METS: Multimodal Learning Analytics of Embodied Teamwork Learning

Conference Paper · January 2023

DOI: 10.1145/3576050.3576076

11 authors, including:

Linxuan Zhao
Monash University (Australia)

Zachari Swiecki
University of Wisconsin–Madison

Dragan Gasevic
The University of Edinburgh

Lixiang Yan
Moansh University

METS: Multimodal Learning Analytics of Embodied Teamwork Learning

LINXUAN ZHAO, Monash University, Australia
ZACHARI SWIECKI, Monash University, Australia
DRAGAN GAŠEVIĆ, Monash University, Australia
LIXIANG YAN, Monash University, Australia
SAMANTHA DIX, Monash University, Australia
HOLLIE JAGGARD, Monash University, Australia
ROSIE WOTHERSPOON, Monash University, Australia
ABRA OSBORNE, Monash University, Australia
XINYU LI, Monash University, Australia
RIORDAN ALFREDO, Monash University, Australia
ROBERTO MARTINEZ-MALDONADO, Monash University, Australia

Fig. 1. Embodied teamwork in an immersive healthcare simulation where a team of students constantly reconfigure themselves into sub-groups (e.g., see simultaneous, coded dialogue unfolding at the left and right of the learning space) to complete a joint task.

Embodied team learning is a form of group learning that occurs in co-located settings where students need to interact with others while actively using resources in the physical learning space to achieve a common goal. In such situations, communication dynamics can be complex as team discourse segments can happen in parallel at different locations of the physical space with varied team member configurations. This can make it hard for teachers to assess the effectiveness of teamwork and for students to reflect on their own experiences. To address this problem, we propose METS (Multimodal Embodied Teamwork Signature), a method to model team dialogue content in combination with spatial and temporal data to generate a signature of embodied teamwork. We present a study in the context of a highly dynamic healthcare team simulation space where students can freely move. We illustrate how signatures of embodied teamwork can help to identify key differences between high and low performing teams: i) across the whole learning session; ii) at different phases of learning sessions; and iii) at particular spaces of interest in the learning space.

CCS Concepts: • **Applied computing → Collaborative learning**; *Computer-assisted instruction.*

Additional Key Words and Phrases: Healthcare simulation, Collaborative learning, Communication, Teamwork, Multimodality

Manuscript submitted to ACM

ACM Reference Format:
Linxuan Zhao, Zachari Swiecki, Dragan Gašević, Lixiang Yan, Samantha Dix, Hollie Jaggard, Rosie Wotherspoon, Abra Osborne, Xinyu Li, Riordan Alfredo, and Roberto Martinez-Maldonado. 2023. METS: Multimodal Learning Analytics of Embodied Teamwork Learning. In *LAK23: 13th International Learning Analytics and Knowledge Conference (LAK 2023), March 13–17, 2023, Arlington, TX, USA*. ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3576050.3576076

1 Introduction

Learning how to collaborate in co-located settings remains an irreplaceable practice in many professional sectors, despite online teamwork practices becoming more common [58]. Highly embodied team learning is one such setting where students need to interact with others while actively using the physical learning space [18, 21]. Embodied teamwork is commonplace in high-risk sectors, such as firefighting [1], emergency healthcare [27], and naval training [52], where inadequate teamwork skills are associated with failures and safety issues [30]. In such situations, communication dynamics can be complex as discourse segments can happen in parallel at different locations of the physical space with varied team member configurations (as in Figure 1). It may be possible for researchers to apply current learning analytics (LA) innovations that model dialogue in small groups, but such innovations do not commonly consider the space where dialogue happens in parallel [20]. Teachers may also find it difficult to assess in detail the effectiveness of embodied teamwork as critical events may happen concurrently at different locations. Students may also find it difficult to reflect on their own team dynamics given that they can only observe the events happening around themselves [49].

With the rapid improvements in multimodal sensing and analysis [38], an emerging body of LA research is focusing on studying and supporting learning as it unfolds in co-located group settings by mapping from low-level multimodal trace data to higher-order collaboration constructs [34]. For example, co-located dialogue content has been modelled to explore (a) teachers' communications with students to improve the quality of their teaching [12], (b) students' performance on collaborative problem-solving tasks [43], and (c) visualisations of communication behaviours used directly by group members [42]. However, analysing the dialogue of teams of students in highly dynamic physical learning environments based on theories of teamwork(e.g., [48]) has yet to be fully explored in LA.

In addition to team dialogue, spatial behaviours of team members have been found to contribute to effective teamwork. For example, Yan et al. [65] relied on positioning data from students engaged in a team task to differentiate between high and low performing teams by identifying associations among the tasks students could perform depending on where they stood. Similarly, Echeverria et al. [13] provided end-user visualisations of team positioning data to let teachers and students interpret and reflect on their spatial dynamics. Yet, to the best of our knowledge, no previous work has focused on modelling both the content of team dialogue and their spatial dynamics in co-located settings.

To address the above gaps, we propose METS (Multimodal Embodied Teamwork Signature), a method for modelling team dialogue in combination with spatial and temporal data to generate a signature of embodied teamwork. We present a study in the context of highly dynamic healthcare team simulations where students can freely move and interact with each other at different spaces in a specialised classroom (shown in Figure 1). Positioning and audio data from 60 students, grouped in teams of four, were used in the study. Epistemic networks were used as signatures to identify the differences among high and low performance teams. Through this study, we illustrate how signatures of embodied teamwork can help to identify key differences between high and low performing teams: i) across the whole learning session; ii) at different phases of learning sessions; and iii) at particular spaces of interest (SoIs) in the learning space.

2 Background

2.1 Small-group communication analysis

Digital traces of students' communication in physical learning spaces have been used to demonstrate constructs such as symmetry of contributions [4] and expertise [40]. Basic properties from audio signals – such as speech volume, number of utterances and communication overlaps – have been used to measure constructs such as the contribution to the group task [4, 50], level of task expertise [40], and quality of collaboration [6]. However, the effectiveness of those summary properties of communication has been questioned since they can only capture superficial communication behaviours [26] and are difficult to make sense of in educational terms [67].

A more robust approach to model the co-located inter-personal communication of learners is to analyse the content of dialogue [45]. This has a long tradition in the learning sciences for research purposes [20]. Yet, analysing the content of conversations in learning analytics, such as for providing immediate feedback, imposes additional technical challenges related to the automated transcription of dialogue [55]. Nonetheless, there is a body of emerging work which is exploiting the rapid advancements in sensing and automated speech recognition services to provide analytics to support improved communication in educational contexts. For example, Jensen et al. [22] created a model to automatically detect teacher-student instances of communication for providing feedback to improve teachers' discourse skills in the classroom. Focusing also on teacher-student dialogue, Kelly et al. [24] built a software system to detect the level of *authority* of teachers' questions posed to students. Pugh et al. [43] designed a software system to provide analytics on students' collaborative problem-solving (CPS) skills. Worsley and Blikstein [62] investigated several markers of expertise based on communication among students to support students' learning progression. These two previous works [43, 62] focused on analytics for research purposes. In contrast, Praharaj et al. [42] created data visualisations for a group collaboration task that portrayed the usage of words, such as frequency or connection between certain keywords, to demonstrate collaboration quality and provoke reflection. Similarly, Southwell and colleagues [55] provided a dashboard for classrooms that automatically coded discourse segments among groups of students into constructs.

Overall, existing studies demonstrate that automated analysis of group communication content is important for a wide range of educational scenarios and gaining traction. In most of these studies, the role of theory (e.g., CPS [55]) has been key in interpreting students' utterances. However, none of this work targeted highly dynamic situations where student discourse is distributed in the physical space and can happen in parallel.

