The Multimodal Matrix as a Quantitative Ethnography Methodology

Chapter · October 2019

DOI: 10.1007/978-3-030-33232-7_3

3 authors:

Simon Buckingham Shum
University of Technology Sydney

Vanessa Echeverria
Monash University (Australia)

Roberto Martinez-Maldonado
Monash University (Australia)

Making Sense of Collocated Team Activity: The Multimodal Matrix as a Quantified Ethnography Methodology

Simon Buckingham Shum[1], Vanessa Echeverria[1,2] and Roberto Martinez-Maldonado[1]

[1] *Connected Intelligence Centre, University of Technology Sydney, PO Box 123 Broadway, Sydney, Australia*
[2] *Escuela Superior Politécnica del Litoral, ESPOL, Campus Gustavo Galindo Km 30.5 Vía Perimetral, Guayaquil, Ecuador*

Abstract: This paper seeks to contribute to the emerging field of Quantitative Ethnography (QE), by demonstrating its utility to solve a complex challenge now facing educational data scientists: how can we build real time Learning Analytics delivering feedback to collocated teams and their coaches, combining multimodal data streams with human observational data? We define some key requirements that extend current conceptions of QE in order to operationalise it for the Learning Analytics design process. We then introduce, and demonstrate, the *Multimodal Matrix* in the context of high-performance nursing teamwork. The Multimodal Matrix integrates theoretical concepts about teamwork with contextual insights about specific work practices, and enables the analyst to map these codes to low level sensor data, or indeed, manually performed qualitative analyses. In order to close the feedback loop to students and educators in a timely way, this is implemented in software as a workflow for rapid data modelling, analysis and visualization. We propose that this exemplifies how a QE methodology can underpin collocated activity analytics, at scale, with in principle applications to embodied, collocated activities beyond our case study.

Introduction and Background

Quantitative Ethnography (QE) is a methodological approach that respects the insights into specific cultural practices gained from the interpretive disciplines developed in ethnographic and other qualitative traditions, but seeks to apply the power of statistical and other data science techniques to qualitatively coded data, such as observational fieldnotes, interviews, or video analysis [17]. To date QE has been exemplified primarily by the development of the Epistemic Network Analysis (ENA) tool, to examine the *relationships* between coded data elements, e.g. [3, 9, 16]; In this paper, we propose that a modelling methodology and analytics workflow, developed to process and visualize multimodal data from nursing team simulations, also exemplifies QE principles. The intended contribution is thus twofold: (i) an articulation of key QE principles as we understand them, which we extend in two key respects to operationalise them for our learning analytics application, and (ii) a demonstration of QE's value in implementing a multimodal learning analytics pipeline and end-user tool, which has been piloted with nursing academics and students. To our knowledge, this is the first time that QE has been applied to the analysis of multimodal sensor data, combined with human observational data, for face-to-face teamwork, rather than online mediated interaction.

In the next section we introduce the educational challenge, namely, how to better inform the debriefings for coaching high performance, collocated nursing teams. We then introduce the *Multimodal Matrix* as a modelling methodology, inspired by Quantitative Ethnography principles in combination with a teamwork activity theory. We illustrate the application to nursing simulation data, and briefly describe the visualisations it enables, designed to meet the needs of students and staff.

Nursing Teamwork Simulations

Nursing 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 UTS Bachelor of Nursing experience many hypothetical scenarios across different stages of their professional development. In these scenarios students, acting as Registered Nurses (RNs), 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 [10].

The manikin ("Mr. Lars") was programmed by the teacher to deteriorate over time, dividing the task into two phases. In phase one a group of four students assess and treat Mr Lars for chest pain. These RNs in different roles communicate with Mr. Lars, apply oxygen, assess his pain, perform vital sign observations, administer Anginine according to the six rights, connect him to an ECG, identify his cardiac rhythm, document appropriately and call for a clinical review. In phase two, the same group of students takes over Mr. Lars's care at which point he loses consciousness due to a fatal cardiac rhythm, and the team must perform basic life support. Each simulation lasted an average of 9.5 minutes. Fuller procedural details are provided elsewhere (Echeverria et al, 2019; Echeverria, In Prep).

