Learning and Instruction 102 (2026) 102302

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Physiological synchrony amongst medical residents during crisis management simulation training and a video-based assessment of leaders' performance

Lucia Patino Melo [a], Sebastian Wallot [b], Keerat Grewal [a], Matthew Moreno [a], Osamu Nomura [a], Jason M. Harley [a,c,d,e,*]

[a] Department of Surgery, McGill University, Montreal, Canada, H3A 0G4
[b] Institute of Psychology, Leuphana University, Lüneburg, 21335, Germany
[c] Institute of Health Sciences Education, McGill University, Montreal, Canada, H3A 0G4
[d] Research Institute of the McGill University Health Centre, Montreal, H3A 0G4, Canada
[e] Steinberg Centre for Simulation and Interactive Learning, McGill University, Montreal, H2X 3P9, Canada

ARTICLE INFO

Keywords:
Physiological synchrony
Medical education
Electrodermal activity
Recurrence analysis
MdRQA

STRUCTURED ABSTRACT

Background: Complex learning dynamics exist in ecologically valid collaborative learning environments, such as medical simulation training. Although medical scenarios are known to elicit emotional responses, the role of these emotions in shaping team performance is often overlooked. Physiological synchrony (PS) offers a promising lens to explore these emotional dynamics and, when integrated with multimodal learning analytics, can enhance our understanding of collaborative processes in medical teams, including the team leaders' performance**.**
Theoretical framework: This study was guided by socially-shared regulated learning and the control-value theory of achievement emotions. We used these theories to operationalize and situate physiological arousal as an emotional expression component, draw insight from the context simulation training occurred in, and interpret findings.
Aims: To examine PS among medical residents undergoing crisis resource management (CRM) simulation training and explore its relationship with team leaders' performance.
Sample: Participants were 29 medical trainees grouped into 9 teams participating in team-based simulations.
Methods: A field-based, quantitative approach was employed. Electrodermal activity (EDA) was used to assess participants' emotional arousal, and Multidimensional Recurrence Quantification Analysis (MdRQA) was used to evaluate PS. Team leaders' performance was scored using the Ottawa Global Rating Scale.
Results: PS was observed among team members during both pre-simulation and simulation phases. Preliminary results revealed that PS was positively correlated with team leader performance.
Conclusion: Findings offer novel insights into the existence and role of PS in ecologically-valid team-based medical simulation education. PS may reflect important aspects of teams' emotional arousal and, while preliminary, team leaders' performance.

1. Introduction

Medical training is cognitively, physically, and emotionally demanding, requiring years of rigorous preparation, increasing levels of responsibility, and a lifelong engagement in learning (McConnell, 2024). Unlike many other educational domains, medical training uniquely combines knowledge acquisition, technical skill development, and the intricate realm of human emotional experiences. Trainees regularly encounter emotionally charged situations, ranging from life-threatening emergencies and patient suffering to moments of professional triumph and compassionate care. These experiences can evoke heightened emotional responses when compared with most other educational fields (LeBlanc et al., 2015). In emergency medical settings, the stakes are particularly high. Medical professionals must integrate

This article is part of a special issue entitled: Implementing Multimodal Learning Analytics published in Learning and Instruction.

* Corresponding author. Department of Surgery, McGill University, Montreal General Hospital, 1650 Cedar Ave, R1.112, H3G 1A4, Canada.
E-mail address: jason.harley@mcgill.ca (J.M. Harley).

https://doi.org/10.1016/j.learninstruc.2025.102302
Received 7 November 2024; Received in revised form 11 December 2025; Accepted 16 December 2025
Available online 9 January 2026

cognitive, affective, and metacognitive (CAM) processes in real time, making rapid, high-pressure decisions while working within a team (Lingard, 2016). However, research has shown that acute stress and high emotional arousal can contribute to medical errors and adverse events due to communication challenges, diagnostic uncertainty, novelty of cases, and unique distractions in these contexts (Dias & Scalabrini-Neto, 2017; Harvey et al., 2010; Karnatovskaia et al., 2020; LeBlanc, 2009). Despite the importance of emotions in medicine and medical training, little is known about how these psychological processes operate at the group level.

We aim to help address this gap in the literature using physiological synchrony (PS) to explore the emotional dynamics of teams in an ecologically valid, team-based medical simulation training. Leveraging Multimodal Learning Analytics (MMLA) and an advanced quantitative technique, Multidimensional Recurrence Quantification Analysis, our findings are the first to examine the presence of physiological synchrony (PS) amongst teams of medical residents. Accordingly, our second aim to conduct a preliminary investigation of whether the presence of PS amongst team members was a predictor of the team leaders' performance is also novel. Our findings enhance understanding of collaborative processes in medical teams and advance the MMLS methods used in educational research with medical residents.

1.1. Multimodal learning analytics

Emergencies in medical simulation settings are complex and dynamic collaborative learning contexts, where cognitive, affective, and behavioral processes interact continuously. Multimodal learning analytics (MMLA) has emerged as a powerful approach for capturing and analyzing these dynamics by integrating multiple data sources, including self-reports, behavioral observations, and physiological signals, to create a richer understanding of how learning and teamwork unfold in real time (Cohn et al., 2025; Schneider, 2024). Unlike traditional methods that rely solely on self-assessments or expert ratings, MMLA allows for a more objective and fine-grained examination of the learning process by tracking physiological and behavioral indicators of engagement, attention, and emotions (Worsley et al., 2021; Yan et al., 2024).

For example, MMLA can combine objective measures, like physiological signals, with subjective measures, including self-reports, complementing one another by providing varying levels of granularity and temporality (Järvelä et al., 2021). Furthermore, MMLA can leverage various data channels as sources of convergent validity, for example, to examine alignment between different expression components of emotions (Authors, 2015). This is valuable because of variations between expression components. Behavioral expressions of emotions are culturally framed (von Suchodoletz & Hepach, 2021), whereas, emotional arousal is a psychophysiological response regulated by the autonomic nervous system that cannot be consciously modulated, allowing researchers to have a look into an individual's internal emotional arousal (Cacioppo et al., 2000, pp. 173–191; Harley et al., 2019a,b; Horvers et al., 2021; Liu & Du, 2018; van Dooren et al., 2012; Zhai & Barreto, 2006). Psychophysiological signals are therefore responses that reflect the degree of activation of the autonomic nervous system (ANS) and the excitatory sympathetic (SNS) and inhibitory parasympathetic (PNS) branches secondary to an individual's emotional states.

MMLA can also contribute to predictive algorithms for specific analyses such as using behavioral self- and socially-shared regulated learning codes from videos to predict trainees emotional arousal levels (Authors, 2025). This enables researchers to better identify patterns that may help explain and predict phenomena observed in learning compared to unimodal techniques. For example, Olsen and colleagues (Olsen et al., 2020) found that models predicting students' learning gains were more accurate when using multimodal data from log files, eye-gaze patterns, and audio data than from models using only one of these data channels.

Ecologically valid learning environments allow the findings from MMLA to be more representative of real-world environments as compared to lab-based studies (Cukurova et al., 2020). For example, classroom-based studies have been conducted with students working on collaborative tasks, combining their physiological activation, self-report, and behavioral data to better understand learners' metacognitive processes, such as socially-shared and self-regulation (Järvelä et al., 2023; Sobocinski et al., 2022; Törmänen et al., 2021). For example, Sobocinski and colleagues (Sobocinski et al., 2022) found that learners in collaborative activities adapted their learning processes (e.g., strategies, goals, perceptions) more often when engaging in defining a task than when engaging in the task. Learners also exhibited synchrony in their physiological responses throughout the collaborative activity, indicating that group members experienced a feeling of unity throughout the activity. However, MMLA in ecologically valid settings can be challenging as there can be barriers to using multiple measures such as cost and feasibility (Schneider, 2024) especially in embodied learning settings where learners engage physically with their learning (Cohn et al., 2024), such as medical simulation education.

Medical simulation training provides a realistic, high-fidelity, and ecologically valid setting for examining team-based learning and performance while offering an opportunity to understand trainees' learning processes and factors that may influence their performance, as it closely mirrors real-world clinical environments where trainees must coordinate under pressure, adapt to dynamic scenarios, and regulate their emotions effectively (Gaba, 2007). Unfortunately, research using MMLA in medical simulation education is sparse despite its potential to help uncover the underlying physiological and cognitive processes that shape teamwork, decision-making, and emotion regulation.

Previous MMLA research with the same sample used in this article has indicated that different phases of medical simulations can elicit significant changes to emotional arousal from electrodermal activation (EDA) and the types of regulatory responses that trainees autonomically respond to (Authors, 2024). EDA has been linked to autonomic emotional processes, making it a valuable index for assessing emotional processing (Dindar et al., 2019; Schuurmans et al., 2020)) and is broadly used in research as an indicator of emotional states and physiological responses (Liu & Du, 2018; Wilson et al., 2021). Authors (2024) found that the "beginning" of a medical simulation is more associated with anticipatory EDA responses that require the collection of information and developing familiarity with the case, which in turn drives regulatory responses to optimize the collection and management of new data (Authors, 2024). Likewise, the "escalation" or development of a simulation is where trainees are required to take that information that they have acquired and begin to formulate hypotheses that can be tested to ensure patient survival, showing increasing EDA arousal patterns. Finally, the "peak" of a simulation is where trainees are at their most dynamic and activated, where they must work, often under the threat of patient duress, to regulate their teamwork and interpersonal dynamics to ensure patient survival (Authors, 2024). These patterns of regulation and emotional arousal share a relationship with each other throughout different phases of medical simulation. Additional research extended findings from Authors (2024) by applying supervised machine learning to examine whether medical trainees' different behaviors, self-regulatory and socially shared regulatory actions, and self-reported regulation responses predicted changes in their EDA during three different segments of a medical simulation that they were completing (Authors, 2025). The results of three Lasso ML models determined that this multimodal data could indeed predict changes in emotional arousal. The current study extends our research with emotional arousal by investigating whether (1) medical residents whom were working together had similar or different EDA by examining teams' physiological synchrony and (2) investigating whether teams' physiological synchrony was associated with team leaders' performance.