2.2 Spatial behaviour analysis

Spatial behaviours are also increasingly being considered in LA research according to recent reviews [41, 66]. Small-group collaborative learning that usually happens in a constrained space, such as around a table, is one learning setting that has been explored. Computer vision techniques have been used, for example, to extract students' spatial behaviours in their intimate space while students collaborate around a table to estimate the level of individual expertise in a CPS session [39]. In a similar learning setting, Spikol et al. [56] designed a machine learning model that applied the spatial distance between students as a key indicator of success.

Some studies have focused on learning situations where students and teachers can freely move in the learning space by tracking their position using computer vision or wearables. For example, Saquib et al. [51] developed a system to analyse social interactions of pre-school kids to help teachers develop their instructional capabilities regarding spatial behaviours. Martinez-Maldonado et al. [35] focused on team teaching by developing a system to characterise teaching approaches according to the learning design based on their positional traces in the classroom. Yan et al. [63] analysed the socio-spatial behaviours of teachers and students in a large open learning space to identify students who may become socially isolated using social network analysis. Chng et al. [11] demonstrated how the students' density and connections among the collaboration activities in an open fabrication lab can influence their performance.

Similar to our approach, others have tracked the positions of team members in the context of simulation-based learning. Echeverria et al. [13] pioneered this exploration by visualising critical locations of students around a simulated hospital bed to provoke reflection during particular phases of a learning task. In a similar setting, Fernandez-Nieto et al. [14] adapted a form of epistemic network analysis to code spatial locations of students into meaningful SoIs (i.e., where only a set of relevant actions can happen) to provide a more succinct overview of where the students were in the physical space, instead of showing all the positioning traces on a map. In a more complex setting, Yan et al. [65] mapped students' spatial behaviours to task-related constructs using epistemic network analysis, demonstrating how the spatial data alone can serve as an indicator for evaluating students' performance for learning designs where the space is associated with specific tasks that students can perform.

2.3 Research gap and research questions

The prior work discussed above suggests there is increased interest in exploring positioning and proximity data to investigate spatial behaviours during learning tasks that involve embodied teamwork. However, none of this work has considered both the team dialogue and the spatial behaviours together to model embodied group communication. Ignoring the relationship between these two modalities is potentially problematic. In settings with complex communication dynamics, embodied team dialogue usually happens in close proximity (such as less than 1.5 meters in a hospital ward [54]) to have effective dialogue without interference from other dialogues happening in parallel. In turn, analytic methods should be sensitive to the co-location of teammates if their results are to be valid.

As suggested by theories of communication in healthcare teams [16, 17], the co-occurrence of communication behaviours can identify highly-effective teamwork. If techniques measuring these co-occurrences do not consider the co-location of team members, spurious relationships can be measured—e.g., measuring the co-occurrence of behaviours occurring in distinct locations where different team members are not attending to one another. Current co-occurrence-based models, such as epistemic network analysis (ENA) [53], can account for the co-location of team members. However, until now, this has been done only using information from the task design (e.g., a variable that indicates when teammates are presumed to be working together, such as a specific phase of an activity). Here, we present a method for identifying co-located communication based on dynamic spatial behaviour and integrate this information into a subsequent co-occurrence model using ENA to create multimodal embodied teamwork signatures (METSs).

To test our method, we formulated the following three research questions to explore the extent to which it can account for the complexity of discourse segments happening in parallel at different locations by modelling digital traces of dialogue (audio traces) and spatial behaviours (indoor positioning traces) into METSs. The first question explores whether the METSs can help uncover major differences among students across an entire embodied team learning session: **RQ1 –** To what extent can the METSs serve to identify key differences between high and low performing teams participating in an embodied teamwork task? Since students may develop different communication dynamics as the learning activity progresses [25], the second question explores the potential contribution of METS to investigate such differences across time: **RQ2 –** To what extent can the METSs serve to identify key differences between high and low performing teams at various phases of the learning design? Finally, since students in a highly dynamic learning space can perform different tasks at different locations [65], the third question investigates the potential of METS to investigate such differences at various locations as follows: **RQ3 –** To what extent can the METSs serve to identify key differences between high and low performing teams in different SoIs?

3 Method

3.1 The learning situation

Our method is illustrated through a set of team simulations about emergency healthcare. These simulations were a compulsory component of an undergraduate nursing degree at Monash University. The simulations were held over four weeks between August and September in 2022. Multimodal data (see details below) from 228 consenting students (aged 20 to 23), grouped in 57 teams of four, were collected. Yet, due to limitations of the microphone hardware (e.g., microphone disconnections) and other practical challenges (e.g., students accidentally moving microphones away from their faces or turning them off), data from 60 students grouped in 15 teams were used in this study. Three nursing teachers, who were the designers of the simulations, assessed participants' performance. They also monitored the simulation classroom from a control room built behind a single-direction mirror and remotely controlled the patient manikins to progress the simulation.

This simulation required students to demonstrate highly-effective teamwork, communication, and spatial skills to identify which patient out of four in a hospital ward required close attention (see four patient beds in the simulation classroom in Figure 1). The learning design of the simulation had four phases. In the **phase 1 (Initial Handover)**, two students would enter the simulation room with a teacher who explained the current situation of the four patients to the two students. The teacher leaving the room marked the beginning of **phase 2 (Initial Assessment)**. The students were expected to make a plan, allocate tasks, and then start enacting their plan. At the middle of this phase, the patient on bed 4 would start to deteriorate, which meant they entered into an emergency. Students were expected to escalate the situation by calling the "Medical Emergency Team (MET)" for help. Then, the other two students in the team would enter the simulation marking the beginning of **phase 3 (Resolving Emergency)**. The four students were expected to collaborate to help the deteriorated patient on bed 4 (defined by teachers as the **Primary task**) while taking care of other stable patients in beds 1, 2, and 3 (the **Secondary tasks**). Once the students completed specific tasks, one teacher would enter the room enacting the role of a doctor to support the students, marking the beginning of **phase 4 (Emergent Diagnosis)**. The doctor would ask the students for critical medical information, guide them to determine the cause of deterioration, resolve it, and finalise the simulation. Commonly, the simulation would take between 15 and 30 minutes. The average duration of simulations was 20.25 minutes (standard deviation = 8.13 minutes). Teachers focused on assessing students' team dynamics centred on phases 2 and 3.

3.2 Data collection

Audio signal, spatial data, and team evaluation results were collected from the simulations and used in this study. For audio data, a highly-portable wireless headset microphone was provided to each consenting student to collect their voices. These were Xiaokoa microphones that had a wireless receiver embedded directly on the headset. A multi-channel (TASCAM US-16×08) audio interface was also used to receive and synchronise the audio streams and store them into individual files. For spatial data, a waist bag containing a Pozyx positioning tag (a sensor device to collect spatial data) was provided to each student. The data collected by Pozyx tag included body orientation and x-y spatial coordinates. To evaluate team performance, we provided a questionnaire to teachers who participated in this study. The questionnaire aimed to assess three critical learning objectives about: 1) teamwork effectiveness, 2) communication effectiveness, and

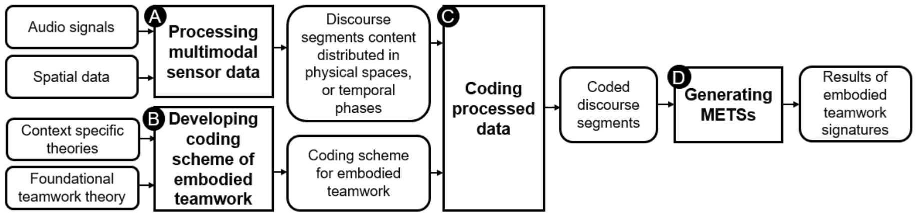

Fig. 2. Procedures and artefacts of multimodal embodied teamwork signature (METS). The rectangles containing bolded text mean the procedures and the rounded rectangles mean the artefacts used or created by procedures.