Instrumenting simulations to detect teamwork

Several sensors and equipment were utilised to track interactions, summarised below:

Indoor localisation: Students' movement around the manikin was captured automatically through ultra-wideband (UWB) wearable badges (Pozyx.io[1]). This system is composed of a set of anchors to sense the physical space, which are mounted on the walls, and several wearable tags or badges attached to people or objects (such as the resus trolley). Figure 1 illustrates the distribution of the anchors across the simulation room (blue squares).

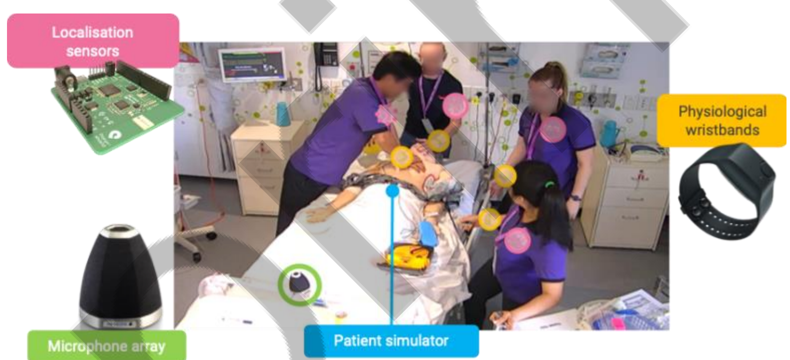

Figure 1: Data collection from a nursing simulation in a lab scenario, using a range of sensors.

Patient Simulator: Some student and patient actions were automatically logged by the high-fidelity Laerdal SimMan 3G[2] manikin 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. Proprietary Laerdal Software exported the actions and their timestamp in a .txt file.

Microphone array: A six-channel high-quality USB microphone array (Microcone) was located at the base of the patient's bed to detect nurses' conversations. Microcone Recorder application for MacOS was used to automatically track multiple people speech. Six .wav files were saved at the end of each session, one per channel. In addition, the application generated a .csv file including the total duration of the session, start and end timestamps where speech was detected and the person who was speaking (previously configured in the application).

Physiological wristbands: Empatica E4[3] wristbands included a photo plethysmography (PPG) sensor to measure Heart Rate 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 physical activity.

[1] Pozyx developer kit and a multitag-positioning system: https://www.pozyx.io
[2] Laerdal simulation manikins: https://www.laerdal.com/nz/products/simulation-training/emergency-care-trauma/simman-3g
[3] Empatica wristbands: https://www.empatica.com/en-int/research/e4

Each wristband exported 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.

In addition to these sensors and equipment, all the sessions were recorded by the video camera system installed in the lab room, comprising three fixed cameras and several microphones in the ceiling.

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 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 derive multimodal observations for further analysis.

Defining our requirements for a QE modelling methodology

In our reading of Shaffer (2017), central to QE's 'DNA' is the goal 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 seems to us to be important, and distinctive, clearly moving beyond mixed methods, pivotal to enabling ethnography, and the social sciences more broadly, to move into 'big data' and real time analytics.

To this summary we make two additions that reflect our design context, which gives us a working definition of our desired QE analysis process and infrastructure:

> Applying Quantitative Ethnography to design real time 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**.

Firstly, we emphasise *co-design* because we are developing feedback tools for use by real users (educators and students), so simply from sound user-centred design principles, this is good practice: we need to understand what feedback will be of most value, and elsewhere we detail the use of co-design tools to elicit such information from non-technical stakeholders [4, 15]. Moreover, human-centred design goes deeper than good user interfaces: it shapes the data we gather, and how we model it. The analytics challenge requires us to devise a way to model the sensor data in ways that respect, and will enable, culturally meaningful interpretations of work practices when visualized. Co-design provides us with a way to understand work practices in great specificity: the experiences of UTS students and academics in the specific simulations run in the Health faculty's facilities.

Secondly, we emphasise *automated analytics workflow* because this is the only way to make sense of large data sets in order to serve our purposes, namely, to close the feedback loop to educators and students in a timely manner. We need to deploy data science and information visualization for immediate input to post-simulation debriefings.