1.2. Physiological synchrony

Accordingly, the primary aim of this study was to determine whether physiological synchrony (PS), as an indicator of shared emotional arousal, could be detected amongst team members training in ecologically valid, highly realistic medical simulation learning environments, such as Crisis Resource Management (CRM) simulation scenarios. PS refers to the similarities in physiological processes (e.g., EDA, heart rate) among two or more individuals over time. PS has been found to be associated with team cohesion, collaboration quality, and overall performance (Dindar et al., 2022; Tomashin et al., 2022). Preliminary evidence has also been reported where PS is associated with paramedic group cohesion during simulation through LIWC analysis, particularly using first-person plural pronoun use (Misal et al., 2020, October). PS has also been found to be an indicator of successful monitoring of collaboration; seen as the byproduct of shared understanding during goal-setting (Haataja et al., 2018), and may be a possible indicator of when groups are challenged to overcome a shared set of problems (Malmberg et al., 2019). These findings have indicated that PS does occur within collaborative learning. Further research by Yan and colleagues (Yan et al., 2023, pp. 602–614) has explored the heartrate PS of nursing students and discovered that PS predicted stress, group performance and task performance during collaborative work. However, research on PS using EDA within medical trainees or in simulation training remains limited, especially in authentic training environments such as CRM simulations, despite the potential for PS to play a crucial role in team cohesion and performance.

1.3. Physiological synchrony and CRM performance

A secondary aim of this article was to conduct a preliminary investigation of whether PS amongst team members could be a predictor of performance. In the context of CRM scenarios adapted to medical teams, the team leader, typically one physician, is responsible for making key decisions, assigning tasks, managing communication, and setting the tone for collaborative work. The leader's actions directly influence the performance and coordination of the entire team. Therefore, team effectiveness is closely tied to the leader's ability to manage the crisis and leverage team resources effectively. The team leader's CRM performance is therefore integral to the team's functioning and outcomes. Accordingly, we used an expert rating instrument broadly adopted in CRM training and with evidence of validity for assessing team leaders to assign a score to team leaders in our study. Therefore, both clinical and instrument selection considerations drove us to focus our assessment of performance to team leaders.

CRM are a set of skills outside of medical knowledge that can be applied to a variety of crisis environments, including those encountered in medicine (Carne et al., 2012; Kim et al., 2009). Medical emergencies are highly challenging for healthcare professionals who need extensive medical knowledge, and effective self-regulatory and leadership skills. Leadership, communication, problem-solving, resource utilization, and situational awareness are CRM skills that are also paramount in clinical practice.

CRM skill training programs have been developed to promote dynamic decision-making, interpersonal skills, and team management in medical emergency training CRM training utilizes medical simulations with high-fidelity scenarios allowing trainees to practice critical skills in a risk-free environment while developing team-based competencies and interpersonal skills necessary to be a good leader, communicator, and collaborator, essential for healthcare professionals working in multidisciplinary teams (Carne et al., 2012). Given that CRM emergency training programs unfold in real-time with unpredictable outcomes, further research is needed to understand learning processes and team performance measurement in these dynamic settings (Duffy et al., 2015). Additionally, there is a gap in the literature on the role of shared emotions in shaping teamwork and learning within these trainings (Ahn et al., 2023).

1.4. Self- and socially-shared regulated learning

Given the potential consequences, not only for health professionals but for patients' outcomes, medical trainees must develop self- and team regulation strategies to maintain performance under pressure. Further, effective performance in these settings depends not only on clinical knowledge but also on the team's ability to communicate, coordinate, and support one another under emotionally-demanding conditions. Despite this, and although compassion, respect, and collegiality are expected values in a physician, most medical curricula do not explicitly address self- and other-regulation even though regulating emotions is known to influence decision-making, collaboration, and clinical outcomes (McConnell & Eva, 2022). Research on emotion regulation in health professions education, especially in ecologically-valid educational contexts is especially sparse (LeBlanc et al., 2024; LeBlanc & Posner, 2022). We drew on self- and socially-shared regulated learning and achievement emotions literature and theories to guide our study and interpretation of our findings.

Self-regulated learning (SRL) is the process where an individual manages their learning through a process that involves preparatory planning, performance, and appraisal of their learning tasks (Jansen et al., 2019; Zimmerman, 2002). To be an effective self-regulated learner, one must establish goals, engage in metacognitive monitoring of their performance, and have a self-reflection mechanism to be able to make necessary adaptations throughout the learning process (Winne & Hadwin, 2008). SRL can occur within an individual at the cognitive-affective (Ben-Eliyahu & Linnenbrink-Garcia, 2013; Harley et al., 2019b; Li et al., 2021), behavioral (Kazemitabar et al., 2019; Lajoie et al., 2021), and motivational levels (Efklides, 2019; Järvenoja et al., 2018) and can fluctuate over time given the demands and duration of learning.

Specific components of SRL, such as the affective domain, also have their own regulation models, including the Emotion Regulation in Achievement Situations (ERAS) model (Harley et al., 2019a,b; Harley & Pekrun, 2024). ERAS outlines how emotions are generated, when individuals regulate their own or others' emotions during the emotion generation process, and the emotion regulation (ER) strategies that may be used. Optimal emotional states can enhance learning and performance (LeBlanc & Posner, 2022; Madsgaard et al., 2022), making it valuable for learners to manage their own emotions as part of their SRL skills.

Beyond the need to manage and regulate one's own learning and emotions, team members need to be able to collaborate and manage the regulation of members in their team (Hadwin, Järvelä, & Miller, 2017). To do this, socially shared regulated learning (SSRL) is needed to collaboratively regulate learning among multiple individuals (Hadwin & Oshige, 2011; Järvelä & Hadwin, 2013; Miller & Hadwin, 2015; Hadwin et al., 2017). This requires the team to negotiate personal interests, set common goals, and manage the team as members progress in accomplishing these goals. Specifically, team members need to have a baseline understanding of their individual regulatory abilities and the skills to meaningfully contribute to the team's goals (Isohätälä et al., 2018, 2020). Mastering SSRL requires prerequisite SRL abilities, which are then further developed in team-based medical practice.

SSRL theory and research is helpful to highlight the importance of group members and group processes, like shared monitoring, in learning. While we did not measure the latter in this article, the concept of shared monitoring and other SSRL processes help us interpret and situate our findings. Most importantly, SSRL contributes the proposition that shared emotional responses can be indicators of adaptive learning (Dindar et al., 2019; Haataja et al., 2018). This proposition is important to our hypothesis and interpretation of the findings.

Table 1
Participants' demographics.

		N	*M* ± *SD*	*Min-Max*
Age		24	28.5 ± 3	24–33
	Missing	5		
Gender	Female	16		
	Male	13		
Ethnicity	Caucasian	7		
	Visible Minority	19		
	Missing	3		
Program	Internal Medicine	14		
	Emergency Medicine	9		
	Anesthesia	3		
	Critical Care	3		
Training Level (PGY)	Residency year	27	2.27 ± 1.62	1–6
	Missing	2		
Role	Leader	9		
	Exam	9		
	Assisting	11		

Note. Role refers to the role assigned to the participants during the simulations. Exam refers to the participant assigned to perform the physical examination or any other physical intervention on the patient. Assisting refers to the participant assigned to perform activities other than the physical exam or interventions, such as helping with the procedures, making phone calls, and gathering information. These roles were categorized by the research team based on the activities that the team leader assigned the team members to do. Occasionally, some team members temporarily changed roles, from examining to assisting or vice versa. Regardless, the role in which the participant spent most of their time during the scenario was their assigned role prior to the simulation.

1.5. Emotions

Given the emphasis on emotional activation in our research questions we drew on the Control-Value Theory of Achievement Emotions (CVT; Pekrun, 2006; Pekrun, 2024) as an additional theoretical framework for our study. CVT is an integrated theory that provides several important propositions for understanding and measuring EDA.

First, CVT, like Russel's circumplex model (Barrett & Russell, 1998), proposes that emotional arousal (i.e., degree of activation) is a key dimension of emotions (along with valence and object foci). This is important for EDA as this physiological signal provides information about emotional arousal. Enjoyment, anxiety, and frustration are examples of high arousal discrete emotions while boredom and relief are examples of low arousal emotions. This brings us to another dimension Russel and Pekrun agree upon: valence. Valence is the degree to which something is pleasant or unpleasant. Typically, positively-valanced discrete emotions like joy are pleasant to experience whereas anxiety and boredom are not. Pekrun (2024) notes that achievement emotions can be grouped according to valence and physiological arousal. The third taxonomical dimension of emotions in CVT is object focus which refers to (1) whether the object of the emotion is an outcome or an activity and (2) the temporal relations between the object and the person: whether the object is occurring in the future, present, or past. Theories are helpful not just in guiding us to interpret what we are measuring, but also in drawing our attention to what we are *not* measuring and, accordingly, need to consider when interpreting our results. In our case, we do not explicitly measure valence nor do we ask participants what the object of their emotional arousal is. Accordingly, we need to exercise caution making claims along these emotional dimensions, including what discrete emotions medical residents may be experiencing.