3) role comprehension. A 7-points Likert scale was used to score each learning objective. As our focus in this study was on teamwork effectiveness, we divided the 15 teams into seven low-performing and eight high-performing teams based on the median value of teamwork effectiveness score. All data collection devices were controlled by a software that automatically synchronised the collected data. Ethical approval was obtained from Monash University Human Research Ethics Committee.

3.3 The METS modelling

Figure 2 depicts the four procedures (A, B, C, and D) and artefacts (i.e., team theory, sensor data, and processed data) required for extracting METS.

3.3.1 A – Processing multimodal sensor data The first procedure (A) focuses on processing the *Audio Signals* and *Spatial Data* to create *discourse segments distributed across the physical learning space* for each phase in the learning design. Utterance timing (i.e., when an utterance was made and its duration) and dialogue content were extracted from audio signals. The utterance timing is then used for determining the phases in which an utterance was made and organising utterances into discourse segments. The utterance timing was automatically extracted by a voice activity detection script written via the Python library *py-webrtcvad*. Dialogue content was manually acquired by a third-party transcription service and then verified by researchers.

For METS, spatial data are needed to organise discourse segments across the multiple locations where students interacted. Based on the body orientation and spatial coordinates, instances of co-location were automatically detected by applying the theoretical notion of f-formations. An f-formation arises whenever two or more people, in close proximity, sustain a spatial and orientational relationship in which the space between them is one to which they have equal, direct, and exclusive access [33]. A previous study in healthcare [54] suggested that communication between healthcare professionals commonly happens in close proximity (less than 1.5 meters). We applied an f-formation detection algorithm [67] to determine who was in the same conversation with whom. This algorithm determined whether students were in the same f-formation if they were in close proximity (i.e., 1.5 m distance) and their bodies were not divergent (e.g., if they were oriented towards the same space or were working side by side). In addition to defining conversations, the spatial data were used to detect SoIs and filter discourse segments to focus on conversations that occur in a particular location of the learning space. The SoIs we identified in the data were those related to the *primary tasks* (occurring around the bed of the patient at risk) and those related to the *secondary tasks* (occurring around the beds of the stable patients).

3.3.2 B – A coding scheme for embodied teamwork To analyse the processed data, we developed a coding scheme for this embodied teamwork context by adapting previously developed schemes for communication in healthcare

Table 1. Coding scheme for teamwork communication and inter-rater reliability of each communication behaviour

Teamwork construct	**Description**	**Communication behaviours**	**Kappa**
Shared leadership	Team members jointly take the responsibility of directing the team, such as making decisions and guiding new team members.	Task allocation [44]	0.744
		Provision of handover information [23]	0.781
Situation awareness	Team members show the awareness of change in patient states and react to the change.	Situation assessment [19]	0.747
		Escalation [8]	0.853
Shared mental model	A shared understanding of the current situation and plan among team members.	Information sharing [59]	0.744
		Information requesting [2]	0.794
		Responding to request [2]	0.804
		Planning [57]	0.896
Closed-loop communication	Verbal activities to double-check or complete the exchange of information.	Agreement [17]	0.858
		Disagreement [17]	0.856
		Checking-back [17]	0.922

[5, 37, 44] and Salas et al.'s big five of teamwork framework [48]. Table 1 shows the four high-level constructs and eleven corresponding communication behaviours we applied.

The first construct is *shared leadership*. Although centralised leadership is commonly observed in healthcare teams [47], shared leadership is more frequently observed in this learning setting, as no specific leader is assigned [9]. This construct includes two communication behaviours: task allocation [44] and provision of handover information [23]. *Task allocation* occurs when a student explicitly assigns a task to others or proactively self-allocates a task. This is an example utterance coded as task allocation: "You do the medical observation, and I will do the discharge for the bed three patient". *Task allocation* can also be done through a polite question (e.g., "Would you like to do the medical observation?"). *Provision of handover information* occurs when a student proactively updates a team member regarding a task to which they have not been exposed, such as describing the situation of a patient to a new team member [10]. An example utterance for this code is the following: "Imani (one of the manikin patients) is day-one post total hysterectomy. She has got a history of heart disease".

The second construct, *situation awareness* captures team members' communication about identifying and reacting to the emergency that happened to patients [44] through two coded behaviours. *Situation Assessment* occurs when a student explicitly announces or warns others about the improvement or decline of a patient's state [19]. This is commonly observed when a student finishes doing some medical assessment (e.g., "The patient's oxygen saturation is very low"). *Escalation* occurs when a student informs others that the situation goes beyond their capabilities and they need extra help [8]. This behaviour would commonly be observed when an emergency happened to a patient (e.g., "I think we need to call the MET team (the two students waiting outside) to help us do the checklist".

The third construct, *shared mental model*, captures the communication aimed at synchronising the context information and mutual understanding among team members [61] and includes four coded behaviours. *Information sharing* is observed when students proactively share information, such as readings from a device, or patient status, to others without being asked about this information. This behaviour is usually observed as call-outs from students [59] (e.g., "Her wound is dry and intact. There is no concern now"). *Information requesting* is observed when a student asks someone else a question to get information [2]. Commonly, this is done by asking a question (e.g., "was she very drowsy when she was handed over, as well?"). *Responding to request* captures cases when a student provides information corresponding to a previously asked question. An example utterance coded this way would be "No, she does not need that. It is fine", to answer a previously asked question: "Is the ceftriaxone necessary for this patient?" (which would be coded as Information requesting). *Planning* is captured when students list a series of tasks remaining to be done to other students for provoking subsequent task allocation [57]. This behaviour can commonly be observed at the beginning of a simulation (e.g., "We should go through what they need. That patient is due for antibiotics, pain medicines, and the checklist, and we also need to get the phone to call their family").

The last construct is *Closed-loop communication* which refers to students double-checking information or ending a communication interchange where no extra information would be added [37] and includes three behaviours. *Agreement* and *Disagreement* are observed when students explicitly express their agreement or disagreement to another student. It can be observed in simple "yes" or "no", and any phrase that conveys the same meaning. *Check-back* captures cases when a student double-checks the information or instructions that another student provides [17]. An example can be student A asks student B: "Can you help me to give her some oxygen?", and student B double-checks what student A asked: "Do you want me to give her some oxygen?".

3.3.3 C – Coding processed data The dialogue content was coded at an utterance level. An utterance can have multiple codes, since a student could express multiple communication behaviours in one utterance. For example, an utterance of a student can be: *"Yes, her oxygen saturation is dropping, so I think we need to call the other two nurses"*; which demonstrates a situation where the student shows agreement ("*Yes*"), situation assessment ("*her oxygen saturation is dropping*"), and escalation ("*so I think we need to call the other two nurses in*") in one utterance. The coding was completed by two researchers. Twenty per cent of data was selected to be coded by both researchers (756 utterances from 3 simulation sessions). Cohen's kappa was used as the metric for inter-rater reliability and we chose a value of 0.6 as a threshold for suitable agreement as suggested by a previous study on inter-rater reliability evaluation [36]. For each code we achieved a kappa greater than 0.7 (see Table 1 – column 4). The remaining 80 per cent of data was coded by one researcher, and then the coded data was used for further analysis.

