 2018. Only the activity occurring around one bed was recorded in each class, to allow 1 team to opt in for the study. Thus, the study involved six teams of 5 students each (teams 1-6), with a total of 30 students (27 female and 2 male) and 2 teachers. Due to limitations imposed by the authenticity of the context and privacy agreements, no other personal information was recorded.

Learning design. The coordinator of the unit designed a scenario that involved a patient experiencing an adverse drug reaction. Students were expected to perform the following tasks: i) assessment of chest pain symptoms, ii) administration of medication, iii) management of adverse drug reaction and iv) conducting an ECG. 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.

Apparatus. The manikins available in this setting (Laerdal Nursing Anne) are similar to the one used in Study 1 but of lower fidelity (the range of actions that it can log are limited to detect whether nurses check the patient's vital signs). To overcome this technical limitation, we developed a mobile observation tool, synchronised with the other sensors, for a researcher to log all the critical actions, as defined by the teacher, including: placing the oxygen mask, preparing intravenous (IV) fluids, administering IV fluids, reading blood pressure, writing observations in charts, attaching the ECG device, stopping the IV fluids and calling the doctor. Each of the four active students in the scenario worn a localisation badge and a physiological wristband as in Study 1. Individual audio was captured using lapel microphones. The student enacting the patient was not tracked as it was commonly sitting on a chair at the bed side. The teacher was also tracked but her/his data were not used for the purpose of this analysis.

Data gathering and analysis. Two researchers were present in each classroom session while one of the two teachers delivered the regular class. In addition to the multimodal data gathering, each team was invited to participate in a 30-minutes semi-structured follow-up interview a week after their class (week 4). Out of the 30 students, 18 participated in these sessions in which they explored one of the collaboration proxies (proxy 4 to be presented in section 5.4). We structured the interview into three main parts, using three key constructs from social translucence [87] to structure the interview: 1) *visibility*, in which students were expected to re-construct their activity based on the proxy and they 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 3) *accountability*, in which students were asked about who else should have access to the proxy and for what purposes. All group discussions were audio recorded for further analysis.

4.3 Giving meaning to multimodal data

Each data stream captured by the sensors and devices in our studies was encoded into columns in the multimodal matrix based on meaning elicited from subject matter experts, the learning design, or literature. The multimodal observations used in our studies, and their relationship with the dimensions of collaboration, are depicted in Figure 3, and described as follows:

Embodied strategies and proximity. Embodied strategies during high-stakes teamwork scenarios are critical in healthcare education [51]. Nurses are expected to be positioned in strategic areas when it comes to an emergency. Based on interviews with four nursing teachers, and related work [86], we identified five meaningful zones which are commonly associated with a range of actions nurses commonly perform (see Figure 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 (for i and v) and localisation (for ii, ii and iv) data captured during study sessions 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 'bagging' zone).

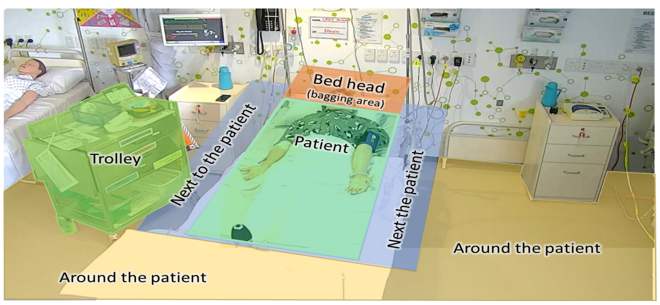

Figure 4: Zones of interest around the patient

Intensity of physical activity. From the literature [6] we know that nurses' physical activity varies from light (e.g. walking, talking, manipulating medical tools) to moderate (e.g. performing a CPR) intensity. We defined three levels of physical intensity for our study: *static*, *light* and *moderate,* where the highest intensity corresponds to performing CPR. Inspired by related work [51], to determine the level of intensity we first applied a moving average filter on the raw acceleration data streams from the wristbands to remove signal noise (window sample = 32). Then, through visual inspection of these data, we defined thresholds 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 in each cell.

Actions. Based on the learning design, we matched the set of expected actions and procedures (see sections 4.1 and 4.2.) expected to be performed during the simulation 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).

Speech activity and interaction. Verbal communication clearly plays an important role in the management and coordination of patient care. Nurses are encouraged to coordinate tasks, anticipate actions, and report information to the medical team in order to construct awareness [86]. Non-verbal metadata about speech, in the form of speech onset/offset, may shed light on the team performance [56, 77]. We attempted to capture hands-free speech activity via a microphone array in Study 1. Given the nature of the simulation, it was hard to obtain clean streams of voice that could be automatically analysed through speech detection algorithms due to noise generated by clinical instruments. For this reason, we used lapel microphones in Study 2. In the proxies to be presented in the next section, we reconstructed the dataset of Study 1 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 of speech (0) per second. We also added one column per nurse and for the patient to indicate who was listening (*RN(#)_listening, PT_listening)*, if they were in close proximity to the person speaking. 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.*

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 [57]. 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 [8]. When a change in the level of arousal is produced, physiological responses are activated in our body (e.g. increasing sweat production, heart rate, and blood pressure) [9]. We automatically identified peaks in EDA data using a tool called EDA Explorer [79]. A peak in skin conductance was defined by a minimum increase of 0.03 μs as suggested in the literature [5]. We added a column to the multimodal matrix called *EDA peaks.* Each cell contains a value of "1" when a peak was detected or "0" otherwise.