Second, emotions are understood to be multicomponential, meaning they have multiple expression components, including experiential (typically measured through self-report, interview, or think-aloud), behavioral (typically measured through facial expressions and posture), and physiological, including EDA, as per our study. Pekrun (2024) advocates for the use of multiple channels for measuring emotions (see Pekrun, 2023 for proposed guidelines), especially physiological arousal and behavioral components that self-report is not well suited to assess. This proposition compliments physiological arousal being recognized as a core emotional dimension and helps theoretically justify our interpretation of EDA as a component of emotion.

Aside from providing a theoretical anchor for our interpretation of EDA data and providing guidance on what dimensions are versus are not captured about emotions, CVT can be used to make theory-informed inferences about the emotional states of medical residents by drawing on what we know about their context and how we can expect that to influence the antecedents of their emotions. CVT describes antecedents of emotions, including the learner's perceived control over the learning situation and the value the learner places on the learning situation (Pekrun, 2024). CVT proposes that learners who appraise a situation as being higher in value are more likely to have a higher level of emotional arousal. This is a useful proposition for us to interpret our findings because of our specialized population and learning context: senior medical trainees, residents, who are performing crisis resource management simulations to practice handling complex patient cases they need to be able to safely and successfully manage to save lives. Therefore, while measuring discrete emotions is beyond the scope of this article which uses physiology, CVT helps us interpret our EDA findings with information we have about their achievement situation (CRM simulation). There are two relevant theory and context-informed hypotheses we can draw about residents' discrete emotional states during the simulation. First, that our medical residents likely have high appraisals of value and that this appraisal pattern arouses emotions like anxiety, especially when appraisals of control are uncertain, which is also likely given their junior medical status.

Another important context-driven hypothesis is that all team members share at least some important goals, such as saving the patient. Therefore, deterioration of the patient or improvements to the patients' condition in the simulation should lead to parallel changes in the emotions of team members that may further be shared through emotional arousal (most often, probably anxiety) becoming visible on team members' faces or in their tone of voices. Therefore, we anticipate physiological synchrony through shared changes in team's own appraisals linked to changes in the simulated patients' status. For example team members' appraisals of control are expected to go down when patient declines. More specifically, uncertainty of the patient surviving can rise and appraisals of one's own and one's team having the capacity to save the patient may decline. We also anticipate that physiological synchrony could arise from appraisals about the team's ability to save the patient from attending to emotional information from team members. For example, a team member may lose confidence (lower appraisal of control) from observing a team member express anxiety through visible perspiration.

1.6. Research objectives

The primary aim of this study was to determine whether PS, as an index of shared emotional arousal, could be detected amongst medical learners (residents) training in ecologically valid, team-based settings using MMLA. Furthermore, this study sought to conduct a preliminary investigation of whether the presence of PS amongst team members was a predictor of the team leaders' performance. The following research questions were examined:

Research Question 1. Can PS be detected before or during a CRM training scenario? *Hypothesis:* PS will be detected amongst medical residents during CRM simulated training as they attempt to collaboratively stabilize and save the life of a simulation patient. We anticipate detecting PS not only because it has been observed in prior literature with different learner populations (Järvelä et al., 2023; Yan et al., 2023, pp. 602–614), but because of authentic and therefore hypothesized shared goals and high-value (appraisal) nature of the task.

Research Question 2. Is the PS level before or during the simulation scenarios associated with the performance of the team leader? *Hypothesis:* The level of PS will positively correlate with the team leaders'

Table 2
Crisis Resource Management (CRM) training scenarios descriptions and roles by specialty and team.

Team	Specialty	Roles Filled	Case Description
1	**Anesth.**	Lead Exam Lead Asst Exam Asst	*A hospitalized pregnant woman, treated for high blood pressure during her pregnancy, presents sudden shortness of breath, chest pain, and rapid clinical deterioration. The first team, (OBGYN team) responds the nurse call. The leader must gather information from the patient and nurse. The patient then presents cardiac arrest, so the leader needs to activate the Code Blue protocol, which is the cue for the Anesthesiology team to enter the simulation and lead the resuscitation of the patient, while the OBGYN team performs an urgent C-section simultaneously. Both teams need to work together to maintain, both mother and fetus alive, and decide further treatment and/or interventions.
2	**Emerg. Med.**	Lead Exam Lead Asst Exam Asst	A young female is brought to the Emergency Department (ED) as trauma activation case after being bucked off her horse and presenting head and pelvic trauma. Resident Leader needs to effectively lead the trauma team during the complex scenario delegating tasks, to recognize the need to intubate and the brain herniation, and timely consult the proper specialties for definite management.
3	**Internal Med.**	Lead Exam Lead Asst Lead. Asst_2 Asst. Asst_2 Asst. Exam Asst_2. Exam	Hemodialysis patient, admitted for gastrointestinal bleed, has a witnessed fall in the hospital, and does not feel well since. Patient's potassium is high and besides their hemodialysis catheter, there are not other intravenous lines to deliver treatment. Resident Leader and team need to detect the hyperkalemia (High potassium blood levels) as a medical emergency promptly, start medical treatment, and call the medical teams that can provide a solution to the catheter issue to promptly deliver medications and prevent further clinical deterioration of the patient
4	**Internal Med.**	Lead Exam Lead Asst Exam Asst	Residents are called to a room to perform CPR on a hospitalized patient. The nursing team has already started the protocol, and the residents need to take over the resuscitation efforts and diagnose the etiology of the cardiac arrest to treat it.
5	**Anesth.**	Lead Exam Lead Asst Exam Asst	Residents are called to ED to care for a child brought after drowning for several minutes. The team needs to carry the resuscitation, gather information from the patient's relatives, and progress in the treatment sequence to minimize potential negative effects of sustained hypoxia on the patient's brain.
6	**Emerg. Med.**	Lead Exam Lead Asst Exam Asst	An elderly man presents to the ED with palpitations and malaise. The patient had an LVAD (Left ventricular assist device) placed within the last month, which makes complicated to obtain vital signs from the patient. The Resident Leader and team need to find a way to assess the patient's vitals. The patient LVAD line site has infection signs, and electrocardiographic signs of hyperkalemia. The patient progress to cardiac arrest and the team needs to resuscitate the patient and manage the causes of it.
7	**Internal Med.****	Lead Exam Lead Asst Lead. Asst_2 Asst. Asst_2 Asst. Exam Asst_2. Exam	**A young pregnant woman, admitted for hypertension (high blood pressure) of pregnancy. The team is called because the patient complains of chest pain and shortness of breath and is found to have hypotension (low blood pressure) on exam, and neurological deterioration. The fetus also presents signs of suffering. The patient shows progressive clinical deterioration. The Resident Leader and team need to identify the possible diagnosis and start treatment. They need to be able to progress in the complexity of the treatment sequence for the dyad mother-fetus.
8	**Emerg.Med. ****	Lead Exam Lead Asst Exam Asst	
9	**Internal Med.**	Lead Exam Lead Asst Exam Asst	See description for Team 4.

Notes. Anesth. = Anesthesiology. Emerg. Med. = Emergency Medicine. Medical Resident roles were based on the four roles or activities performed by the residents during the simulations: Leader (*Lead)*: the appointed leader of the team and focus of the CRM skills leader performance score; Examinator (*Exam)*: the participant commissioned by the leader to perform the physical examination or any procedure on the patient; Assistant 1 and Assistant 2 (*Ass1* and *Ass2*): Participants assigned to other activities such as gathering information, making phone calls, assisting procedures, etc.

performance because shared emotional responses amongst a team can be an indicator of adaptive learning and performance (Dindar et al., 2019; Haataja et al., 2018 Mayo et al., 2021; Misal et al., 2020, October).

2. Methodology

Institutional Review Board ethics approval was obtained prior to the beginning of the study. Before collecting any data, all participants were informed about the study and their right to freely withdraw consent at any time during the data collection. Consent was obtained via signature and written approval.

2.1. Participants

Medical residents (N=29) belonging to different residency programs participated in this study (Table 1).

2.2. Study design

This was a field-based, cross-sectional observational study. We collected data from medical residents while they took part in CRM simulated training sessions from March 2022 to January 2023 at a North American hospital-affiliated simulation centre. The residents were divided into groups and each group had an appointed leader. In total, data was collected from 9 groups. The CRM simulation training was part of residents' curriculum. Accordingly, the research team did not influence the scheduling of the scenarios, selection of the participants, teams, or role assignments, or the content or duration of the simulations.

Data collection sessions consisted of a member of our team obtaining informed consent from the residents. As residents were preparing to enter the simulation room, psychophysiological wristbands (Empatica E4) were placed on their wrists to capture their EDA signals, and they filled out a pre-simulation survey. Participants then took part in the simulation. After the simulation, residents filled out a post-simulation survey before beginning a debriefing session. All participants consented to their inclusion on the study.

It is important to note that the duration of the pre-simulation data collection window was variable across participating teams. This variability was due to the real-world nature of the training environment, where the research team had no control over the scheduling, timing, or organization of the simulation scenarios. As such, the research team leveraged the unique opportunity for data collection within these ecologically valid conditions, which inherently limited the consistency of the pre-simulation recordings.