The Multimodal Matrix as a QE modelling methodology

In order to address the above requirements, we have developed a modelling approach and data representation named the *Multimodal Matrix* (Figure 2), comprising the following conceptual elements: *dimensions of collaboration, multimodal observations, segments,* and *stanzas*.

This matrix provides the common, integrating representation to hold data and analysis results from qualitative and quantitative methods: it can be populated with categorical data *automatically* from a full sensor → analytics pipeline, *semi-automatically* in which human input augments the analysis and/or workflow, or manually from conventional qualitative or quantitative data analysis. Qualitative codes are modelled by combining events from multiple sources (columns) into *segments*, and by combining multiple segments. Temporally dependent codes can be modelled into meaningful *stanzas* by combining *segments* (rows of events).

The matrix provides a representation to make sense of low-level sensor data through the introduction of qualitative coding derived from top-down (theory) and bottom-up (context-specific phenomena) sources:

- *Theory:* a framework for collaborative activity (ACAD) was used to define key constructs for combining lower level events into higher order codes (see below). Obviously, this could be replaced by any other theory/framework that served the analyst's interests and stakeholders' needs.
- *Insights into work practices:* multiple sources: (i) insights from nursing professionals about what makes a nurse's position meaningful when performing different tasks (informing the definition of spatial zones: Fig.3); (ii) what information they would like to see captured to inform post-simulation debriefing (informing which sensors are deployed); and (iii) information that staff and students said would assist post-simulation debriefing (from co-designing visualization prototypes).

Application of the Multimodal Matrix to nursing team simulations

Each data stream captured by the sensors and devices listed above was encoded into columns in the multimodal matrix based on meaning elicited from subject matter experts, the learning design, or literature [AIED, CHI]. The data streams were manually synchronised at a 1 Hz, down sampling data streams from sensors that had a higher frequency. The multimodal observations used in our studies, and their relationship with the dimensions of collaboration, are depicted in Table 1, and described below.

Theoretically informed dimensions of collaboration

Our methodology enables us to introduce theoretical perspectives. We draw on the Activity-Centred

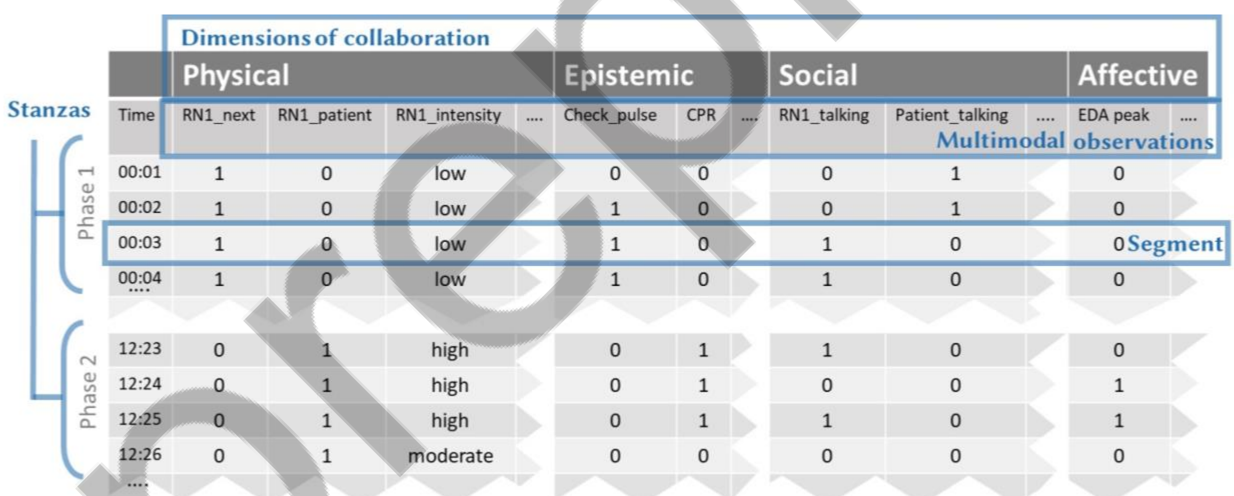

Figure 2: Schematic design of the Multimodal Matrix

Analysis & Design (ACAD) framework [13], which defines *physical, epistemic and social* dimensions as critical. To this we add *affective* states of engagement, worry or anticipation, this being particularly important in the healthcare professions. As can be seen from the edited excerpt from a teamwork nursing simulation in Table 1, these are reflected in the four principal column headings: every row has events classified under these.