5 DESIGN AND ANALYSIS

Having established the multimodal matrix modelling methodology, we now present exemplar collaboration proxies that can be generated to help make collaboration translucent. Proxies 1-3 present information from Study 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 Study 2. For each proxy, we explain its design, the dimensions of collaboration that are made visible; the columns in the multimodal matrix used for the proxy; and evidence from observations and interviews with the nursing students.

5.1 Proxy 1: collocated social interaction

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.

Design. Figure 5 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 [2, 15, 22, 38], using a typical undirected network representation [25]. 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. *Analysis.* The social proxies of phase 1 (Figure 5, top) 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, 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, we can notice 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, we show two excerpts in teams A and B 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.

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)*

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, 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, we confirmed that the other nurse (RN3) was standing away from the bed, either waiting for instructions or just

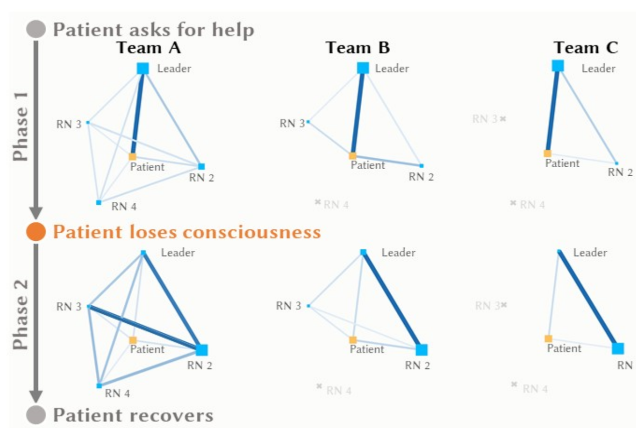

Figure 5: Social proxies in the physical space. The orange node represents the patient and blue nodes the nurses. Edges represent communication among nurses and with the patient.

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).

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

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 is for at least one nurse to remain in close proximity to the patient who is asking for help.

Design. Figure 6 shows the physical localisation and proximity for the three teams in Study 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

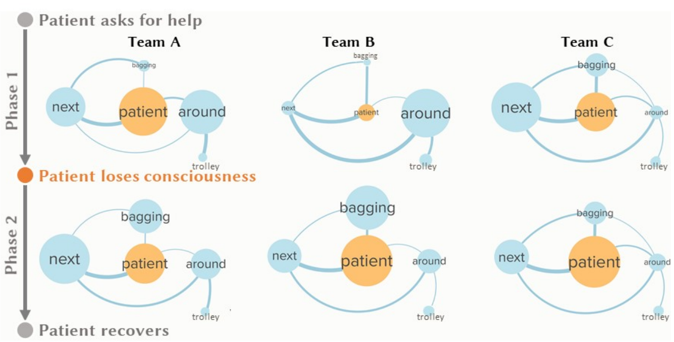

Figure 6: Physical localisation and proximity proxies. Each circle represents one zone of interest around the patient's bed, while the links are the transitions among zones.

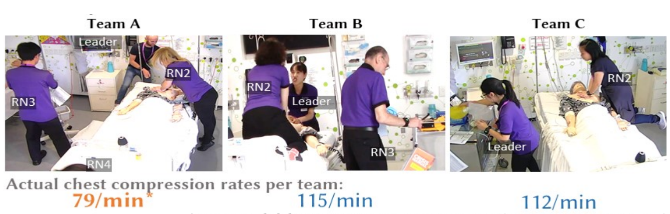

Figure 7: Three different ways in which nurses performed chest compressions: by the bed (team A), on top of the bed

transitions from one zone to another. This proxy can be generated per individual or for the whole team.

Analysis. The proxies of the teams during phase 1 (Figure 6, top) suggest that members of teams A and C were closer to the patient compared with team C. For team B, nurses were mostly *around* and further away from the patient, also showing more transitions between the *next* and *around* zones. Interestingly, we can triangulate evidence from the social proxy (described above) and suggest 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 6, 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 7, left), a suboptimal posture that may result in a poor CPR [64]. By contrast, teams B and C performed CPR on top of the patient or kneeling on the bed (Figure 7, centre and right), postures associated with better quality CPR. We triangulated this information 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 which one of the nurses should have been clearing the airway as other teams did (see Figure 7, left and centre). This issue was raised by the teacher during the debrief.

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

5.3 Proxy 3: electrodermal activity peaks

This proxy aims to help nurses reflect on *affective* traits that they may have experienced, based on detected peaks in their EDA [69], to consider potential coping strategies [57]. The proxy draws on both EDA peaks and levels of physical activity, since high physical activity may decrease the reliability of EDA modelling [67]. 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

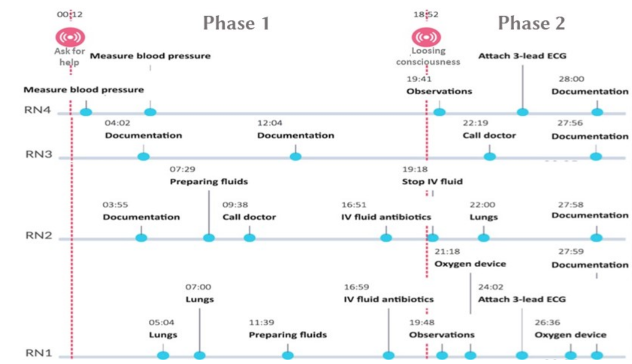

Figure 8: The team timeline as an epistemic proxy, depicting each student's actions in the simulation.

to intense activity, rather than to authentic sources of arousal.