2.3. CRM simulation training

The CRM training sessions consisted of nine high-fidelity (i.e., high degree of realism) medical simulation scenarios. The simulations employed high-fidelity manikins to serve as patients, mimicking physiological activity (i.e., breath sounds, heart sounds, blinking, etc.) and interacting with the residents through speakers controlled by simulation educators. Medical residents were divided into 9 teams: 7 groups of 3 participants and 2 groups of 4 participants. A leader was appointed for each team by the simulation facilitator. The research team had no control over the scheduling, selection of the medical trainees, or design of the simulations. Group size was determined by logistical factors outside the researchers' control, specifically based on the residency program cohort sizes and institutional educational simulation scheduling constraints. All simulations were audio and video recorded to synchronize with participants' physiological data, provide contextual qualitative descriptions of each scenario, and enable performance coding. Scenarios lasted between 10 and 20 min. See Table 2 for a description of the CRM healthcare simulation cases medical resident teams were assigned by their educational coordinator (not the research team). Cases and team size varied by medical resident program. CRM simulations reflect skills and scenarios medical residents in different specializations (e.g., internal medicine, emergency medicine) need to practice providing life-saving patient care. While the learning objectives varied across scenarios, the overarching aim of the simulation training was the practice and consolidation of CRM skills. These included the use of a systematic approach to clinical assessment, the prompt recognition and anticipation of clinical challenges, and the implementation of appropriate interventions. Medical knowledge was important within the context of each clinical case, but was adapted to the participants' training level and played a secondary role relative to the CRM learning objectives. This design principle allowed for the assessment of team performance independently of the participants' postgraduate year, as CRM skills can be developed by trainees across all levels and medical programs.

2.4. Measures

Ottawa Global Rating Scale (GRS): This scale assesses team leaders' CRM performance by dividing CRM skills into five categories: problem solving, situational awareness, leadership, resource utilization, and communication. Each category is measured on a 7-point anchored ordinal scale with a description of every point to avoid personal bias. A score of 1 corresponds to the performance of a novice. A score of 3 corresponds to the performance of a novice with some CRM and resuscitation experience. A score of 5 corresponds to the performance of a physician with sufficient CRM and resuscitation experience to manage critical events competently, and a score of 7 corresponds to the performance of an experienced physician in CRM and resuscitation.

Validity evidence for the Ottawa Global Rating Scale (GRS) has been established across medical emergency contexts, residency levels, and specialties (Kim et al., 2009). Given that the CRM training emphasized leadership and communication skills, the research team adapted the Ottawa GRS by removing categories not directly aligned with these learning objectives (overall performance, problem solving, situational awareness, and resource utilization). This decision was supported by Kane's validity framework (Tavares et al., 2018), which allows for the adaptation of assessment tools to suit specific evaluation contexts without undermining their validity. As the scale was not being used for summative clinical evaluation of learners, omitting components that were not essential to the research objectives allowed for greater focus and efficiency during scoring.

Every category and item on the modified CRM score were assessed by two experienced Emergency Medicine physicians, members of the research team. In the first iteration, both physicians scored each video individually, focusing on the team leader's performance, guided by learning objectives of each scenario. After the first iteration, both raters reviewed the videos and their ratings, and negotiated the differences in scores. Cohen's Kappa was 0.77 before negotiation, representing moderate to strong agreement with perfect agreement reached after deliberation. Once the scoring was completed, all the item scores in every category of both *Leadership* and *Communication* were averaged to get one global CRM score.

2.5. Collection and pre-processing of the physiological data

Participants' EDA was collected using *Empatica E4* wristbands both before and during the simulation scenarios. The Empatica E4 wristband provides four data points per second (sampling rate: 4 Hz) measured in micro-Siemens (uS). Each participant included in this article wore an Empatica E4 wristband and consented to have their EDA data collected and analyzed. The Empatica E4 is widely used to measure participants' physiological signals and has been tested against gold-standard devices, highlighting the reliability of the E4 wristband as a non-obtrusive tool to measure psychophysiological signals in research (Hu et al., 2024; Schuurmans et al., 2020; Wilson et al., 2021). Given the robustness of Empatica E4 devices in research contexts, its use was ideal for capturing EDA in high-intensity scenarios like medical emergencies.

Importantly, we did not implement specific *in situ* controls for movement artifacts as providing directions to trainees during simulations could compromise the immersive nature of their training or their ability to perform technical skills (e.g., chest compressions). However, additional steps were taken during *post-hoc* data preprocessing to address potential issues related to signal quality and artifacts that may have been missed by automatic approaches such as the application of signal processing filters without prior visual inspection, which may risk either overlooking motion artifacts or inadvertently distorting useable data. (e.g., via MATLAB and Ledalab). This allowed us to improve the accuracy of the subsequent analyses. All EDA recordings were visually inspected by the researchers in the Empatica platform to explore the integrity of the signals, identify artifacts (including those relating to movement) or data loss.

EDA has two components: The tonic, which represents slow changes in conductivity, expressed as skin conductance level (SCL), and the phasic, which denotes the changes of conductivity from the baseline with rapid responses, known as skin conductance response (SCR). For this study, we analyzed the entire EDA signal rather than deconvolving it into its components (SCL, SCR). Synchrony focuses on the similarity of patterns or recurrence of the EDA signal rather than the intensity of the signal itself, hence, we could potentially observe little activation and yet, observe synchronization among individuals (Wallot et al., 2016a,b).

After inspection, EDA recordings were synchronized with the times in the video recordings. Once the simulation window was defined, a 3-min window immediately preceding the scenario was extracted and used as the baseline state of each participant, acknowledging the previously noted variability in pre-simulation recording lengths due to real-world logistical constraints (Tomashin et al., 2022). The baseline period was a short window of time prior to the start of the simulation. Simulation time is expensive and extremely limited, especially when research studies are taking place around a scheduled educational activity. Prior to the simulation medical residents received instructions for the simulation, consenting medical residents completed a brief self-report measure about their emotions and demographics and some medical residents also chatted with one another. The research team did not capture medical resident behavior during this window, therefore some variance in activities during this period is expected. Accordingly, while this was used as a baseline relative to the simulation task it was not possible within this ecologically valid context to have a traditional, lab-based baseline. By this we mean a baseline where participants are assigned a task, such as watching something that is associated with neutral affect.

Finally, all the raw data was transformed into z-scores, using RStudio (version 2022.12.0 + 353, packages *tidyverse, ggplot2,* and *stats4*), then preprocessed using the Ledalab function (version V3.4.9) from MATLAB (version R2023a) by applying a *Butterworth* low-pass filter (0.03) to attenuate high-frequency noise, and manually removing artifacts using the linear interpolation method (Dindar et al., 2020; Kazi et al., 2021; Liu et al., 2021).

To examine the presence of PS, we used the Multidimensional Recurrence Quantification Analysis (MdRQA) (Wallot et al., 2016a; Wallot & Leonardi, 2018b) method, which is an extension of cross-recurrence analysis, which measures the degree of synchrony between two or more time series. MdRQA can handle multivariate time series data, which we are dealing with in this study, and is not limited to bi-variate analyses.

To perform recurrence analysis, the time series data are transformed into a phase space using time-delayed embedding. The time series are then plotted against themselves with a time lag determined by a delay parameter, and with a dimension parameter that determines the number of times the data points are plotted against themselves. It is particularly helpful when dealing with complex and stochastic signals like EDA due to its robustness to non-normally distributed data, and outliers (Wallot et al., 2016a; Wallot & Leonardi, 2018a).

MdRQA produces various output measures that explore various aspects of PS. However, for the purposes of this study, we focused on the recurrence rate (%REC) outcome measure, since it allows one to assess the proportion of recurrent points in the phase space, indicating the presence of synchrony. While other output measures capture system dynamics, our study's main objective was to determine the presence of PS rather than explore finer characteristics. Thus, focusing on %REC aligns with our research goals (Wallot et al., 2016a; Wallot & Leonardi, 2018b).

We conducted the analysis in RStudio (version 2022.12.0 + 353, packages *crqa, entropy, nonlinearTseries, plot3D, SDMTools, tserieschaos, matrix)* as per Wallot et al.'s guidelines (Wallot et al., 2016a; Wallot & Leonardi, 2018b). We aimed to capture teams' PS levels. To conduct this analysis, we applied the following parameters: 1) Groups: *emb:* 1, *del* = 21, *rad*=0.2, *norm:* Euclidean. Additionally, the *radius (rad)* parameter was defined running iterations of the MdRQA with different threshold values until obtaining %REC outputs ranging from 5 to 20 %, as suggested for stochastic systems, and the same value was applied throughout all the data analysis. This was done to procure the comparability of the MdRQA results across data sets (Wallot & Leonardi, 2018a).

Finally, to answer RQ1, we generated surrogate data sets to compare against real data sets. The surrogate data sets were false data matrices created by randomizing the EDA z-scored values of each participant within a group. Using the MdRQA function with the same parameters as the real data sets, we obtained surrogate %REC values for both the pre-simulation and simulation windows. This process was applied to all teams (see Supplemental Materials B). These surrogate data sets served as reference values to assess synchrony levels in the real data sets, as they were expected to lack significant synchrony or %REC due to their randomized nature.

Although creating surrogate datasets with non-group members (e.g., randomized dyads across teams) is a strategy used in related work (Ahonen et al., 2018), this was not feasible with our data due to differences in recording lengths and sample size limitations. Truncating simulation data to enforce standardization would have led to excessive data loss and diminished the ecological validity of the training sessions. Our approach was shaped by these practical constraints within a quasi-naturalistic simulation environment. This method allowed us to confirm that no meaningful synchrony was present in the surrogate data, validating that synchrony detected in the real data exceeded chance levels.