Multimodal Observations

From the data collected, as explained in the previous sections, we were now in a position to associate multimodal observations, optionally in combination, with one or more dimensions of collaboration. Space precludes a very detailed description, which can be found elsewhere [4, 7]. Our goal in this paper is to convey the way in which the coding of data works.

Segments: Segments are considered the smallest unit of meaning. Thus, for this particular example in teamwork nursing simulations, we took a segment of one second. This small value was selected because we needed to analyse moment-to-moment critical reactions from nurses during the performance of the activity, this being a high-stakes activity.

Stanzas: Segments can be grouped according to criteria to show meaningful relationships. In the simulations, stanzas were defined to capture key phases in the collaborative task (e.g. see rows grouped by phase in Figure 1). For this particular example, two stanzas were defined, based on two critical actions in the learning design: *i) when the patient asks for help* and *ii) when the patient loses consciousness*.

Table 1 illustrates the Multimodal Matrix with some data, with commentary below.

Table 1. Edited excerpt from a teamwork nursing simulation encoded in the Multimodal Matrix

	Physical						**Epistemic**			**Social**						**Affective**	
time	**RN1.patient_bed**	**RN1.next_to_patient**	**RN1.around_patient**	**RN1.bed_head**	**RN1.trolley_area**	**RN1.pysical_intensity**	**RN1.check_pulse**	**RN2.check_pulse**	**RN1.compressions**	**RN1.speaking**	**RN2.speaking**	**patient.speaking**	**RN1.listening**	**RN2.listening**	**patient.listening**	**RN1.EDA_peak**	**RN2.EDA_peak**
03:22.0	0	1	0	0	0	L	0	0	0	0	0	1	1	1	0	0	0
03:22.1	0	1	0	0	0	L	0	0	0	0	0	1	1	1	0	1	0
03:22.2	0	1	0	0	0	L	0	0	0	0	0	1	1	1	0	0	0
03:22.3	0	1	0	0	0	L	0	0	0	0	0	1	1	1	0	0	0
03:22.4	0	1	0	0	0	L	0	0	0	0	0	1	1	1	0	0	0
03:22.5	0	1	0	0	0	L	0	1	0	0	0	0	0	0	0	0	0

ACAD Physical dimension. Embodied strategies during high-stakes teamwork scenarios are critical in healthcare education [12]. This is an example of how qualitative insights into the work practice shape the quantitative modelling. We must first understand what makes position *meaningful* for *these stakeholders* in *this simulation* and its intended *purpose* — because in other simulations, position might take on other significances, or with more advanced students (for instance) there might be other learning outcomes, which will focus on other key behaviours. Based on interviews with four nursing teachers [6], we identified five meaningful zones which are associated with a range of actions nurses must perform: i) *the patient's bed*, for cases in which nurses were

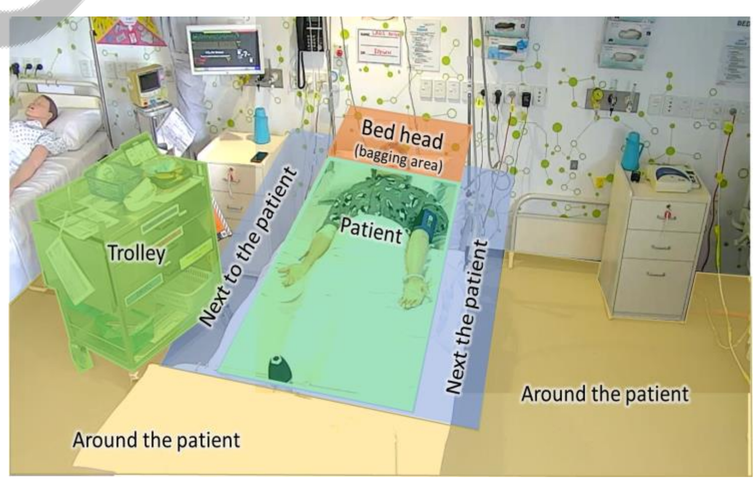

Figure 3: Nurses' positions were classified into zones, reflecting insights from subject matter experts regarding what makes position significant in teamwork.