3.3.4 D – Generating the METSs As suggested by theories of communication in healthcare teams [16, 17], the co-occurrence of communication behaviours can serve to identify highly-effective teamwork. For example, the co-occurrence of codes such as *task allocation*, *checking-back*, and *agreement* can be indicative of effective close-loop communication (CLC) [17]. Such evidence can also demonstrate students' understanding of formal communication protocols teams are meant to follow [17]. To account for the co-occurrence of behaviours in the context of embodied teamwork, we analysed the data using ENA [53].

Prior to analysis with ENA, the data had to be ordered and segmented as *lines*–the fundamental units of meaning in the data; *stanza windows*—spans of lines that form the recent common ground of the discourse; and *conversations*—collections of lines that are potentially meaningfully related—e.g., all of the lines for a given team on a given day. The ENA algorithm creates a network model for each line in the data by counting the co-occurrences between each pair of codes within that line's stanza window (i.e, a pre-defined span of prior lines). Networks are aggregated across lines within each unit of analysis and each conversation to obtain each unit's cumulative network. Next, a dimensional reduction via Singular Value Decomposition (SVD) is applied to the collection of all networks to produce a low dimensional space that maximally distinguishes units in terms of the co-occurrence between codes in their data. The networks are represented both as points in this space and weighted network graphs whose nodes are the codes and edges correspond to the relative frequency of co-occurrence between codes. The nodes of these networks are positioned to correspond with the loadings matrix of the dimensional reduction and can thus be used to interpret the dimensions of the space.

For this analysis, we used both the R implementation of ENA [31] and the web-tool version [32]. We selected the unique combinations of teams, phase, and SoI as the units of analysis, utterances (i.e., turns of talk) as lines, a stanza window size of four lines, and co-located discourse segments identified by the f-formation algorithm as conversations. We grouped teams according to high and low performance as defined in Section 3.2. The dimensional reduction employed

Table 2. Regression analysis results (estimated marginal means). All confidence intervals are non-overlapping, suggesting that METS can differentiate high and low performing teams across the whole learning session, by learning phase, and by space of interest.

RQ1: Whole session											
Perf.	mean	SE	df	lower.CL	upper.CL						
High	0.092	0.035	56	0.022	0.161						
Low	-0.105	0.037	56	-0.179	-0.031						
RQ2: Learning phase = *initial assessment*						**Learning phase = *resolving emergency***					
Perf.	mean	SE	df	lower.CL	upper.CL	Perf.	mean	SE	df	lower.CL	upper.CL
High	0.033	0.043	56	-0.053	0.119	High	0.150	0.043	56	0.064	0.236
Low	-0.163	0.045	56	-0.253	-0.073	Low	-0.046	0.045	56	-0.136	0.043
RQ3: SoI = *primary tasks*						**SoI = *secondary tasks***					
Perf.	mean	SE	df	lower.CL	upper.CL	Perf.	mean	SE	df	lower.CL	upper.CL
High	0.096	0.043	56	0.011	0.182	High	0.087	0.043	56	0.001	0.173
Low	-0.100	0.045	56	-0.190	-0.010	Low	-0.110	0.045	56	-0.199	-0.020

by the ENA algorithm was constrained to find the maximal difference between these two groups on the first dimension of the ENA space in terms of their co-occurrence patterns [7]. Prior to the analysis, we removed data from the first (*Initial Handover*) and last phases (*Emergent Diagnosis*) because little communication happened among students in these phases, as the communication was mainly led by the teachers.

3.4 Analysis

To address our research questions, we regressed the position of each unit of analysis on the first dimension of the ENA space on categorical predictors for performance (high/low), phase (2/3) and SoI (primary/secondary). For this regression, the outcome variable was nested within teams–that is, there was one outcome for each combination of team, phase, and SoI. However, confidence intervals for the intraclass correlation coefficient that described the nesting of outcomes into teams suggested that a multi-level model was not necessary. Thus, we conducted a standard regression analysis.

To address RQ1, we used the regression model to calculate the *estimated marginal means* on the first dimension of the ENA space for high and low performing teams and compared those means using 95% confidence intervals. Similarly, to address RQ2 we compared the estimated marginal means of the high and low performing teams for each phase, averaged over the levels of SoI. To address RQ3, we compared the estimated marginal means of the high and low performing teams for each SoI, averaged over the phases. To interpret any differences in these means, we inspected the ENA network subtractions for each comparison. These diagrams subtract the averaged edge weights for each group (i.e., high/low performers) to show the co-occurrences that were more frequent in one group compared to the other.

4 Results

4.1 RQ1 - METs characterising high and low performing teams

The estimated marginal means from the regression analysis are presented in Table 2. For RQ1, the means of the high and low performing teams were compared and averaged over the levels of phase and SoI. The confidence intervals for these means do not overlap, suggesting that the high and low performing teams are statistically significantly different from one another on the first dimension of the ENA space.

Figure 3 shows the mean network subtraction for the high and low performing teams. Compared to high-performing teams (blue), the low-performing teams (red) had more prominent connections among *task allocation*, *information sharing*, and *agreement* and between *planning* and *agreement*. The connections among *task allocation*, *information sharing*, and *agreement* may demonstrate the low-performing teams tended to spend more time allocating tasks to collect information about the scenario. For example, as evidenced in the data, student X politely asked student Y to

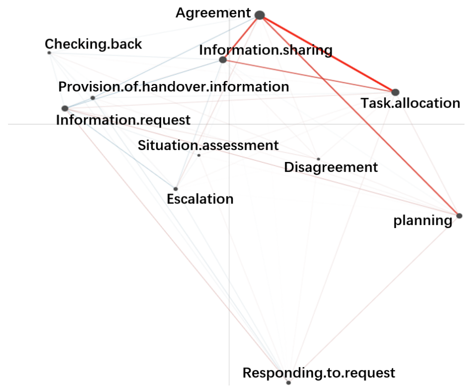

Fig. 3. Mean Network subtraction for low (red edges) and high (blue edges) performing teams across the entire learning session.

check the morphine status: "Do you want to do the morphine check real quick?" (*task allocation*). Then, student Y responded to student X as follows – "Yes, okay" (*agreement*) – and provided some information – "The morphine is running." (*information sharing*). Finally, student X acknowledged that they received the information – "okay" (*agreement*). The connection between the codes *planning* and *agreement* also suggests that low-performing teams tended to spend considerable time planning. This may suggest that these teams did not necessarily plan efficiently.

4.2 RQ2 - characterising embodied teamwork across phases of the learning design

The regression results in Table 2 show that METSs could differentiate high and low performing teams in different learning phases. For both phases, the confidence intervals for these means did not overlap, suggesting that the high and low performing teams were statistically significantly different from one another on the first dimension of the ENA space when controlling for the learning phase.

Figure 4 shows the mean network subtractions between high and low performing teams in the *initial assessment* phase (left) and the *resolving emergency phase* (right). In the initial assessment phase, the goal for students was to identify which patient required urgent attention. Highly effective teams would be expected to rapidly exchange information and explicitly escalate the situation to ask for help. Indeed, high-performing teams demonstrated their more effective team dynamics through the stronger connection between two information-related communication codes, *information requesting* and *information sharing* as well as stronger connections to the code *escalation*. In contrast, the prominent connections between *task allocation* and *agreement*, and between *planning* and *agreement*, in low-performing teams suggest that they may have not identified the urgency of the critical task, and thus may have spent more time in planning and allocating themselves to non-critical tasks.