2.6. Analysis

Data were analyzed using IBM SPSS 29. Normality was assessed with the *Shapiro Wilk*'s test, and some of the variables were not normally distributed. Significance was defined at the $p < .05$ level. A description of the study variables' definitions and their descriptive statistics can be found in Table 3.

Using GPower 3.1, we calculated the number of groups needed to meet statistical power. A retrospective power analysis indicated that a sample size of 9 groups, with an effect size of 0.03 (dz; computed using GPower 3.1 mean differences calculator), an error probability of 0.05, produced a nominal power of 0.05 for the Wilcoxon signed rank test (matched pairs) with a normal parent distribution. This test was used to compare surrogate %REC and real data %REC values of teams (RQ1). Next, another retrospective power analysis was conducted for RQ2 which examined correlations between %REC levels during PreSim and Sim windows and the CRM scores. As GPower 3.1 does not support power calculations for Kendall's tau correlations, a two-tailed bivariate normal model (Pearson correlation) was used as a substitute to estimate power. This indicated that with a sample size of 9 groups, an error probability of 0.05, and a ρH1 of 0.54 (obtained via r statistic), that the power was 0.34. Power calculations for both research questions 1 and 2 were therefore underpowered (see Supplementary Material B).

Table 3
Descriptive statistics for all study variables.

	N		*M*	*SD*	Range	*Min*	*Max*
	Valid	Missing					
1. Surrogate PreSim%REC	9	0	0.69	0.25	0.6	0.25	0.85
2. PreSim %REC	9	0	7.23	3.61	10.66	3.6	14.26
3. Surrogate Sim%REC	9	0	0.71	0.27	0.62	0.23	0.85
4. Sim%REC	9	0	6.72	2.66	6.87	3.92	10.79
5. CRM Score	9	0	6.24	0.78	2.36	4.51	6.87

1. Surrogate PreSim %REC: Percentage of recurrence resulted from MdRQA of surrogate data sets during the pre-simulation (PreSim) window.
2. PreSim %REC: Percentage of recurrence resulted from MdRQA of participants' real data during PreSim window.
3. Surrogate Sim %REC: Percentage of recurrence from surrogate data sets of the simulation (Sim) window.
4. Sim %REC: Percentage of recurrence from participants' real data of the Sim window.
5. CRM Score: Team's performance score on a Likert scale from 1 to 7 (1 = novice, 7 = clearly superior).

Table 4
Difference between surrogate data and real data during PreSim and During-Sim in group level.

Wilcoxon Signed Ranks Test Statistics		
	PreSim %REC - Surrogate PreSim %REC	Sim%REC- Surrogate Sim %REC
Z	−2.666[a]	−2.666[a]
Asymp. Sig. (2-tailed)	0.008	0.008

1. PresSim %REC and Sim %REC refers to Pre-simulation and During simulation real data.
2. Surrogate PreSim %REC and Surrogate Sim %REC refers to PreSimulation and During Simulation surrogate data, respectively.
[a] Based on negative ranks.

3. Results

3.1. Research question 1: is there PS before or during the simulation scenarios?

After running the MdRQA, the presence of PS was found across all teams in both, pre-simulation, and simulation windows. Therefore, to determine if PS levels were significant, we conducted a Wilcoxon signed-ranked test comparing the %REC levels of the pre-simulation and simulation windows, between the surrogate datasets of each team, and each teams' real datasets (Table 4).

The Wilcoxon signed-rank test result of the pre-simulation window analysis revealed that there was a statistically significant difference between surrogate %REC and real data %REC PreSim values of all teams, with the real teams exhibiting higher %REC than surrogate data (Surrogate %REC and real %REC: $W = -2.66$, $p = .008$, $r = 0.89$: large effect).

Likewise, the Wilcoxon signed-rank test result of the simulation window analysis revealed that there was a statistically significant difference between surrogate and real data %REC Sim values of all teams, with the real teams exhibiting higher %REC than surrogate data (Surrogate %REC and real %REC: $W = -2.66$, $p = .008$, $r = 0.89$: large effect). Hence, we can conclude that there was physiological synchrony among team members before and during the simulation scenarios. Supplemental paired sample t-tests reported in Supplemental Material A replicated these results.

Table 5
Kendall's tau correlations between study variables.

Correlations					
		1	2	3	4
Kendall's tau_b	1. Surrogate PreSim %REC	–			
	2. PreSim%REC	−0.056	–		
	3. Surrogate Sim%REC	0.444	0.278	–	
	4. Sim%REC	−0.111	0.611[a]	0.333	–
	5. CRM Score	−0.085	0.31	0.254	0.535[a]

1. Surrogate PreSim %REC: Percentage of recurrence resulted from MdRQA of surrogate data sets during the pre-simulation (PreSim) window.
2. PreSim %REC: Percentage of recurrence resulted from MdRQA of participants' real data during PreSim window.
3. Surrogate Sim %REC: Percentage of recurrence from surrogate data sets of the simulation (Sim) window.
4. Sim %REC: Percentage of recurrence from participants' real data of the Sim window. 5. CRM Score: Team's performance score on a Likert scale from 1 to 7 (1 = novice, 7 = clearly superior).
5. CRM Score: Team's performance score on a Likert scale from 1 to 7 (1 = novice, 7 = clearly superior).
[a] Correlation is significant at the 0.05 level (2-tailed).

3.2. Research question 2. does PS correlate with the overall performance of the team?

To answer our second question, we ran Kendall's tau correlation to determine the relationship between %REC levels during PreSim and Sim windows and the CRM scores. We found a statistically significant, moderate, positive correlation between During Sim real data %REC and CRM score ($\tau_b = 0.535$, $p = .046$). We conducted assumption tests to determine if a predictive test could be done, however, four of the assumptions to conduct regressions were not met, thus no regressions were conducted (Table 5).

4. Discussion

This study aimed to investigate PS among medical residents during ecologically valid, team-based CRM simulated training and its potential correlation with team leaders' performance. Teams leaders' performance was assessed using the CRM score scale, which assigned scores based on the leadership and communication skills exhibited by the teams' leader during the training session.

The key finding of this study is the successful application of the MdRQA technique to medical education in an ecologically valid medical training setting and the empirical evidence of PS among medical residents both before and during CRM training. This finding provides novel evidence that medical residents' experience shared emotional arousal fluctuations before and during simulation training. This finding replicates similar empirical findings in different populations and settings, strengthening the generalization of PS in teams (Yan et al., 2023, pp. 602–614). Educationally and clinically, these finding have practical implications: if emotional arousal fluctuations are shared, then it is imperative that individuals training for high-stakes procedures monitor and consider taking steps to adaptively manage their own but also help support their team members' emotions. For example, team leaders might reflect that if they are feeling anxious, members of their team may very well be too. Knowing that such states have the potential to get in the way of successful patient care (simulated training or real patient) they might take a moment to down-regulate a climbing negative, arousing state like anxiety or frustration within themselves and perhaps others before continuing. Our findings about PS prior to the start of the simulation offer evidence that considerations around self- and team emotion monitoring and management may apply not only to the simulation itself, but prior to its start when medical residents may be thinking about it, even if they are engaged in other tasks.

We believe that our PS findings may be attributed to the nature of the simulation scenarios, which involved complex, challenging medical cases medical trainees would want to (high appraisal of value) but likely find challenging to (uncertain appraisal of control) effectively manage (Pekrun, 2024). Environments such as medical simulations are particularly likely to elicit strong arousal due to them mirroring situations medical trainees could find themselves in shortly. We hypothesized, based on our knowledge of the educational context, medicine, and simulation training, this leads to high appraisals of value (LeBlanc, 2019). Our medical simulation context and medical resident population distinguishes our study from prior PS research conducted in lower-stakes educational contexts (Yan et al., 2023, pp. 602–614). In our case, the immersive and emotionally charged nature of the CRM training may have contributed to increased PS among team members. These findings contribute to supporting CVT propositions, albeit with inferred rather than measured appraisals, in an under-studied context, previously unexamined population, and through collective team fluctuations in emotional arousal. The latter is especially valuable as most evidence for CVT propositions stems from individual and often self-reported accounts of emotion. Our context is fertile ground for future investigations of shared emotional arousal and its role in team dynamics, including communication, coordination, and shared understanding (Pekrun, 2024).

Our secondary and preliminary finding is that statistically significant correlations were observed between team leaders' performance scores and PS at the group level. This finding echoes prior research associating PS positively with performance in various domains (Dindar et al., 2022; Tomashin, 2022). Notably, some prior studies in education have reported inconsistent results, linking PS to relationship quality, monitoring activities, and perceived task difficulty during learning activities rather than performance itself (Haataja et al., 2018; Järvelä et al., 2021; Mayo et al., 2021; Wallot et al., 2016a,b). While this finding should be interpreted with caution due to low power and a small sample it supports SSRL theory and prior findings that experiencing PS can be a good thing for learning and performance outcomes.

Three pathways could potentially help explain the direction of our correlational findings between PS and leadership: *Leader-driven synchrony* where a high-performing leader creates a more emotionally synchronized team (Leader Performance → PS). Second, *Synchrony-enabled leadership* (PS → Leader Performance) where a team that is already emotionally in sync creates an environment that makes it easier for the leader to perform well. *Third-variable effects*: another factor, such as the difficulty of the scenario, simultaneously driving both high PS and impacting the leader's performance score. While future studies will be needed to provide empirical evidence of which pathways are most influential the literature would suggest and we hypothesize that all three pathways likely play a role to some extent.