Design. Figure 9 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.

Analysis. 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 9, 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, we qualitatively analysed the videos. A researcher 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 9, 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 9, 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 9, 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.

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 (Study 2) for supporting post-hoc reflection. This proxy utilises the *actions* columns from the multimodal matrix.

Design. We generated a timeline proxy that includes the actions, the time when the actions were enacted and who performed each. Figure 8 depicts an example of a timeline of team 5. 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.

Analysis. As detailed above, we analysed our findings grouped around the three themes from social translucence: *visibility, awareness* and *accountability* as per in [62].

Visibility. *In* all the six teams, students individually

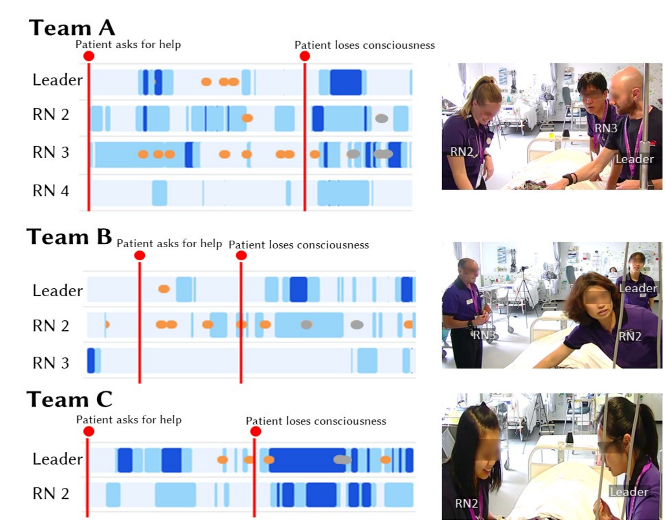

Figure 9: Timeline proxies of EDA indicating: EDA peaks (orange dots); b) physical intensity (represented by different shades of blue); and c) EDA peaks that may be affected by intense physical activity (grey dots).

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).

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 4 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 in 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 3 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). 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.

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.

6 DISCUSSION AND CONCLUSION

As Dewey [11] famously noted, *"We do not learn from experience... we learn from reflecting on experience."* (p.78). Once an experience can be examined from different perspectives, and reflected upon, it can be improved. One can do this as a personal activity in the mind's eye, although memory is imperfect and benefits from memory aids. Moreover, when that experience is cognitively and emotionally intense, and a function of complex teamwork, only some of which one had control over, there is arguably a critical role for external aids to help replay that experience in 'slow motion', ideally highlighting salient features that deserve closer attention. Elite teams have several coaches who call on advanced analytics to augment their capacity. For the majority, the dedicated coaching team will remain a dream, but collaboration analytics may have an increasingly important and interesting role to play.

This paper has extended the concept of *social translucence* for *online systems*, by operationalising the new construct of *collaboration translucence* for *collocated teamwork*, using multimodal traces from interactions between people and artefacts. The goal is to move such activity from being ephemeral, and largely opaque to computational analysis, to a translucent phenomenon from which selected features of interaction can be captured and rendered visible, for the purposes of learning.

This paper looks beyond the important multimodal data-fusion challenge of integrating information from disparate sensors and devices [42]. Assuming this can be accomplished, the *interpretive* challenge remains: how can someone make sense of all the data? Our approach is to ground the data *theoretically* (in this paper, along physical, epistemic, social and affective dimensions [30, 68]), and *contextually* (in these studies, using concepts from healthcare simulation). The mapping from low-level data to meaningful constructs is *modelled* using a multimodal matrix (Figure 3), from which collaboration proxies can be generated, providing views of what has taken place in a session. This has given us a way to operationalise a concept such as *patient-centeredness* in the way a team engages on the ward. Different visualisations (proxies 1-4), attending to different aspects of the activity, invite exploration and interpretation.

In sum, other researchers and designers can build on our approach by following the next steps: i) defining the higher order constructs that groups are aimed to develop or that can serve to identify 'good' groupwork or teamwork practice for the particular context; ii) 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; iii) after fusing and synchronising the multimodal data streams, each modality can be encoded into one or more meaningful *multimodal observations,* defining the columns of the matrix; iv) selecting the unit of analysis, for example, utterances, a time window or critical incidents, that will define the meaning of each row (*segments*) of the matrix; v) selecting how segments relate to each other (to group them into *stanzas*) based on the learning design or by expert knowledge to facilitate associating multimodal information with higher order constructs for a specific part of the activity; v) 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).

This work raises several issues for discussion.