In the *resolving emergency* phase (Figure 4, right), the goal for students was to provide close care to the patient who needs more help before the emergency doctor arrives. In this phase, two new students entered the simulation room, so the students need to provide medical help to the patient at risk, while quickly organising themselves to also complete medical tasks for the stable patients. Therefore, codes related to task allocation and collecting information were expected as well as the provision of handover information to the new team member and other codes related to requesting or providing information. However, low-performing teams may not have correctly recognised the key information required for helping those patients, so they spent more time assigning tasks to collect information (see most prominent connections for low-performing teams among *task allocation*, *information sharing*, and *agreement* in Figure 4, right) which may have contributed to failure in demonstrating effective teamwork. In contrast, high-performing

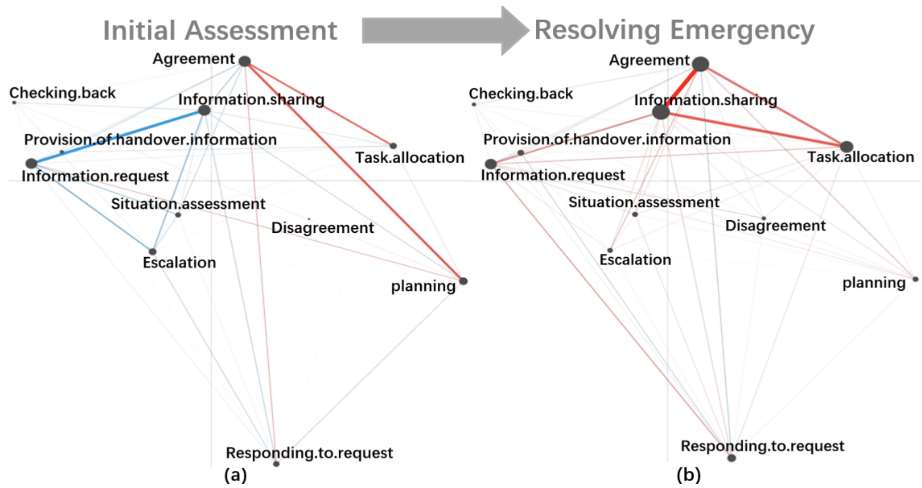

Fig. 4. Mean network subtractions for low (red edges) and high (blue edges) performing teams during the *initial assessment* phase (left) and *resolving emergency* phase (right).

teams may have recognised the key information faster, resulting in fewer and more balanced connections among *task allocation*, *information sharing*, and *agreement* in relation to other codes.

4.3 RQ3 - characterising embodied teamwork across spaces of interest

The regression results in Table 2 show that METSs could differentiate high and low performing teams in different SoIs. For both the primary and secondary task locations, the confidence intervals for these means did not overlap, suggesting that the high and low performing teams were statistically significantly different from one another on the first dimension of the ENA space when controlling for SoI.

Figure 5 shows the mean network subtractions between high and low performing teams at the critical locations for primary task (left) and secondary tasks (right). Regarding the primary task location, high-performing teams had stronger connections among codes related to exchanging information, namely *information requesting*, *responding to request*, *information sharing*, and *provision of handover information*. In contrast, low-performing teams had stronger connections between *task allocation* and *agreement*, and a stronger connection to *planning*. The stronger connections of high-performing teams related to information-related communication codes illustrate that they spent more time than low-performing teams exchanging information about the patient at risk. On the other hand, the stronger connections between *task allocation* and *agreement* and stronger connections to *planning* suggest that low-performing teams spent more time assigning tasks and planning for the high-risk patient. In turn, low-performing teams may have overemphasised the coordination of the tasks, reducing the time spent providing medical support to the patient at risk.

Regarding the secondary task location (Figure 5, right), compared to high-performing teams, low-performing teams had stronger connections between *agreement*, and *task allocation*, and stronger connections among several communication behaviours related to exchanging information—namely *information requesting*, *responding to request*, and *information sharing*. This result suggests that low-performing teams spent more time assigning tasks and collecting/exchanging information related to secondary tasks than high-performing teams. In turn, low-performing teams may have overemphasised tasks for stable patients and neglected tasks for the patient at risk.

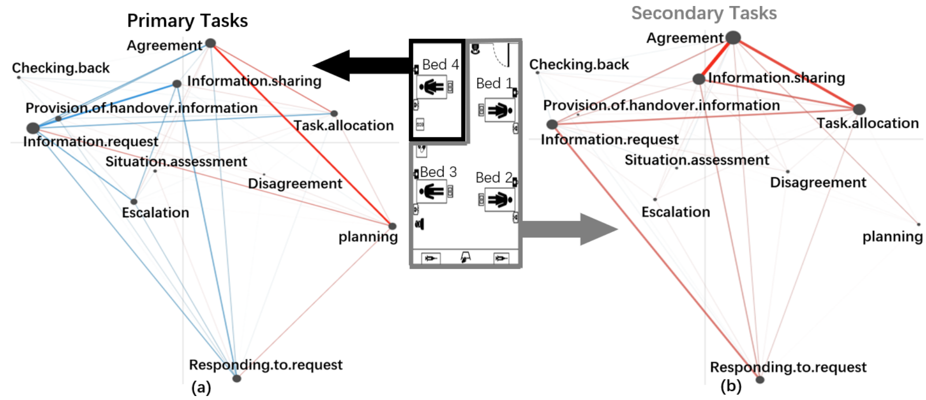

Fig. 5. Mean network subtractions for low (red edges) and high (blue edges) performing teams at the primary task location (left) and secondary task locations (right).

5 Discussion

5.1 Summary of results

In this study, we illustrated METS, a method based on ENA to model team dialogue content in highly dynamic and embodied team learning settings. In such settings, discourse segments can happen in parallel at different locations of the physical space with varied team member configurations. The key contribution of the method is the identification of co-located discourse segments (via the f-formation algorithm) to appropriately constrain the co-occurrences measured by ENA. We addressed our RQs by exploring to what extent METSs can help to differentiate between high and low performing teams participating in an embodied teamwork task while accounting for specific variables related to the task design. Our results showed that METSs can be used to differentiate high and low performing teams over the whole learning session, in two learning phases, and at two spaces of interest in the context of a healthcare simulation. Furthermore, we described and interpreted the differences in the communication behaviours of high and low performing teams through three groups of ENA mean network subtractions.

The results can be summed up as follows. First, the low-performing teams focused more on coordination, compared to high-performing teams, demonstrated by stronger connections between the codes *planning* and *task allocation* observed in four out of the five network comparisons (Figures 3, 4 (left), 4 (right), and 5 (left)). Effective task coordination is essential to improve teamwork efficiency in emergency healthcare [30]. However, these results suggest that low-performing teams were overly focused on coordination at the expense of other important behaviours.

A second key finding is related to the critical ability of nursing students to prioritise their care activities (e.g., recognise the patient at risk) while omitting or delaying other secondary tasks. Errors or delays in identifying and prioritising patients who may need urgent medical attention can contribute to negative patient outcomes in hospital wards [15]. Our results showed that high-performing teams focused more on exchanging information in the initial assessment phase, in which teams needed to quickly assess which patient was at risk (demonstrated through the stronger connections between *information requesting* and *information sharing* and connections to *escalation*). This supports the idea that highly effective teamwork in this context involves quick task allocation and longer conversations among team members to gather evidence in order to recognise whether a situation must be escalated.