Concerning Leader-driven synchrony, attending to the emotions of others is widely recognized in leadership literature, including clinical literature, as a key ingredient (Stoller, 2021). Indeed, an effective leader can remain calm, maintain a global perspective, anticipate crisis early on, delegate tasks strategically, sets task priorities, is open to suggestions, and encourage input from team members (Lei & Palm, 2025). Therefore, we hypothesize that high-performing leaders are more likely to be able to create an emotionally synchronized team than lower-performing ones.

Using CVT we could hypothesize that PS can reflect all team members focusing on a common object foci and with similar appraisals. Having everyone focused on the simulation activity rather than thinking about the outcomes of the simulation activity and having similar, realistic (e.g., not too over or under-confident) appraisals of control could make it easier for the leader to anticipate and respond to how others are feeling (synchrony-enabled leadership).

If a simulated patient presents with symptoms that are severe and complex to treat (third-order variable effects) than a high-performing team leader may anticipate and respond to both expressed and suppressed anxiety from team members. In doing so, they may attempt to make the whole team, and themselves in turn, feel better by acknowledging the difficulty of the case, but reminding the team that they all know the procedure to follow. This example illustrates that these three pathways are likely not mutually exclusive or one-directional.

We further hypothesize that a good leader is likely to better navigate a situation where there is low PS than a poor leader is to benefit from high PS in their team. A good leader is more likely to recognize and effectively respond to the emotions—shared or not—of their team. A poor leader may not recognize shared emotions in others and this may not only be a missed opportunity but potentially create worse outcomes. For example, they might fail to recognize and respond to their own and others' anxiety, especially if that anxiety is shared. A good leader, on the other hand, might recognize that a single team member appears to be anxious where the others are calm (or the other way around) and react accordingly. In other words: we hypothesize that a good leader is more likely to be able to (1) leverage PS when high, (2) to create a more emotionally synchronized team when PS is low, and (3) mitigate potentially undesirable third-variable effects than a weak leader. We also hypothesize that team members' abilities to communicate, collaborate, and support one another and the leader, instrumentally and emotionally, will influence both the emotionally synchrony of the team as well as performance.

5. Limitations and future directions

While this study offers novel insights into physiological synchrony during simulation-based medical training, several limitations and future directions should be noted to contextualize the findings and guide future research.

First, a larger sample size is required to increase confidence in our preliminary findings for research question 2, which should therefore be interpreted with caution. Robust replication of this finding will be important because underpowered analyses risk both failing to detect true effects (false negatives) and reducing the reliability of statistically significant findings (Button et al., 2013). That said, small sample sizes are a common limitation in studies involving highly specialized populations such as medical residents, who are limited in number and have complex scheduling constraints. For example, a study of health sciences education literature reported that the median sample size of 627 studies was of N = 25 for non-intervention studies (Cook & Hatala, 2015), comparable to our sample of N = 29. Additionally, our observational design limited our recruitment efforts to only the residents who were required to undergo CRM simulation training as part of their curriculum and learning objectives. Moreover, studies involving multimodal physiological data face high risk of data loss across channels, further impacting sample size. Future work should aim to robustly replicate our findings with larger samples to enhance statistical power, confidence, and generalizability. This work could also help test the proposed causal pathway between PS and leaders' performance.

Although CRM skills are designed to be independent of knowledge level or postgraduate year, a larger sample would allow future researchers to explore whether PS-performance associations differ by specialty or training stage. Our current group-level sample size limited the ability to run regression analyses, though the significant correlations observed remain promising, even if they should be interpreted with caution on account of sample size.

Another limitation is that we only measured the performance of the team leader, reflecting the design and learning objectives of the CRM simulations. While leaders' leadership and communication are essential to team outcomes and our preliminary findings are highly novel, future research should both incorporate additional CRM dimensions as well as other robust performance metrics that include assessments for all team members (e.g., Cooper et al., 2010; TEAM) to better capture the complexity of collaborative performance. In other words, future research must ask how we can turn a team of experts into an expert team, which entails capturing individual member competence, followership, emergent team processes, and the collective execution of tasks, (Salas et al., 2006). Future research should also examine other team constructs, such as team climate, group cohesion, and other measures to assess shared emotions, including shared discrete emotions.

We acknowledge that not decomposing the EDA channel can risk having the tonic (SCL) component to drift, for example on account of temperature changes or hydration. That said, we expect drifting to be limited in our study given that simulation sessions ranged from as short as 10-min to a maximum length of 20-min in a climate-controlled environment. Future PS analyses, especially during longer sessions and in less controlled environments, may be better served by decomposing the EDA signal. Another limitation was that the percentage of artifacts removed or interpolated data was not tracked and this limits the interpretability and replicability of our current findings. Future research should track and report this information.

We also acknowledge that differences in simulation session time could have influenced the extent of PS in groups with longer or short sessions. For example, teams that worked for longer together may have had more time to observe and respond (engage in SSRL) with others' emotions which research has shown is associated with PS. Future research, especially conducted in more controlled and less authentic settings, would be well positioned to examine the potential influence time might have on teams' PS.

Of our 9 teams, 7 had 3 members while 2 had 7 members. A visual inspection of Supplemental Material B reveals that the percentage of recurrence or synchrony did not seem to vary more between teams 3 and 7 and teams with fewer members for either the simulation or pre-simulation window (see %REC[4] columns). More systematic comparisons of the potential role of team size on PS is a worthwhile future direction, especially for research in more controlled and less ecologically-valid contexts. We also recommend that future studies examine the potential impact of both different postgraduate levels and prior acquaintance with team members on PS as well as team leader performance.

We recommend that future studies integrate self-report data and other multimodal indicators which could offer empirical evidence of validity between synchrony and outcomes like learning, well-being, and team functioning. The use of complementary frameworks, such as the Framework for Regulation of Emotions in Collaborative Learning (Lobczowski, 2022), could support this line of inquiry and provide additional theoretical grounding.

Future studies could use an MdRQA derived analyses to examine PS within teams, such as cross-recurrence quantification analysis (CRQA) and diagonal cross-recurrence profile (DCRP), which focus on bivariate dynamics could shed light on co-regulated (DCRP: leader-follower interactions) and shared-regulated (CRQA: interaction between peers others than leaders) learning mechanisms (Wallot & Leonardi, 2018b). Findings from such work may reflect underlying team cohesion, though cohesion was not directly measured, contrasting previous findings from Yan and colleagues (Yan et al., 2023, pp. 602–614) which utilized heart rate to model PS, highlighting the importance of using various physiology measures to determine if PS exists. Future research could explore this connection further.

Finally, while we did not employ surrogate datasets composed of non-group members, due to recording length variability and sample size constraints, we used within-group randomized surrogate datasets as a pragmatic and ecologically valid alternative. These datasets showed no meaningful synchrony, validating their utility as a baseline for detecting real PS. Although initially adopted to address analytic constraints, we propose this approach as a scalable strategy for future research conducted in naturalistic educational environments, where strict time-series standardization across participants may not be feasible.

6. Conclusions

In conclusion, this study provides novel insights into the potential of using PS to identify shared emotional arousal among medical teams during high-fidelity simulation training. We also found preliminary evidence of a correlational relationship between PS and team leader performance. Notably, this is the first study to apply Multidimensional Recurrence Quantification Analysis (MdRQA) to examine PS in this context, demonstrating the feasibility and utility of this MMLA method in medical education. Within and beyond medical education, this approach offers promising applications for enhancing our understanding of team collaboration, emotional arousal, and performance, including in ecologically-valid environments and as part of multimodal studies.

7. Participant consent statement

Prior to data collection, participants provided informed consent and were informed of their right to freely withdraw consent at any time during the data collection. Consent was obtained via a signature and written approval on a consent form approved by our institutional research ethics board.

CRediT authorship contribution statement

Lucia Patino Melo: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Sebastian Wallot:** Writing – review & editing, Software, Methodology, Formal analysis. **Keerat Grewal:** Writing – review & editing, Software, Project administration, Methodology, Investigation, Data curation. **Matthew Moreno:** Writing – review & editing, Methodology, Investigation, Data curation. **Osamu Nomura:** Writing – review & editing, Formal analysis. **Jason M. Harley:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Data statement

Data is not available due to being outside the scope of IRB approval and reflects a small, highly specialized professional population.

Sources of funding

This work was supported by a Partnership Development Grant from the Social Sciences and Humanities Research Council of Canada [grant number: 890-2020-0060] awarded to the senior author and P.I., J.M. Harley.

Conflict of interest statement

The authors report no conflicts of interest.

Acknowledgments

The authors thank the staff at the Steinberg Centre for Simulation and Interactive Learning (SCSIL), McGill University, and the McGill University Health Centre for Interprofessional Simulation (MUHC-i-SIM) where data was collected. The authors also thank the undergraduate research assistant volunteers, K.H. and K.R, for their help and Sayed Azher for support with power calculations.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.learninstruc.2025.102302.

Data availability

The authors do not have permission to share data.

References

Ahn, B., Maurice-Ventouris, M., Bilgic, E., Yang, A., Lau, C. H.-H., Peters, H., Li, K., Chang-Ou, D., & Harley, J. M. (2023). A scoping review of emotions and related constructs in simulation-based education research articles. *Advances in Simulation, 8*(1), 22. https://doi.org/10.1186/s41077-023-00258-z

Ahonen, L., Cowley, B. U., Hellas, A., & Puolamäki, K. (2018). Biosignals reflect pair-dynamics in collaborative work: EDA and ECG study of pair-programming in a classroom environment. *Scientific Reports, 8*(1), 3138. https://doi.org/10.1038/s41598-018-21518-3

Barrett, L. F., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology, 74*(4), 967–984. https://doi.org/10.1037/0022-3514.74.4.967

Ben-Eliyahu, A., & Linnenbrink-Garcia, L. (2013). Extending self-regulated learning to include self-regulated emotion strategies. *Motivation and Emotion, 37*(3), 558–573. https://doi.org/10.1007/s11031-012-9332-3

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*(5), 365–376. https://doi.org/10.1038/nrn3475

Cacioppo, J., Berntson, G., Larsen, J., Poehlmann, K., & Ito, T. (2000). *The psychophysiology of emotion*.