Technical infrastructure. In our healthcare teamwork examples, all proxies can be automatically generated by applying rule-based algorithms on the multimodal data streams encoded into the multimodal matrix (see preliminary work in [19]). The only exception is proxy 1. For this proxy, while the transcript was manually created and analysed, we anticipate that this step will be automatable in the near future, with the use of personal microphones for personally attributable, and rapid improvements in speech-to-text services [54]. The infrastructure described is experimental, but an explicit part of the project is to examine strategies to embed it economically in the regular infrastructure.

Combining proxies. Aligning views of communication and proximity (proxies 1-2) might provide more complete information, clarifying the locational context in which nurses talked with each other, and their proximity to patients when talking to them.

Multimodal matrix. This modelling approach has helped us apply the insights from "quantitative ethnography" [74] to our challenge, enabling us to automate the coding of quantitative, low-level data into qualitative, higher order categories that are grounded in generic features of all teamwork, and the specifics of the particular activity. Additional columns could be added manually to the matrix, from conventional post hoc qualitative analysis that codes combinations of events.

Implications for nurse education. Our proxies address a gap identified in the literature on simulation for nursing education. Supports reflective professional practice [45] and facilitating evidence-based assessment [16] are two areas that will be explored by contextualising the proxies in relation to 'good' clinical practice.

Interpretation and assessment. This work has been undertaken in the context of healthcare simulation, but we are optimistic that this could generalise to other contexts such as professional development. The approach is fundamentally about improving a team's ability to reflect on its work. We were intrigued to see how students read meaning into the team timeline, perhaps raising the issue for some readers of whether they were 'correct'. We note first of all that in the original work on social translucence, there was never the intention that a proxy had one correct meaning, and we adhere to that principle in this work. In highly complex forms of human interaction such as this, the purpose of proxies is to provide a meaningful sense of what has been taking place, but it is not reasonable to expect a human – far less a machine – to know for sure what this signifies.

To conclude, there is more going on than is visible to the machines. Our goal is to improve the resources for interpretation where none existed before, in order to provoke more productive reflection and discussion, grounded for the first time in empirical evidence. In educational terms, therefore, collaboration proxies should serve as scaffolds for *formative* assessment (feedback for ongoing learning) rather than *summative* assessment (which assigns a grade). Conceivably, as the evidence base grows in size and quality, patterns in the data may be so highly correlated with strong or weak teamwork that they come to serve as proxies for quality. Until then, the priority is to further evaluate how students and educators engage with the representations, and how they can be improved.

REFERENCES

[1] S. S. J. Alhadad (2018). Visualizing Data to Support Judgement, Inference, and Decision Making in Learning Analytics: Insights from Cognitive Psychology and Visualization Science. Journal of Learning Analytics, 5, (2), 60-85. DOI: http://dx.doi.org/10.1007/s11528-017-0226-9.

[2] K. Bachour, F. Kaplan and P. Dillenbourg (2010). An Interactive Table for Supporting Participation Balance in Face-to-Face Collaborative Learning. IEEE Transactions on Learning Technologies, 3, (3), 203-213. DOI: https://doi.org/10.1109/TLT.2010.18.

[3] J. E. Bardram (2005). Activity-based computing: support for mobility and collaboration in ubiquitous computing. Pers Ubiquit Comput, 9, (5), 312-322. DOI: https://doi.org/10.1007/s00779-004-0335-2.

[4] M. Bilandzic and M. Foth (2012). A review of locative media, mobile and embodied spatial interaction. International Journal of Human-Computer Studies, 70, (1), 66-71. DOI: http://doi.org/10.1016/j.ijhcs.2011.08.004.

[5] J. J. Braithwaite, D. G. Watson, R. Jones and M. Rowe (2013). A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. Psychophysiology, 49, (1), 1017-1034.

[6] S. E. Chappel, S. J. J. M. Verswijveren, B. Aisbett, J. Considine and N. D. Ridgers (2017). Nurses' occupational physical activity levels: A systematic review. International Journal of Nursing Studies, 73, (August), 52-62. DOI: https://doi.org/10.1016/j.ijnurstu.2017.05.006.

[7] M. Cukurova, R. Luckin, E. Millán and M. Mavrikis (2018). The NISPI framework: Analysing collaborative problem-solving from students' physical interactions. Computers & Education, 116, 93-109. DOI: https://doi.org/10.1016/j.compedu.2017.08.007.

[8] C. W. Darrow (1964). The rationale for treating the change in galvanic skin response as a change in conductance. Psychophysiology, 1, (1), 31-38. DOI: https://doi.org/10.1111/j.1469-8986.1964.tb02618.x.

[9] M. E. Dawson, A. M. Schell and D. L. Filion. 2007. The Electrodermal System. In *Handbook of Psychophysiology*, Gary Berntson, John T. Cacioppo and Louis G. Tassinary (Eds.). Cambridge University Press, Cambridge, 159-181.

[10] Y. O. de Lima and J. M. de Souza. 2017. The future of work: Insights for CSCW. In *Proceedings of the IEEE International Conference on Computer Supported Cooperative Work in Design, CSCWD'17* 42-47. DOI: https://doi.org/10.1109/CSCWD.2017.8066668.

[11] J. Dewey. 1997. *How we think*. Courier Corporation, Chelmsford, USA.