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 during CPR; and v) *trolley area*, for cases in which nurses were getting medication or equipment (a localisation badge was attached to the trolley). Indoor localisation data was automatically encoded into these meaningful zones. A Kalman filter was applied to remove noisy data points, and a cluster analysis was performed (k=16) to assign one meaningful zone to each point. The first five columns in the Physical dimension group from Table 1 illustrate the meaningful zones for RN1 (e.g. RN1.patient_bed). Each cell has a value of "1", if that zone is occupied by a RN, or "0" otherwise. For instance, the row in the first second [0,1,0,0,0] means that RN1 was next to the patient. In addition to movement, nurses' *physical intensity* is studied in the literature [2], ranging from low (e.g. walking, talking, manipulating medical tools) to high (e.g. performing a CPR). We defined *low (L)*, *medium (M)* and *high (H)* levels, where high = performing CPR. The last Physical column shows that RN1's physical intensity at 03:22.0 was low (L).

ACAD Epistemic dimension. In the matrix, each column represents who performed an action (e.g. RN1.check_pulse). For example, the first column in the Epistemic dimension group is RN1.check_pulse = 0, meaning that RN1 did not check the pulse at time 03:22.0, while RN2.check_pulse = 1 at 03:22.5.

ACAD Social dimension. Verbal communication plays an important role in the management and coordination of patient care and teamwork strategies [19]. From the video recordings, we manually transcribed and synchronised the speech for each nurse and the patient using NVivo software. Start and ending points were annotated, along with the speaker and listener identification to further model the interactions. With this information, we created a sparse matrix, with 1's when a nurse was speaking/listening at a specific time, or 0 otherwise. The first column in the Social dimension group shows RN1.speaking, RN2.speaking and patient.speaking, and the following three columns are listening interactions. We can observe how the patient-nurse speaking-listening interaction is represented for the first four seconds: row [0,0,1]; [1,1,0] means that the patient is speaking while RN1 and RN2 are listening.

Affective dimension. Physiological data can be effectively used to aid nurses in recalling confronting experiences in order to develop coping strategies [14]. 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. We automatically identified peaks in EDA data as a minimum increase of 0.03 μs [1], using EDA Explorer [18]. Each cell contains a value of "1" when a peak in a specific time was detected or "0" otherwise. For example, RN1.EDA_peak column shows that RN1 had an EDA peak at 03:22.0.

Conclusions

In this paper, we have proposed a Quantitative Ethnographic methodology and analytics platform, to tackle the problem of making multimodal data streams meaningful for interpreting collocated teamwork. We accomplish this using a data representation called the Multimodal Matrix, which organises quantitative data in relation to codes derived from two sources: qualitative insights from stakeholders into their work practices (i.e. undertaking and coaching nursing simulations), plus theoretical insights into how collaborative teamwork can be analysed (ACAD framework). Rows in the matrix may be populated automatically or semi-automatically, and with results from manual data analyses. This modelling approach thus enables us to build a principled bridge between multimodal logs generated from sensors, and higher order constructs such as a curriculum outcome (e.g. *'competence in patient-centred teamwork'*), and its constituent skills (e.g. *'patient-centred talk'; 'correct positioning during CPR'*). When those signals are combined meaningfully, they may serve as proxies for these competencies, and once visualized, may provoke deeper reflection and discussion in debriefing sessions.