Carne, B., Kennedy, M., & Gray, T. (2012). Review article: Crisis resource management in emergency medicine. *Emerg Med Australas, 24*(1), 7–13. https://doi.org/10.1111/j.1742-6723.2011.01495.x

Cohn, C., Davalos Anaya, E., Vatral, C., Fonteles, J., Wang, H., & Ma, M. (2024). Multimodal methods for analyzing learning and training environments: A systematic literature review. https://doi.org/10.48550/arXiv.2408.14491.

Cohn, C., Snyder, C., Fonteles, J. H., Ts, A., Montenegro, J., & Biswas, G. (2025). A multimodal approach to support teacher, researcher and AI collaboration in STEM + C learning environments. *British Journal of Educational Technology, 56*(2), 595–620.

Cook, D. A., & Hatala, R. (2015). Got power? A systematic review of sample size adequacy in health professions education research. *Advances in Health Sciences Education: Theory and Practice, 20*(1), 73–83. https://doi.org/10.1007/s10459-014-9509-5

Cooper, S., Cant, R., Porter, J., Sellick, K., Somers, G., Kinsman, L., & Nestel, D. (2010). Rating medical emergency teamwork performance: Development of the Team Emergency Assessment Measure (TEAM). *Resuscitation, 81*(4), 446–452. https://doi.org/10.1016/j.resuscitation.2009.11.027. Epub 2010 Feb 1. PMID: 20117874.

Cukurova, M., Giannakos, M., & Martinez-Maldonado, R. (2020). The promise and challenges of multimodal learning analytics. *British Journal of Educational Technology, 51*(5), 1441–1449. https://doi.org/10.1111/bjet.13015

Dias, R. D., & Scalabrini-Neto, A. (2017). Acute stress in residents playing different roles during emergency simulations: A preliminary study. *International Journal of Medical Education, 8*, 239–243. https://doi.org/10.5116/ijme.5929.60f1

Dindar, M., Järvelä, S., & Haataja, E. (2020). What does physiological synchrony reveal about metacognitive experiences and group performance? *British Journal of Educational Technology, 51*(5), 1577–1596. https://doi.org/10.1111/bjet.12981

Dindar, M., Järvelä, S., Nguyen, A., Haataja, E., & Çini, A. (2022). Detecting shared physiological arousal events in collaborative problem solving. *Contemporary Educational Psychology, 69*, Article 102050. https://doi.org/10.1016/j.cedpsych.2022.102050

Dindar, M., Malmberg, J., Järvelä, S., Haataja, E., & Kirschner, P. A. (2019). Matching self-reports with electrodermal activity data: Investigating temporal changes in self-regulated learning. *Education and Information Technologies, 25*(3), 1785–1802. https://doi.org/10.1007/s10639-019-10059-5

Duffy, M., Azevedo, R., Sun, N.-Z., Griscom, S., Stead, V., Crelinsten, L., Wiseman, J., Maniatis, T., & Lachapelle, K. (2015). Team regulation in a simulated medical emergency: An in-depth analysis of cognitive, metacognitive, and affective processes. *Instructional Science, 43*, 401–426. https://doi.org/10.1007/s11251-014-9333-6

Efklides, A. (2019). Gifted students and self-regulated learning: The MASRL model and its implications for SRL. *High Ability Studies, 30*(1–2), 79–102.

Gaba, D. M. (2007). The future vision of simulation in healthcare. *Simulation in Healthcare, 2*(2), 126–135. https://doi.org/10.1097/01.SIH.0000258411.38212.32

Haataja, E., Malmberg, J., & Järvelä, S. (2018). Monitoring in collaborative learning: Co-occurrence of observed behavior and physiological synchrony explored. *Computers in Human Behavior, 87*, 337–347. https://doi.org/10.1016/j.chb.2018.06.007

Hadwin, A., Järvelä, S., & Miller, M. (2017). Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In *Handbook of self-regulation of learning and performance* (pp. 83–106). Routledge.

Hadwin, A., & Oshige, M. (2011). Self-regulation, coregulation, and socially shared regulation: Exploring perspectives of social in self-regulated learning theory. *Teachers College Record, 113*(2), 240–264.

Harley, J., Jarrell, A., & Lajoie, S. (2019a). Emotion regulation tendencies, achievement emotions, and physiological arousal in a medical diagnostic reasoning simulation. *Instructional Science, 47*. https://doi.org/10.1007/s11251-018-09480-z

Harley, J. M., Pekrun, R., Taxer, J. L., & Gross, J. J. (2019b). Emotion regulation in achievement situations: An integrated model. *Educational Psychologist, 54*(2), 106–126.

Harley, J. M., & Pekrun, R. (2024). Managing emotions in education: The emotion regulation in achievement situations model. In J. J. Gross, & B. Q. Ford (Eds.), *Handbook of emotion regulation* (3rd ed., pp. 536–544). The Guilford Press.

Harvey, A., Nathens, A. B., Bandiera, G., & LeBlanc, V. R. (2010). Threat and challenge: Cognitive appraisal and stress responses in simulated trauma resuscitations. *Medical Education, 44*, 587–594. https://doi.org/10.1111/j.1365-2923.2010.03634.x

Horvers, A., Tombeng, N., Bosse, T., Lazonder, A. W., & Molenaar, I. (2021). Detecting emotions through electrodermal activity in learning contexts: A systematic review. *Sensors, 21*(23). https://doi.org/10.3390/s21237869

Hu, X., Sgherza, T. R., Nothrup, J. B., Fresco, D. M., Naragon-Gainey, K., & Bylsma, L. M. (2024). From lab to life: Evaluating the reliability and validity of psychophysiological data from wearable devices in laboratory and ambulatory settings. *Behavior Research Methods, 56*(7), 1–20. https://doi.org/10.3758/s13428-024-02387-3

Isohätälä, J., Näykki, P., & Järvelä, S. (2020). Cognitive and socio-emotional interaction in collaborative learning: Exploring fluctuations in students' participation. *Scandinavian Journal of Educational Research, 64*(6), 831–851.

Isohätälä, J., Näykki, P., Järvelä, S., & Baker, M. J. (2018). Striking a balance: Socio-emotional processes during argumentation in collaborative learning interaction. *Learning, Culture and Social Interaction, 16*, 1–19.

Jansen, R. S., Van Leeuwen, A., Janssen, J., Jak, S., & Kester, L. (2019). Self-regulated learning partially mediates the effect of self-regulated learning interventions on achievement in higher education: A meta-analysis. *Educational Research Review, 28*, Article 100292.

Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist, 48*(1), 25–39.

Järvelä, S., Malmberg, J., Haataja, E., Sobocinski, M., & Kirschner, P. A. (2021). What multimodal data can tell us about the students' regulation of their learning process? *Learning and Instruction, 72*, Article 101203. https://doi.org/10.1016/j.learninstruc.2019.04.004

Järvelä, S., Nguyen, A., Vuorenmaa, E., Malmberg, J., & Järvenoja, H. (2023). Predicting regulatory activities for socially shared regulation to optimize collaborative learning. *Computers in Human Behavior, 144*, Article 107737. https://doi.org/10.1016/j.chb.2023.107737

Karnatovskaia, L. V., Johnson, M. M., Varga, K., Highfield, J. A., Wolfrom, B. D., Philbrick, K. L., Ely, E. W., Jackson, J. C., Gajic, O., Ahmad, S. R., & Niven, A. S. (2020). Stress and fear: Clinical implications for providers and patients (in the time of COVID-19 and beyond). *Mayo Clinic Proceedings, 95*(11), 2487–2498. https://doi.org/10.1016/j.mayocp.2020.08.028

Kazi, S., Khaleghzadegan, S., Dinh, J. V., Shelhamer, M. J., Sapirstein, A., Goeddel, L. A., Chime, N. O., Salas, E., & Rosen, M. A. (2021). Team physiological dynamics: A critical review. *Human Factors, 63*(1), 32–65. https://doi.org/10.1177/0018720819874160

Kim, J., Neilipovitz, D., Cardinal, P., & Chiu, M. (2009). A comparison of global rating scale and checklist scores in the validation of an evaluation tool to assess performance in the resuscitation of critically ill patients during simulated emergencies (abbreviated as "CRM simulator study IB"). *Simulation in Healthcare, 4*(1), 6–16. https://doi.org/10.1097/SIH.0b013e3181880472

Lajoie, S. P., Zheng, J., Li, S., Jarrell, A., & Gube, M. (2021). Examining the interplay of affect and self regulation in the context of clinical reasoning. *Learning and Instruction, 72*, Article 101219.

LeBlanc, V. R. (2009). The effects of acute stress on performance: Implications for health professions education. *Academic Medicine, 84*(10 Suppl), S25–S33. https://doi.org/10.1097/ACM.0b013e3181b37b8f

LeBlanc, V. R. (2019). The relationship between emotions and learning in simulation-based education. *Simulation in Healthcare, 14*(3), 137–139. https://doi.org/10.1097/SIH.0000000000000379

LeBlanc, V. R., Brazil, V., & Posner, G. D. (2024). More than a feeling: Emotional regulation strategies for simulation-based education. *Advances in Simulation, 9*(1), 53.