[12] D. Di Mitri, J. Schneider, M. Specht and H. Drachsler (2018). From signals to knowledge: A conceptual model for multimodal learning analytics. Journal of Computer Assisted Learning, 34, (4), 338-349. DOI: https://doi.org/10.1111/jcal.12288.

[13] D. Di Mitri, J. Schneider, R. Klemke, M. Specht and H. Drachsler. 2019. Read Between the Lines: An Annotation Tool for Multimodal Data for Learning. In *Proceedings of the International Conference on Learning Analytics and Knowledge, LAK'19*, to appear.

[14] Y. Dich, J. Reilly and B. Schneider. 2018. Using Physiological Synchrony as an Indicator of Collaboration Quality, Task Performance and Learning. In *Proceedings of the International Conference on Artificial Intelligence in Education, AIED 2018*, 98-110. DOI: https://doi.org/10.1007/978-3-319-93843-1_8.

[15] J. M. DiMicco, K. J. Hollenbach, A. Pandolfo and W. Bender (2007). The impact of increased awareness while face-to-face. Human-Computer Interaction, 22, (1), 47-96. DOI: https://doi.org/10.1080/07370020701307781.

[16] R. G. dos Santos Almeida, B. M. Jorge, V. D. Souza-Junior, A. Mazzo, J. C. A. Martins, E. C. Negri and I. A. C. Mendes (2018). Trends in research on simulation in the teaching of nursing: an integrative review. Nursing education perspectives, 39, (3), E7-E10. DOI: http://dx.doi.org/10.1097/01.NEP.0000000000000292.

[17] S. Dumais, R. Jeffries, D. M. Russell, D. Tang and J. Teevan. 2014. Understanding User Behavior Through Log Data and Analysis. In *Ways of Knowing in HCI*, S. Judith Olson and A. Wendy Kellogg (Eds.). Springer, New York, 349-372. DOI: https://doi.org/10.1007/978-1-4939-0378-8_14.

[18] V. Echeverria, R. Martinez-Maldonado, K. Chiluiza and S. B. Shum. 2017. Dbcollab: Automated Feedback for Face-to-Face Group Database Design. In *Proceedings of the International Conference on Computers in Education, ICCE'17*, 156 - 165.

[19] V. Echeverria, R. Martinez-Maldonado, R. Granda, K. Chiluiza, C. Conati and S. B. Shum. 2018. Driving data storytelling from learning design. In *Proceedings of the International Conference on Learning Analytics and Knowledge, LAK'18*, 131-140. DOI: https://doi.org/10.1145/3170358.3170380.

[20] T. Erickson, D. Smith, W. Kellogg, M. Laff, J. Richards and E. Bradner. 1999. Socially translucent systems: social proxies, persistent conversation, and the design of "babble". In *Proceedings of the SIGCHI Conference on Human Factors in Computer Systems, CHI'99*, 72-79. DOI: https://doi.org/10.1145/302979.302997.

[21] T. Erickson and W. A. Kellogg (2000). Social translucence: an approach to designing systems that support social processes. ACM Transactions on Computer-Human Interaction, 7, (1), 59-83. DOI: https://doi.org/10.1145/344949.345004.

[22] T. Erickson, C. Halverson, W. A. Kellogg, M. Laff and T. Wolf (2002). Social translucence: designing social infrastructures that make collective activity visible. Communications of the ACM, 45, (4), 40-44. DOI: https://doi.org/10.1145/505248.505270.

[23] A. C. Evans, J. O. Wobbrock and K. Davis. 2016. Modeling Collaboration Patterns on an Interactive Tabletop in a Classroom Setting. In *Proceedings of the International Conference on Computer-Supported Cooperative Work, CSCW'16*, 860-871. DOI: https://doi.org/10.1145/2818048.2819972.

[24] P. Fournier-Viger, R. Nkambou and E. M. Nguifo. 2010. Building Intelligent Tutoring Systems for Ill-Defined Domains. In *Advances in Intelligent Tutoring Systems*, Roger Nkambou, Jacqueline Bourdeau and Riichiro Mizoguchi (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 81-101. DOI: https://doi.org/10.1007/978-3-642-14363-2_5.

[25] D. Gašević, S. Joksimović, B. R. Eagan and D. W. Shaffer (2018). SENS: Network analytics to combine social and cognitive perspectives of collaborative learning. Computers in Human Behavior, July, (2018). DOI: https://doi.org/10.1016/j.chb.2018.07.003.

[26] A. Gibson and R. Martinez-Maldonado. 2017. That dashboard looks nice, but what does it mean?: towards making meaning explicit in learning analytics design. In *Proceedings of the Australian Conference on Computer-Human Interaction, OZCHI'17*, 528-532. DOI: https://doi.org/10.1145/3152771.3156171.

[27] E. Gilbert. 2012. Designing social translucence over social networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'12*, 2731-2740. DOI: https://doi.org/10.1145/2207676.2208670.

[28] R. M. Gillies, A. F. Ashman and J. Terwel. 2007. *The teacher's role in implementing cooperative learning in the classroom*. Springer, New York. DOI: https://doi.org/10.1007/978-0-387-70892-8.

[29] P. A. Gloor, F. Grippa, J. Putzke, C. Lassenius, H. Fuehres, K. Fischbach and D. Schoder (2012). Measuring social capital in creative teams through sociometric sensors. International Journal of Organisational Design and Engineering, 2, (4), 380-401. DOI: http://dx.doi.org/10.1504/IJODE.2012.051442.