To deliver real-time, scalable feedback to nursing teams, we have highlighted the need not only for an integrative representation, but also a fully automated workflow. The analytics workflow is now sufficiently automated for a trained person to generate visualizations for student debriefs within a few minutes of completing a simulation, and the Multimodal Matrix has provided the database to generate visualizations to meet stakeholders' requests for debriefing support, which are now being evaluated with staff and students with promising results [5, 8]. Full automation must tackle challenges such as identifying who is performing an action and connecting siloed data sources from different products (e.g. manikin and wristband data) before the scripts can integrate and visualize them. While [11] have described the automation of ENA from online interaction data to provide a real-time dashboard for teachers, we believe that this is the first application of a quantitative ethnographic approach for *collocated* collaborative learning, working from multimodal data streams rather than click streams. On this basis, we conclude that the principles underpinning Quantitative Ethnography have assisted us in making very significant advances in tackling an extremely complex challenge: delivering timely analyses and feedback on embodied groupwork.

References

1. Braithwaite JJ, Watson DG, Jones R, Rowe M (2013) A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments Psychophysiology 49:1017-1034
2. Chappel SE, Verswijveren SJJM, Aisbett B, Considine J, Ridgers ND (2017) Nurses' occupational physical activity levels: A systematic review International Journal of Nursing Studies 73:52-62
3. Csanadi A, Eagan B, Kollar I, Shaffer DW, 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 13:419-438
4. Echeverria V (In prep.) Designing and Validating Automated Feed-back for Collocated Teams Using Multimodal Learning Analytics. University of Technology Sydney (UTS)
5. Echeverria V (In prep.) Designing and Validating Automated Feedback for Collocated Teams Using Multimodal Learning Analytics. University of Technology Sydney (UTS)
6. Echeverria V, Martinez-Maldonado R, Power T, Hayes C, Shum SB (2018) Where Is the Nurse? Towards Automatically Visualising Meaningful Team Movement in Healthcare Education. In: International Conference on Artificial Intelligence in Education, 2018. Springer, Cham, pp 74-78
7. Echeverria V, Martinez-Maldonado R, Shum SB (2019) Towards Collaboration Translucence: Giving Meaning to Multimodal Group Data. In: 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), Glasgow, UK, 2019. ACM, New York, NY, USA, p 16
8. Echeverria V, Martinez-Maldonado R, Shum SB (2019) Towards Collaboration Translucence: Giving Meaning to Multimodal Group Data. In: 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), Glasgow, UK, 2019. ACM, New York, NY, USA, p 16
9. Gašević D, Joksimović S, Eagan BR, Shaffer DW (2018) SENS: Network analytics to combine social and cognitive perspectives of collaborative learning Computers in Human Behavior July
10. Green A, Stawicki SP, Firstenberg MS (2018) Introductory Chapter: Medical Error and Associated Harm-The The Critical Role of Team Communication and Coordination. In: Vignettes in Patient Safety-Volume 3. IntechOpen,
11. Herder T, Swiecki Z, Fougt SS, Tamborg AL, Allsopp BB, Shaffer DW, Misfeldt M (2018) Supporting teachers' intervention in students' virtual collaboration using a network based model. In: 8th International Conference on Learning Analytics and Knowledge, Sydney, Australia, 2018. ACM, pp 21-25
12. Martinez-Maldonado R, Echeverria V, Santos OC, Dos Santos ADP, Yacef K (2018) Physical learning analytics: A multimodal perspective. Paper presented at the International Conference on Learning Analytics and Knowledge, LAK'18,
13. 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 Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, Santa Clara, California, USA,
14. 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, 2011//
15. Prieto-Alvarez CG, Martinez-Maldonado R, Anderson TD (2018) Co-designing learning analytics tools with learners Learning Analytics in the Classroom: Translating Learning Analytics for Teachers
16. Shaffer DW (2006) Epistemic frames for epistemic games Computers & Education 46:223-234
17. Shaffer DW (2017) Quantitative ethnography. Cathcart Press,
18. 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 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'15, 25-29 Aug. 2015

19. Zhang Z, Sarcevic A (2015) Constructing Awareness Through Speech, Gesture, Gaze and Movement During a Time-Critical Medical Task. Paper presented at the European Conference on Computer Supported Cooperative Work, ECSCW'15, Oslo, Norway, 19-23 September 2015,