A third finding emerged in the kind of dialogue that high-performing teams had near the space surrounding the bed with the patient that needed closer care. High-performing teams focused more on exchanging information in the primary task space (illustrated in Figure 5-left), while low-performing teams focused more on exchanging information in the secondary task space (illustrated in Figure 5-right). This suggests that while all teams understood the significance of building a shared mental model through information exchange, which is a key feature of highly-effective teams in healthcare [29], high-performing teams realised they needed to do this for the patient they needed to prioritise. Correctly prioritising primary from secondary tasks is also related to the teamwork construct of effective situation awareness in patient safety (the ability to perceive, prioritise and analyse critical information that can help guide decisions [46])

5.2 Implications

The method presented in this paper has several implications for future learning analytics embodied teamwork and co-located collaborative learning. Our method demonstrated how to combine students' team dialogue and spatial behaviours to identify co-located communication in a learning setting where multiple discourse segments can happen in parallel. This method can benefit the researchers interested in studying students' communication in embodied learning settings with similar complex communication dynamics, such as naval team training [52]. Our method can also benefit researchers interested in analysing how team performance differs in embodied learning settings. Moreover, our method can provide a useful way for other LA researchers to express multimodal data using ENA. This can benefit researchers who want to investigate how multiple modalities can supplement the analysis of discourse, especially, in learning settings where multiple modalities can influence the students' communication behaviours. An example of such a setting is an open learning space where different locations hold different learning activities that may result in different communication behaviours [64]. Besides, this method can also benefit teachers and students eventually. A network subtraction can be created between a team's network and the mean network of high-performing teams. A visual interface could then be developed to support teachers by highlighting salient aspects of team communication and potential mismatches compared to previously recognised high-performing teams. This could be used to provoke reflective discussions among students on aspects they can improve.

5.3 Ethical considerations

The major ethical issue in this study is privacy. The simulations were recorded with audio. Although audio recording is a common practice in this particular educational context [13], the use of recording devices to capture students' data can lead to privacy issues, such as unintended surveillance [3]. To minimise the ethical concerns, we cleared all the identity information from our dataset and solely used colours to differentiate students and strictly control the access to prevent any unintended use [3, 66].

5.4 Limitations and future works

Our study has at least two important limitations. First, our procedures for extracting and coding the dialogue content were not fully automated, which limited the initial scalability of our approach [66]. One way to resolve this is using automated transcription. We tested the feasibility of applying automated transcription in our learning setting. We picked out audio recordings with the best audio quality containing 109 utterances. We used the Google Automatic Speech Recognition service, which has been in previous data-intensive educational studies [12, 55], to transcribe audio recordings. The word error rate (WER) [60] of transcription was 25.8%. This result is similar to the value reported in such previous studies [12, 55]. One of the most frequent automated transcription errors is caused by special words, such as human names and medical terms, so the WER could be reduced if a term dictionary is used. A term dictionary can register those terms into an automated transcription service to reduce the WER. Besides, the coding of dialogue can also be automated. Studies in the natural language processing have demonstrated reliable deep-learning methods for automatically coding conversations [28, 43]. Our future work will explore these techniques for coding. The second limitation is regarding generalisability. Due to microphone hardware issues, our sample size was limited to 15 teams. While the effects reported in this study were strong enough to obtain statistical significance, we do not assume that our results necessarily generalise beyond the specific healthcare context we examined. Nonetheless, our results may be insightful for future studies of similar contexts.

6 Conclusion

In this study, we demonstrated METS for generating signatures of multimodal embodied teamwork. We illustrated the capability of METS through addressing three RQs in a learning setting involving complex communication and spatial dynamics. The research findings illustrate the capability of METS to identify key differences in communication behaviours related to students' embodied teamwork performance. The method in this paper can benefit researchers who study embodied teamwork in learning settings with complex communication dynamics.

Acknowledgments

This research was funded partially by the Australian Government through the Australian Research Council (project number DP210100060). Roberto Martinez-Maldonado's research is partly funded by Jacobs Foundation.

References

[1] A. Adrot and M. B. Figueiredo. 2019. "Lost in Digitization": A Spatial Journey in Emergency Response and Pragmatic Legitimacy. In *Materiality in Institutions*, de Vaujany FX., A. A., B. E., and L. B. (Eds.). Springer, 151–181. https://doi.org/10.1007/978-3-319-97472-9_6

[2] A. Alonso and D. Dunleavy. 2012. Building teamwork skills in healthcare: the case for communication and coordination competencies. In *Improving patient safety through teamwork and team training*, E. Salas and K. Frush (Eds.). Oxford University Press, 41–58.

[3] H. Alwahaby, M. Cukurova, Z. Papamitsiou, and M. Giannakos. 2021. The evidence of impact and ethical considerations of Multimodal Learning Analytics: A Systematic Literature Review. https://doi.org/10.35542/osf.io/sd23y

[4] 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 (2010), 203–213. https://doi.org/10.1109/TLT.2010.18

[5] G. Barton, A. Bruce, and R. Schreiber. 2018. Teaching nurses teamwork: Integrative review of competency-based team training in nursing education. *Nurse Education in Practice* 32 (2018), 129–137. https://doi.org/10.1016/j.nepr.2017.11.019

[6] N. Bassiou, A. Tsiartas, J. Smith, H. Bratt, C. Richey, E. Shriberg, C. D'Angelo, and N. Alozie. 2016. Privacy-Preserving Speech Analytics for Automatic Assessment of Student Collaboration. In *Proc. Interspeech 2016*. 888–892. https://doi.org/10.21437/Interspeech.2016-1569

[7] D. Bowman, Z. Swiecki, Z. Cai, Y. Wang, B. Eagan, J. Linderoth, and D. W. Shaffer. 2021. The Mathematical Foundations of Epistemic Network Analysis. In *Advances in Quantitative Ethnography*, A. R. Ruis and S. B. Lee (Eds.). Springer International Publishing, Cham, 91–105.

[8] P. W. Brady and L. M. Goldenhar. 2014. A qualitative study examining the influences on situation awareness and the identification, mitigation and escalation of recognised patient risk. *BMJ Quality & Safety* 23, 2 (2014), 153–161. https://doi.org/10.1136/bmjqs-2012-001747

[9] J. B. Carson, P. E. Tesluk, and J. A. Marrone. 2007. Shared Leadership in Teams: An Investigation of Antecedent Conditions and Performance. *Academy of Management Journal* 50, 5 (2007), 1217–1234. https://doi.org/10.5465/amj.2007.20159921

[10] W. Chaboyer, A. McMurray, J. Johnson, L. Hardy, M. Wallis, and F. Y. S. Chu. 2009. Bedside handover: quality improvement strategy to "transform care at the bedside". *Journal of nursing care quality* 24, 2 (2009), 136–142. https://doi.org/10.1097/01.NCQ.0000347450.90676.d9

[11] E. Chng, M. R. Seyam, W. Yao, and B. Schneider. 2020. Using Motion Sensors to Understand Collaborative Interactions in Digital Fabrication Labs. In *Artificial Intelligence in Education*, I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, and E. Millán (Eds.). Springer International Publishing, Cham, 118–128. https://doi.org/10.1007/978-3-030-52237-7_10

[12] M. E. Dale, A. J. Godley, S. A. Capello, P. J. Donnelly, S. K. D'Mello, and S. P. Kelly. 2022. Toward the automated analysis of teacher talk in secondary ELA classrooms. *Teaching and Teacher Education* 110 (2022). https://doi.org/10.1016/j.tate.2021.103584

[13] V. Echeverria, R. Martinez-Maldonado, and S. Buckingham Shum. 2019. Towards Collaboration Translucence: Giving Meaning to Multimodal Group Data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. New York, NY, USA, 1–16.