[30] P. Goodyear and L. Carvalho. 2014. Framing the analysis of learning network architectures. In *The architecture of productive learning networks*, Lucila Carvalho and Peter Goodyear (Eds.). Routledge, New York, 48-70.

[31] A. Green, S. P. Stawicki and M. S. Firstenberg. 2018. Introductory Chapter: Medical Error and Associated Harm-The The Critical Role of Team Communication and Coordination. In *Vignettes in Patient Safety-Volume 3*, Stanislaw P. Stawicki and Michael S. Firstenberg (Eds.). IntechOpen, London, UK, 1-12. DOI: https://doi.org/10.5772/intechopen.78014.

[32] E. Haataja, J. Malmberg and S. Järvelä (2018). Monitoring in collaborative learning: Co-occurrence of observed behavior and physiological synchrony

explored. Computers in Human Behavior, 87, 337-347. DOI: https://doi.org/10.1016/j.chb.2018.06.007.

[33] E. Hutchins (2010). Cognitive ecology. Topics in cognitive science, 2, (4), 705-715.

[34] H. Ishii. 2008. Tangible bits: beyond pixels. In *Proceedings of the International Conference on Tangible and Embedded Interaction, TEI'08*, xv-xxv. DOI: https://doi.org/10.1145/1347390.1347392.

[35] H.-C. Jetter, H. Reiterer and F. Geyer (2014). Blended Interaction: understanding natural human–computer interaction in post-WIMP interactive spaces. Pers Ubiquit Comput, 18, (5), 1139-1158. DOI: https://doi.org/10.1007/s00779-013-0725-4.

[36] D. M. Johnson, P. Sutton and J. Poon. 2000. Face-to-Face vs. CMC: Student communication in a technologically rich learning environment. In *Proceedings of the Annual Conference of the Australian Society for Computers in Learning in Tertiary Education, ASCILITE 2000*, 509-520.

[37] C. Kaendler, M. Wiedmann, N. Rummel and H. Spada (2015). Teacher Competencies for the Implementation of Collaborative Learning in the Classroom: a Framework and Research Review. Educ Psychol Rev, 27, (3), 505-536. DOI: https://doi.org/10.1007/s10648-014-9288-9.

[38] T. Kim, A. Chang, L. Holland and A. S. Pentland. 2008. Meeting mediator: enhancing group collaborationusing sociometric feedback. In *Proceedings of the International Conference on Computer Supported Cooperative Work, CSCW'08*, 457-466. DOI: https://doi.org/10.1145/1460563.1460636.

[39] K. Kitto, S. B. Shum and A. Gibson. 2018. Embracing imperfection in learning analytics. In *Proceedings of the International Conference on Learning Analytics and Knowledge, LAK'18*, 451-460. DOI: https://doi.org/10.1145/3170358.3170413.

[40] D. S. Kusunoki and A. Sarcevic. 2015. Designing for Temporal Awareness: The Role of Temporality in Time-Critical Medical Teamwork. In *Proceedings of the International Conference on Computer Supported Cooperative Work, CSCW'15*, 1465-1476. DOI: https://doi.org/10.1145/2675133.2675279.

[41] J. Lazar, J. H. Feng and H. Hochheiser. 2017. *Research methods in human-computer interaction*. Morgan Kaufmann.

[42] M. Liggins II, D. Hall and J. Llinas. 2017. *Handbook of multisensor data fusion: theory and practice*. CRC press, Boca Raton, USA.

[43] K. Littleton, D. Miell and D. Faulkner. 2004. *Learning to collaborate, collaborating to learn: Understanding and promoting educationally productive collaborative work*. NSP, Charlotte, USA.

[44] Z. Liu, Y. Wang, M. Dontcheva, M. Hoffman, S. Walker and A. Wilson (2017). Patterns and sequences: Interactive exploration of clickstreams to understand common visitor paths. IEEE Transactions on Visualization and Computer Graphics, 23, (1), 321-330.

[45] B. Mariani and J. Doolen (2016). Nursing simulation research: What are the perceived gaps? Clinical simulation in nursing, 12, (1), 30-36. DOI: https://doi.org/10.1016/j.ecns.2015.11.004.

[46] N. Marquardt, K. Hinckley and S. Greenberg. 2012. Cross-device interaction via micro-mobility and f-formations. In *Proceedings of the Symposium on User Interface Software and Technology, UIST'12* 13-22. DOI: https://doi.org/10.1145/2380116.2380121.

[47] R. Martinez-Maldonado, A. Clayphan, K. Yacef and J. Kay (2015). MTFeedback: Providing Notifications to Enhance Teacher Awareness of Small Group Work in the Classroom. IEEE Transactions on Learning Technologies, 8, (2), 187-200. DOI: https://doi.org/10.1109/TLT.2014.2365027.

[48] R. Martinez-Maldonado, K. Yacef and J. Kay (2015). TSCL: A conceptual model to inform understanding of collaborative learning processes at interactive tabletops. International Journal of Human-Computer Studies, 83, 62-82. DOI: https://doi.org/10.1016/j.ijhcs.2015.05.001.