LeBlanc, V. R., McConnell, M. M., & Monteiro, S. D. (2015). Predicsup chaos: A review of the effects of emotions on attention, memory and decision making. *Advances in Health Sciences Education: Theory and Practice, 20*(1), 265–282. https://doi.org/10.1007/s10459-014-9516-6

LeBlanc, V. R., & Posner, G. D. (2022). Emotions in simulation-based education: Friends or foes of learning? *Adv Simul (Lond), 7*(1), 3. https://doi.org/10.1186/s41077-021-00198-6

Lei, C., & Palm, K. (2025). Crisis resource management training in medical simulation. In *StatPearls*. StatPearls Publishing Copyright © 2025, StatPearls Publishing LLC.

Li, S., Zheng, J., Lajoie, S. P., & Wiseman, J. (2021). Examining the relationship between emotion variability, self-regulated learning, and task performance in an intelligent tutoring system. *Educational Technology Research & Development, 69*(2), 673–692.

Lingard, L. (2016). Paradoxical truths and persistent myths: Reframing the team competence conversation. *Journal of Continuing Education in the Health Professions, 36*(Suppl 1), S19–S21.

Liu, Y., & Du, S. (2018). Psychological stress level detection based on electrodermal activity. *Behavioural Brain Research, 341*, 50–53. https://doi.org/10.1016/j.bbr.2017.12.021

Liu, Y., Wang, T., Wang, K., & Zhang, Y. (2021). Collaborative learning quality classification through physiological synchrony recorded by wearable biosensors. *Frontiers in Psychology, 12*, Article 674369. https://doi.org/10.3389/fpsyg.2021.674369

Lobczowski, N. G. (2022). Capturing the Formation and regulation of emotions in collaborative learning: The FRECL coding procedure. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.846811

Madsgaard, A., Smith-Strøm, H., Hunskår, I., & Røykenes, K. (2022). A rollercoaster of emotions: An integrative review of emotions and its impact on health professional students' learning in simulation-based education. *Nursing Open, 9*(1), 108–121.

Malmberg, J., Haataja, E., Seppänen, T., & Järvelä, S. (2019). Are we together or not? The temporal interplay of monitoring, physiological arousal and physiological synchrony during a collaborative exam. *International Journal of Computer-Supported Collaborative Learning, 14*(4), 467–490. https://doi.org/10.1007/s11412-019-09311-4

Mayo, O., Lavidor, M., & Gordon, I. (2021). Interpersonal autonomic nervous system synchrony and its association to relationship and performance - A systematic review and meta-analysis. *Physiology and Behavior, 235*, Article 113391. https://doi.org/10.1016/j.physbeh.2021.113391

McConnell, M. M. (2024). Emotions and clinical decision-making. In *Fundamentals and frontiers of medical education and decision-making* (pp. 139–168). https://doi.org/10.4324/9781003316091-7

McConnell, M., & Eva, K. W. (2022). Emotions and learning: Cognitive theoretical and methodological approaches to studying the influence of emotions on learning. In *Researching medical education* (pp. 279–290). https://doi.org/10.1002/9781119839446.ch25

Miller, M., & Hadwin, A. (2015). Scripting and awareness tools for regulating collaborative learning: Changing the landscape of support in CSCL. *Computers in Human Behavior, 52*, 573–588.

Misal, V., Akiri, S., Taherzadeh, S., McGowan, H., Williams, G., Jenkins, J. L., … Kleinsmith, A. (2020). Physiological synchrony, stress and communication of paramedic trainees during emergency response training. In *Companion publication of the 2020 international conference on multimodal interaction* (pp. 82–86).

Olsen, J. K., Sharma, K., Rummel, N., & Aleven, V. (2020). Temporal analysis of multimodal data to predict collaborative learning outcomes. *British Journal of Educational Technology, 5*(5). https://doi.org/10.1111/bjet.12982

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341.

Pekrun, R. (2023). Mind and body in students' and teachers' engagement: New evidence, challenges, and guidelines for future research. *British Journal of Educational Psychology, 93*, 227–238.

Pekrun, R. (2024). Control-Value theory: From achievement emotion to a general theory of human emotions. *Educational Psychology Review, 36*(3), 83. https://doi.org/10.1007/s10648-024-09909-7

Salas, E., Rosen, M. A., Burke, C. S., Goodwin, G. F., & Fiore, S. M. (2006). The making of a dream team: When expert teams do best. In K. Ericsson, N. Anders, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 439–453). New York: Cambridge University Press.

Schneider, B. (2024). Three challenges in implementing multimodal learning analytics in real- world learning environments. Learning: Research and Practice. *10*(1), 103–112.

Schuurmans, A. A. T., de Looff, P., Nijhof, K. S., Rosada, C., Scholte, R. H. J., Popma, A., & Otten, R. (2020). Validity of the Empatica E4 Wristband to measure Heart Rate Variability (HRV) parameters: A comparison to electrocardiography (ECG). *Journal of Medical Systems, 44*(11), 190. https://doi.org/10.1007/s10916-020-01648-w

Sobocinski, M., Malmberg, J., & Järvelä, S. (2022). Exploring adaptation in socially-shared regulation of learning using video and heart rate data. *Technology, Knowledge and Learning, 27*(2), 385–404. https://doi.org/10.1007/s10758-021-09526-1

Stoller, J. K. (2021). Emotional intelligence: Leadership essentials for chest medicine professionals. *Chest, 159*(5), 1942–1948.

Tavares, W., Brydges, R., Myre, P., Prpic, J., Turner, L., Yelle, R., & Huiskamp, M. (2018). Applying Kane's validity framework to a simulation based assessment of clinical competence. *Advances in Health Sciences Education: Theory and Practice, 23*(2), 323–338. https://doi.org/10.1007/s10459-017-9800-3

Tomashin, A., Gordon, I., & Wallot, S. (2022). Interpersonal Physiological Synchrony Predicts Group cohesion. *Frontiers in Human Neuroscience, 16*, Article 903407. https://doi.org/10.3389/fnhum.2022.903407

Törmänen, T., Järvenoja, H., & Mänty, K. (2021). Exploring groups' affective states during collaborative learning: What triggers activating affect on a group level? *Educational Technology Research & Development, 69*(5), 2523–2545. https://doi.org/10.1007/s11423-021-10037-0

van Dooren, M., de Vries, J. J., & Janssen, J. H. (2012). Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology and Behavior, 106*(2), 298–304. https://doi.org/10.1016/j.physbeh.2012.01.020

von Suchodoletz, A., & Hepach, R. (2021). Cultural values shape the expression of self-evaluative social emotions. *Scientific Reports, 11*(1), Article 13169. https://doi.org/10.1038/s41598-021-92652-8

Wallot, S., & Leonardi, G. (2018a). Analyzing multivariate dynamics using cross-recurrence quantification analysis (CRQA), diagonal-cross-recurrence profiles (DCRP), and multidimensional recurrence quantification analysis (MdRQA) - A tutorial in R. *Frontiers in Psychology, 9*, 2232. https://doi.org/10.3389/fpsyg.2018.02232

Wallot, S., & Leonardi, G. (2018b). Analyzing multivariate dynamics using cross-recurrence quantification analysis (CRQA), diagonal-cross-recurrence profiles (DCRP), and multidimensional recurrence quantification analysis (MdRQA) – A tutorial in R. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.02232

Wallot, S., Roepstorff, A., & Mønster, D. (2016a). Multidimensional recurrence quantification analysis (MdRQA) for the analysis of multidimensional Time-Series: A software implementation in MATLAB and its application to group-level data in joint action. *Frontiers in Psychology, 7*, 1835. https://doi.org/10.3389/fpsyg.2016.01835

Wallot, S., Mitkidis, P., McGraw, J.J., & Roepstorff, A. (2016b). Beyond synchrony: Joint action in a complex production task reveals beneficial effects of decreased interpersonal synchrony. *PLoS One, 11*(12), Article e0168306. 10.1371/journal.pone.0168306.

Wilson, C., Chahine, S., Cristancho, S., Aquil, S., Mandurah, M., Levine, M., & Sener, A. (2021). Unusual suspects: Real-time physiological evaluation of stressors during laparoscopic donor nephrectomy. *Can Urol Assoc J, 15*(4), E205–e209. https://doi.org/10.5489/cuaj.6647

Winne, P. H., & Hadwin, A. F. (2008). The weave of motivation and self-regulated learning. In D. H. Schunk, & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). Lawrence Erlbaum Associates Publishers.

Worsley, M., Martinez-Maldonado, R., & D'Angelo, C. (2021). A new era in multimodal learning analytics: Twelve core commitments to ground and grow MMLA. *Journal of Learning Analytics, 8*(3), 10–27.

Yan, L., Echeverria, V., Jin, Y., Fernandez-Nieto, G., Zhao, L., Li, X., … Martinez-Maldonado, R. (2024). Evidence-based multimodal learning analytics for feedback and reflection in collaborative learning. *British Journal of Educational Technology, 55*(5), 1900–1925.

Yan, L., Martinez-Maldonado, R., Zhao, L., Li, X., & Gasevic, D. (2023). *Physiological synchrony and arousal as indicators of stress and learning performance in embodied collaborative learning*. https://doi.org/10.1007/978-3-031-36272-9_49

Zhai, J., & Barreto, A. (2006). Stress detection in computer users through non-invasive monitoring of physiological signals. *Biomedical Sciences Instrumentation, 42*, 495–500.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70.