[14] G. M. Fernandez-Nieto, R. Martinez-Maldonado, K. Kitto, and S. Buckingham Shum. 2021. Modelling Spatial Behaviours in Clinical Team Simulations Using Epistemic Network Analysis: Methodology and Teacher Evaluation. In *LAK21: 11th International Learning Analytics and Knowledge Conference*

(Irvine, CA, USA) *(LAK21)*. New York, NY, USA, 386–396. https://doi.org/10.1145/3448139.3448176

[15] P. C. Gillan, L. J. Delaney, N. Tutticci, and S. Johnston. 2022. Factors influencing nursing students' ability to recognise and respond to simulated patient deterioration: A scoping review. *Nurse Education in Practice* 62 (2022), 103350. https://doi.org/10.1016/j.nepr.2022.103350

[16] K. M. Haig, S. Sutton, and J. Whittington. 2006. SBAR: A Shared Mental Model for Improving Communication Between Clinicians. *The Joint Commission Journal on Quality and Patient Safety* 32, 3 (2006), 167–175. https://doi.org/10.1016/S1553-7250(06)32022-3

[17] M. Härgestam, M. Lindkvist, C. Brulin, M. Jacobsson, and M. Hultin. 2013. Communication in interdisciplinary teams: exploring closed-loop communication during in situ trauma team training. *BMJ Open* 3, 10 (2013). https://doi.org/10.1136/bmjopen-2013-003525

[18] J. Hindmarsh and A. Pilnick. 2007. Knowing Bodies at Work: Embodiment and Ephemeral Teamwork in Anaesthesia. *Organization Studies* 28, 9 (2007), 1395–1416. https://doi.org/10.1177/0170840607068258

[19] J. Holsopple, M. Sudit, M. Nusinov, D. F. Liu, H. Du, and S. J. Yang. 2010. Enhancing situation awareness via automated situation assessment. *IEEE Communications Magazine* 48, 3 (2010), 146–152. https://doi.org/10.1109/MCOM.2010.5434386

[20] I. Howley, E. Mayfield, and C. P. Rosé. 2013. Linguistic analysis methods for studying small groups. In *The international handbook of collaborative learning*. Routledge, 184–202.

[21] M. Ioannou, Y. Georgiou, A. Ioannou, and M. Johnson. 2019. On the understanding of students' learning and perceptions of technology integration in low-and high-embodied group learning. 1 (2019), 304–311. https://doi.org/10.22318/cscl2019.304

[22] E. Jensen, M. Dale, P. J. Donnelly, C. Stone, S. Kelly, A. Godley, and S. K. D'Mello. 2020. Toward Automated Feedback on Teacher Discourse to Enhance Teacher Learning. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376418

[23] C. M. Jorm, S. White, and T. Kaneen. 2009. Clinical handover: critical communications. *Medical Journal of Australia* 190, 11 (2009), 108–109. https://doi.org/10.5694/j.1326-5377.2009.tb02613.x

[24] S. Kelly, A. M. Olney, P. Donnelly, M. Nystrand, and S. K. D'Mello. 2018. Automatically Measuring Question Authenticity in Real-World Classrooms. *Educational Researcher* 47, 7 (2018), 451–464. https://doi.org/10.3102/0013189X18785613

[25] S. Knight, A. Friend Wise, and B. Chen. 2017. Time for Change: Why Learning Analytics Needs Temporal Analysis. *Journal of Learning Analytics* 4, 3 (Dec. 2017), 7–17. https://doi.org/10.18608/jla.2017.43.2

[26] C. N. Lacerenza, S. L. Marlow, S. I. Tannenbaum, and E. Salas. 2018. Team development interventions: Evidence-based approaches for improving teamwork. *American psychologist* 73, 4 (2018), 517. https://doi.org/10.1037/amp0000295

[27] M. Leonard, S. Graham, and D. Bonacum. 2004. The human factor: the critical importance of effective teamwork and communication in providing safe care. *BMJ Quality & Safety* 13, suppl 1 (2004), i85–i90. https://doi.org/10.1136/qshc.2004.010033

[28] J. Lin, S. Singh, L. Sha, W. Tan, D. Lang, D. Gašević, and G. Chen. 2022. Is it a good move? Mining effective tutoring strategies from human–human tutorial dialogues. *Future Generation Computer Systems* 127 (2022), 194–207. https://doi.org/10.1016/j.future.2021.09.001

[29] D. J. Lowe, A. J. Ireland, A. Ross, and J. Ker. 2016. Exploring situational awareness in emergency medicine: developing a shared mental model to enhance training and assessment. *Postgraduate Medical Journal* 92, 1093 (2016), 653–658. https://doi.org/10.1136/postgradmedj-2015-133772

[30] T. MANSER. 2009. Teamwork and patient safety in dynamic domains of healthcare: a review of the literature. *Acta Anaesthesiologica Scandinavica* 53, 2 (2009), 143–151. https://doi.org/10.1111/j.1399-6576.2008.01717.x

[31] Marquart, C. L., Swiecki, Z., Collier, W., Eagan, B., Woodward, R., Shaffer, and D. W. 2017. *rENA: Epistemic Network Analysis*.

[32] C. Marquart, C. Hinojosa, Z. Swiecki, B. Eagan, and D. Shaffer. 2018. Epistemic Network Analysis (Version 1.7. 0)[Software]. *Available from app. epistemicnetwork. org* (2018).

[33] P. Marshall, Y. Rogers, and N. Pantidi. 2011. Using F-Formations to Analyse Spatial Patterns of Interaction in Physical Environments. (2011), 445–454. https://doi.org/10.1145/1958824.1958893

[34] R. Martinez-Maldonado, V. Echeverria, G. Fernandez Nieto, and S. Buckingham Shum. 2020. From Data to Insights: A Layered Storytelling Approach for Multimodal Learning Analytics. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376148

[35] R. Martinez-Maldonado, V. Echeverria, J. Schulte, A. Shibani, K. Mangaroska, and S. B. Shum. 2020. Moodoo: indoor positioning analytics for characterising classroom teaching. In *International Conference on Artificial Intelligence in Education*. Springer, 360–373.

[36] M. L. McHugh. 2012. Interrater reliability: the kappa statistic. *Biochemia medica* 22, 3 (2012), 276–282.

[37] K. MILLER, W. RILEY, and S. DAVIS. 2009. Identifying key nursing and team behaviours to achieve high reliability. *Journal of Nursing Management* 17, 2 (2009), 247–255. https://doi.org/10.1111/j.1365-2834.2009.00978.x

[38] X. Ochoa. 2022. Multimodal Learning Analytics - Rationale, Process, Examples, and Direction. In *The Handbook of Learning Analytics* (2 ed.), C. Lang, G. Siemens, A. F. Wise, D. Gaševic, and A. Merceron (Eds.). SoLAR, 54–65. Section: 6.

[39] X. Ochoa, K. Chiluiza, G. Méndez, G. Luzardo, B. Guamán, and J. Castells. 2013. Expertise Estimation Based on Simple Multimodal Features. In *Proceedings of the 15th ACM on International Conference on Multimodal Interaction* (Sydney, Australia) *(ICMI '13)*. New York, NY, USA, 583–590.

[40] S. Oviatt, K. Hang, J. Zhou, and F. Chen. 2015. Spoken Interruptions Signal Productive Problem Solving and Domain Expertise in Mathematics. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (Seattle, Washington, USA) *(ICMI '15)*. New York, NY, USA, 311–318. https://doi.org/10.1145/2818346.2820743

[41] S. Praharaj, M. Scheffel, H. Drachsler, and M. Specht. 2021. Literature Review on Co-Located Collaboration Modeling Using Multimodal Learning Analytics—Can We Go the Whole Nine Yards? *IEEE Transactions on Learning Technologies* 14, 3 (2021), 367–385.

[42] S. Praharaj, M. Scheffel, M. Schmitz, M. Specht, and H. Drachsler. 2021. Towards Automatic Collaboration Analytics for Group Speech Data Using Learning Analytics. *Sensors* 21, 9 (2021). https://doi.org/10.3390/s21093156

[43] S. L. Pugh, S. K. Subburaj, A. R. Rao, A. E. Stewart, J. Andrews-Todd, and S. K. D'Mello. 2021. Say What? Automatic Modeling of Collaborative Problem Solving Skills from Student Speech in the Wild. *International Educational Data Mining Society* (2021), 55–67.