[49] R. Martinez-Maldonado, P. Goodyear, J. Kay, K. Thompson and L. Carvalho. 2016. An Actionable Approach to Understand Group Experience in Complex, Multi-surface Spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'16*, 2062-2074. DOI: https://doi.org/10.1145/2858036.2858213.

[50] R. Martinez-Maldonado, B. Schneider, S. Charleer, S. B. Shum, J. Klerkx and E. Duval. 2016. Interactive surfaces and learning analytics: data, orchestration aspects, pedagogical uses and challenges. In *Proceedings of the International Conference on Learning Analytics & Knowledge, LAK'16*, 124-133. DOI: https://doi.org/10.1145/2883851.2883873.

[51] R. Martinez-Maldonado, V. Echeverria, O. C. Santos, A. D. P. Dos Santos and K. Yacef. 2018. Physical learning analytics: A multimodal perspective. In *Proceedings of the International Conference on Learning Analytics and Knowledge, LAK'18*, 375-379. DOI: https://doi.org/10.1145/3170358.3170379.

[52] F. Mazzocchi (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.

[53] D. W. McDonald, S. Gokhman and M. Zachry. 2012. Building for social translucence: a domain analysis and prototype system. In *Proceedings of the International Conference on Computer Supported Cooperative Work, CSCW'12*, 637-646. DOI: https://doi.org/10.1145/2145204.2145301.

[54] M. McGregor and J. C. Tang. 2017. More to Meetings: Challenges in Using Speech-Based Technology to Support Meetings. In *Proceedings of the International Conference on Computer Supported Cooperative Work, CSCW'17*, 2208-2220. DOI: https://doi.org/10.1145/2998181.2998335.

[55] S. K. Milligan and P. Griffin (2016). Understanding learning and learning design in MOOCs: a measurement-based interpretation. Journal of Learning Analytics, 3, (2), 88-115. DOI: https://doi.org/10.18608/jla.2016.32.5

[56] A. Mohan, H. C. Sun, O. Lederman, K. Full and A. S. Pentland. 2018. Measurement and Feedback of Group Activity Using Wearables for Face-to-Face Collaborative Learning. In *Proceedings of the IEEE International Conference on Advanced Learning Technologies, ICALT'18*, 163-167. DOI: https://doi.org/10.1109/ICALT.2018.00046.

[57] L. Müller, V. Rivera-Pelayo, C. Kunzmann and A. Schmidt. 2011. From Stress Awareness to Coping Strategies of Medical Staff: Supporting Reflection on Physiological Data. In *Proceedings of the International Workshop on Human Behavior Understanding*, 93-103. DOI: https://doi.org/10.1007/978-3-642-25446-8_11.

[58] P. Müller, M. X. Huang and A. Bulling. 2018. Detecting Low Rapport During Natural Interactions in Small Groups from Non-Verbal Behaviour. In *Proceedings of the International Conference on Intelligent User Interfaces, IUI'18*, 153-164. DOI: https://doi.org/10.1145/3172944.3172969.

[59] Y. I. Nakano, S. Nihonyanagi, Y. Takase, Y. Hayashi and S. Okada. 2015. Predicting participation styles using co-occurrence patterns of nonverbal behaviors in collaborative learning. In *Proceedings of the International Conference on Multimodal Interaction, ICMI'15*, 91-98. DOI: https://doi.org/10.1145/2818346.2820764.

[60] B. A. Nardi and S. Whittaker. 2002. The place of face-to-face communication in distributed work. In *Distributed work: New research on working across distance using technology* Pamela J. Hinds and Sara Kiesler (Eds.). MIT Press, Cambridge, 83-110.

[61] C. Neubauer, J. Woolley, P. Khooshabeh and S. Scherer. 2016. Getting to know you: a multimodal investigation of team behavior and resilience to stress. In *Proceedings of the International Conference on Multimodal Interaction, ICMI'16*, 193-200. DOI: https://doi.org/10.1145/2993148.2993195.

[62] K. Niemantsverdriet, M. Broekhuijsen, H. van Essen and B. Eggen. 2016. Designing for multi-user interaction in the home environment: implementing social translucence. In *Proceedings of the International Conference on Designing Interactive Systems, DIS'16*, 1303-1314. DOI: https://doi.org/10.1145/2901790.2901808.

[63] F. Nihei, Y. Nakano and Y. Takase. 2017. Predicting meeting extracts in group discussions using multimodal convolutional neural networks. In *Proceedings of the International Conference on Multimodal Interaction, ICMI'17*, 421-425. DOI: https://doi.org/10.1145/3136755.3136803.

[64] J. Oh, Y. Chee, T. Lim, Y. Cho and I. Y. Kim (2014). Chest compression with kneeling posture in hospital cardiopulmonary resuscitation: a randomised crossover simulation study. Emergency Medicine Australasia, 26, (6), 585-590. DOI: https://doi.org/10.1111/1742-6723.12307.

[65] J. S. Olson, S. Teasley, L. Covi and G. Olson. 2002. The (currently) unique advantages of collocated work. In *Distributed work: New research on working across distance using technology* Pamela J. Hinds and Sara Kiesler (Eds.). MIT Press, Cambridge, 113-136.

[66] S. Oviatt and A. Cohen. 2013. Written and multimodal representations as predictors of expertise and problem-solving success in mathematics. In *Proceedings of the International Conference on Multimodal Interaction, ICMI'13*, 599-606. DOI: https://doi.org/10.1145/2522848.2533793.

[67] L. A. Pugh, C. R. Oldroyd, T. S. Ray and M. L. Clark (1966). Muscular effort and electrodermal responses. Journal of Experimental Psychology, 71, (2), 241-248. DOI: https://doi.org/10.1037/h0022834.

[68] W.-M. Roth, J. Tenenberg and D. Socha (2016). Discourse/s in/of CSCW. Computer Supported Cooperative Work, 25, (4-5), 385-407. DOI: https://doi.org/10.1007/s10606-016-9258-5.

[69] N. Ruiz-Robledillo and L. Moya-Albiol (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. DOI: https://doi.org/10.1007/s10803-013-1996-3.

[70] N. Rummel, H. Spada and S. Hauser (2009). Learning to collaborate while being scripted or by observing a model. International Journal of Computer-Supported Collaborative Learning, 4, (1), 69-92. DOI: https://doi.org/10.1007/s11412-008-9054-4.

[71] K. Schmidt (2002). The problem with awareness: Introductory remarks on awareness in CSCW. Computer Supported Cooperative Work, 11, (3-4), 285-298. DOI: https://doi.org/10.1023/A:1021272909573.

[72] B. Schneider, K. Sharma, S. Cuendet, G. Zufferey, P. Dillenbourg and R. Pea. 2016. Detecting Collaborative Dynamics Using Mobile Eye-Trackers. In *Proceedings of the International Conference of the Learning Sciences, ICLS'16*, 522-529.

[73] B. Schneider and R. Pea. 2017. Real-time mutual gaze perception enhances collaborative learning and collaboration quality. In *Educational Media and Technology Yearbook*, Springer, 99-125.

[74] D. W. Shaffer. 2017. *Quantitative ethnography*. Cathcart Press.

[75] V. J. Shute and M. Ventura. 2013. *Stealth assessment: Measuring and supporting learning in video games.* MIT Press.
[76] D. Spikol, E. Ruffaldi, L. Landolfi and M. Cukurova. 2017. Estimation of Success in Collaborative Learning Based on Multimodal Learning Analytics Features. In *Proceedings of the IEEE International Conference on Advanced Learning Technologies, ICALT'17*, 269-273. DOI: https://doi.org/10.1109/ICALT.2017.122.
[77] J. Sturm, O. H. V. Herwijnen, A. Eyck and J. Terken. 2007. Influencing social dynamics in meetings through a peripheral display. In *Proceedings of the International Conference on Multimodal Interfaces, ICMI'07*, 263-270. DOI: https://doi.org/10.1145/1322192.1322238.
[78] S. Tausch, S. Ta and H. Hussmann. 2016. A Comparison of Cooperative and Competitive Visualizations for Co-located Collaboration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'16*. DOI: https://doi.org/10.1145/2858036.2858072.
[79] S. Taylor, N. Jaques, W. Chen, S. Fedor, A. Sano and R. Picard. 2015. Automatic identification of artifacts in electrodermal activity data. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'15*, 1934-1937. DOI: https://doi.org/10.1109/EMBC.2015.7318762.
[80] S. D. Teasley (2017). Student Facing Dashboards: One Size Fits All? Technology, Knowledge and Learning, 22, (3), 377-384. DOI: https://doi.org/10.1007/s10758-017-9314-3.
[81] S. A. Viswanathan and K. VanLehn (2017). Using the tablet gestures and speech of pairs of students to classify their collaboration. IEEE Transactions on Learning Technologies, 11, (2). DOI: https://doi.org/10.1109/TLT.2017.2704099.
[82] L. S. Vygotski. 1987. *The collected works of LS Vygotsky. Volume 1: Problems of general psychology, including the volume Thinking and Speech.* Springer New York.
[83] G. Wang, X. Zhang, S. Tang, H. Zheng and B. Y. Zhao. 2016. Unsupervised Clickstream Clustering For User Behavior Analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'16*, 225-236. DOI: https://doi.org/10.1145/2858036.2858107.
[84] A. F. Wise and D. W. Shaffer (2015). Why theory matters more than ever in the age of big data. Journal of Learning Analytics, 2, (2), 5-13. DOI: https://doi.org/10.18608/jla.2015.22.2
[85] M. Worsley and P. Blikstein (2018). A multimodal analysis of making. International Journal of Artificial Intelligence in Education, 28, (3), 385-419. DOI: https://doi.org/10.1007/s40593-017-0160-1.
[86] Z. Zhang and A. Sarcevic. 2015. Constructing Awareness Through Speech, Gesture, Gaze and Movement During a Time-Critical Medical Task. In *Proceedings of the European Conference on Computer Supported Cooperative Work, ECSCW'15*, 163-182. DOI: https://doi.org/10.1007/978-3-319-20499-4_9.
[87] A. Zolyomi, A. S. Ross, A. Bhattacharya, L. Milne and S. A. Munson. 2018. Values, Identity, and Social Translucence: Neurodiverse Student Teams in Higher Education. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'18*, 1-13. DOI: https://doi.org/10.1145/3173574.3174073.