[44] W. Riley, H. Hansen, A. P. Gürses, S. Davis, K. Miller, and R. Priester. 2008. The nature, characteristics and patterns of perinatal critical events teams. *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 3: Performance and Tools)* (2008).

[45] A. P. Roberts, L. V. Webster, P. M. Salmon, R. Flin, E. Salas, N. J. Cooke, G. J. Read, and N. A. Stanton. 2022. State of science: models and methods for understanding and enhancing teams and teamwork in complex sociotechnical systems. *Ergonomics* 65, 2 (2022), 161–187.

[46] E. Salas, C. Prince, D. P. Baker, and L. Shrestha. 1995. Situation Awareness in Team Performance: Implications for Measurement and Training. *Human Factors* 37, 1 (1995), 123–136. https://doi.org/10.1518/001872095779049525

[47] E. Salas, M. A. Rosen, J. D. Held, and J. J. Weissmuller. 2009. Performance Measurement in Simulation-Based Training: A Review and Best Practices. *Simulation & Gaming* 40, 3 (2009), 328–376. https://doi.org/10.1177/1046878108326734

[48] E. Salas, D. E. Sims, and C. S. Burke. 2005. Is there a "Big Five" in Teamwork? *Small Group Research* 36, 5 (2005), 555–599.

[49] E. Salas, R. Stevens, J. Gorman, N. J. Cooke, S. Guastello, and A. A. von Davier. 2015. What Will Quantitative Measures of Teamwork Look Like in 10 Years? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 59, 1 (2015), 235–239. https://doi.org/10.1177/1541931215591048

[50] O. Salinas, F. Riquelme, R. Munoz, C. Cechinel, R. Martinez, and D. Monsalves. 2021. Can Analytics of Speaking Time Serve as Indicators of Effective Team Communication and Collaboration?. In *X Latin American Conference on Human Computer Interaction* (Valparaiso, Chile) *(CLIHC 2021)*. New York, NY, USA, Article 12, 4 pages. https://doi.org/10.1145/3488392.3488404

[51] N. Saquib, A. Bose, D. George, and S. Kamvar. 2018. Sensei: Sensing Educational Interaction. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 161 (jan 2018), 27 pages. https://doi.org/10.1145/3161172

[52] C. Sellberg, O. Lindmark, and H. Rystedt. 2018. Learning to navigate: the centrality of instructions and assessments for developing students' professional competencies in simulator-based training. *WMU Journal of Maritime Affairs* 17, 2 (2018), 249–265.

[53] D. W. Shaffer, W. Collier, and A. R. Ruis. 2016. A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics* 3, 3 (2016), 9–45. https://doi.org/10.18608/jla.2016.33.3

[54] A. Sorokowska, P. Sorokowski, P. Hilpert, K. Cantarero, T. Frackowiak, K. Ahmadi, A. M. Alghraibeh, R. Aryeetey, A. Bertoni, K. Bettache, et al. 2017. Preferred interpersonal distances: a global comparison. *Journal of Cross-Cultural Psychology* 48, 4 (2017), 577–592.

[55] R. Southwell, S. Pugh, E. M. Perkoff, C. Clevenger, J. B. Bush, R. Lieber, W. Ward, P. Foltz, and S. D'Mello. 2022. Challenges and Feasibility of Automatic Speech Recognition for Modeling Student Collaborative Discourse in Classrooms. *thinking* 27 (2022), 29. https://doi.org/10.5281/zenodo.6853109

[56] D. Spikol, E. Ruffaldi, L. Landolfi, and M. Cukurova. 2017. Estimation of Success in Collaborative Learning Based on Multimodal Learning Analytics Features. In *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*. 269–273. https://doi.org/10.1109/ICALT.2017.122

[57] R. J. Stout, J. A. Cannon-Bowers, E. Salas, and D. M. Milanovich. 1999. Planning, Shared Mental Models, and Coordinated Performance: An Empirical Link Is Established. *Human Factors* 41, 1 (1999), 61–71. https://doi.org/10.1518/001872099779577273

[58] P. A. Takizawa, L. Honan, D. Brissette, B. J. Wu, and K. M. Wilkins. 2021. Teamwork in the time of COVID-19. *FASEB BioAdvances* 3, 3 (2021), 175–181. https://doi.org/10.1096/fba.2020-00093

[59] P. Van den Bossche, W. Gijselaers, M. Segers, G. Woltjer, and P. Kirschner. 2011. Team learning: building shared mental models. *Instructional Science* 39, 3 (2011), 283–301. https://doi.org/10.1007/s11251-010-9128-3

[60] Y.-Y. Wang, A. Acero, and C. Chelba. 2003. Is word error rate a good indicator for spoken language understanding accuracy. In *2003 IEEE Workshop on Automatic Speech Recognition and Understanding (IEEE Cat. No.03EX721)*. 577–582. https://doi.org/10.1109/ASRU.2003.1318504

[61] H. K. Westli, B. H. Johnsen, J. Eid, I. Rasten, and G. Brattebø. 2010. Teamwork skills, shared mental models, and performance in simulated trauma teams: an independent group design. *Scandinavian journal of trauma, resuscitation and emergency medicine* 18, 1 (2010), 1–8.

[62] M. Worsley and P. Blikstein. 2011. What's an Expert? Using Learning Analytics to Identify Emergent Markers of Expertise through Automated Speech, Sentiment and Sketch Analysis.. In *EDM*. 235–240.

[63] L. Yan, R. Martinez-Maldonado, B. G. Cordoba, J. Deppeler, D. Corrigan, G. F. Nieto, and D. Gasevic. 2021. Footprints at School: Modelling In-Class Social Dynamics from Students' Physical Positioning Traces. In *LAK21: 11th International Learning Analytics and Knowledge Conference* (Irvine, CA, USA) *(LAK21)*. New York, NY, USA, 43–54. https://doi.org/10.1145/3448139.3448144

[64] L. Yan, R. Martinez-Maldonado, L. Zhao, J. Deppeler, D. Corrigan, and D. Gasevic. 2022. How Do Teachers Use Open Learning Spaces? Mapping from Teachers' Socio-Spatial Data to Spatial Pedagogy. In *LAK22: 12th International Learning Analytics and Knowledge Conference* (Online, USA) *(LAK22)*. New York, NY, USA, 87–97. https://doi.org/10.1145/3506860.3506872

[65] L. Yan, R. Martinez-Maldonado, L. Zhao, S. Dix, H. Jaggard, R. Wotherspoon, X. Li, and D. Gašević. 2022. The role of indoor positioning analytics in assessment of simulation-based learning. *British Journal of Educational Technology* (2022). https://doi.org/10.1111/bjet.13262

[66] L. Yan, L. Zhao, D. Gasevic, and R. Martinez-Maldonado. 2022. Scalability, Sustainability, and Ethicality of Multimodal Learning Analytics. In *LAK22: 12th International Learning Analytics and Knowledge Conference* (Online, USA) *(LAK22)*. New York, NY, USA, 13–23.

[67] L. Zhao, L. Yan, D. Gasevic, S. Dix, H. Jaggard, R. Wotherspoon, R. Alfredo, X. Li, and R. Martinez-Maldonado. 2022. Modelling Co-Located Team Communication from Voice Detection and Positioning Data in Healthcare Simulation. In *LAK22: 12th International Learning Analytics and Knowledge Conference* (Online, USA) *(LAK22)*. New York, NY, USA, 370–380. https://doi.org/10.1145/3506860.3506935