Pijeira-Díaz, Héctor J.; Drachsler, Hendrik; Järvelä, Sanna; Kirschner, Paul A.

Sympathetic arousal commonalities and arousal contagion during collaborative learning. How attuned are triad members?

formal und inhaltlich überarbeitete Version der Originalveröffentlichung in:
formally and content revised edition of the original source in:
Computers in human behavior (2019) 92, S. 188-197

Bitte verwenden Sie in der Quellenangabe folgende URN oder DOI /
Please use the following URN or DOI for reference:
urn:nbn:de:0111-pedocs-180871
10.25656/01:18087

https://nbn-resolving.org/urn:nbn:de:0111-pedocs-180871
https://doi.org/10.25656/01:18087

*Title Page with Author details

Sympathetic arousal commonalities and arousal contagion during collaborative learning: How attuned are triad members?

Héctor J. Pijeira-Díaz[a✉], Hendrik Drachsler[b,c,d], Sanna Järvelä[a], Paul A. Kirschner[a,d]

{Hector.PijeiraDiaz, Sanna.Jarvela}@oulu.fi

{Hendrik.Drachsler, Paul.Kirschner}@ou.nl

[a] Faculty of Education, University of Oulu, Finland, P.O. BOX 2000, 90014

[b] Goethe University Frankfurt, Robert Mayer 10, Frankfurt, Germany

[c] German Institute for International Educational Research (DIPF), Solmsstraße 73-75, 60486 Frankfurt, Germany

[d] Welten Institute, Open University of the Netherlands, Valkenburgerweg 177, 6401 AT Heerlen

✉ Corresponding author

ABSTRACT

This article explores the dynamics of collaborative learning in the classroom from the perspective of the commonalities and interdependence in the degree of physiological activation from the sympathetic nervous system (i.e., sympathetic arousal) of group members. Using Empatica E4 wristbands, electrodermal activity—to derive arousal—was measured in 24 high school students working in groups of three (i.e., triads) during two runs of an advanced physics course. The participants met three times a week over six weeks for lessons of 75 min each. Most of the time (≈ 60–95% of the lesson) the triad members were at different arousal levels, and, when they were on the same level, it was mainly the low arousal (or deactivated) level. Less than 4% of the time were the triad members *simultaneously* in high arousal. Possible within-triad *arousal contagion* cases (71.3%) occurred mostly on a one-to-one basis and with a latency from within a few seconds up to ten min, but usually within one min. This study supports the view that only small parts of group work are collaborative, as far as the synchronicity and coordination which collaboration presupposes. Although exploratory, results also illustrate the affordances of physiological measures to characterize collaborative processes.

Keywords: Collaborative learning; Interpersonal physiology; Sympathetic arousal; Arousal contagion; Biosensors; Electrodermal activity

SYMPATHETIC AROUSAL COMMONALITIES AND AROUSAL CONTAGION DURING COLLABORATIVE LEARNING: HOW ATTUNED ARE TRIAD MEMBERS?

1. Introduction and theoretical background

National education systems increasingly emphasize collaborative pedagogies (e.g., project-based and task/problem-oriented learning) in their curricula so that students develop the collaboration skills needed to thrive in both their academic life and future workplace (OECD, 2017). Collaboration promotes greater learning outcomes if it is properly implemented (cf. Johnson & Johnson, 2002), the individuals are committed to the process, and the task is group worthy (Kirschner, Sweller, Kirschner, & Zambrano, 2018). If improperly implemented, collaboration is not always effective, and actual collaboration will only occur in specific phases of group work (Baker, 2015). Hence, studying collaboration metrics which could assist in distinguishing effective from ineffective or even counterproductive collaboration is of great interest (Barron, 2000; Vogel & Weinberger, 2018).

The complex processes of collaboration present a challenge for consistent, accurate, and reliable measurement across individuals and across user populations (OECD, 2017). Measuring collaborative processes most often relies on the analysis of the communication among group members (K. Sawyer, 2013). Communication provides a window into the cognitive and social processes related to collaborative skills, such as grounding, mutual goal establishment, progress toward goals, negotiation, consensus, sharing perspectives, social states, and judging the quality of solutions generated (Kuhn, 2015). However, conversational processes alone do not explain the effects observed in collaborative learning (Dillenbourg, Baker, Blaye, & O'Malley, 1996). To properly understand the collaborative process, it is also necessary to take into account the features of non-verbal communication which are integral to the collaboration process, such as physiological responses, bodily posture, gestures, etc. (Baker, 2015).

In particular, physiological measures can provide continuous information about participants' cognitive-affective states through their level of arousal, including those states which are outside of awareness and may not be readily observable (Di Mitri, Schneider, Specht, & Drachsler, 2018; Thorson, West, & Mendes, 2018). In general, arousal refers to the degree of physiological activation and responsiveness triggered by an event, object, or situation, during a person's interaction with the environment (De Lecea, Carter, & Adamantidis, 2012; Juvina, Larue, & Hough, 2017). Arousal, however, is a complex physiological response manifested in a variety of bodily systems (e.g., the cardiovascular, electrodermal, muscular, hormonal, and immune systems) (Cattell, 1972; Schwartz, Collura, Kamiya, & Schwartz, 2017), which do not necessarily show similar patterns (Lacey, 1967). Accordingly, different types of arousal have been recognized in the literature (e.g., sympathetic, electrocortical) (Frijda, 1986). This paper focuses on sympathetic arousal, sometimes hereinafter simply referred to as arousal, due to its connections to cognitive-affective processes (Critchley, Eccles, & Garfinkel, 2013; Pecchinenda, 1996; Spangler, Pekrun, Kramer, & Hofmann, 2002), and its accessibility via electrodermal activity (EDA) as a proxy (Boucsein, 2012); a measure with a long history used extensively in psychophysiological research (Dawson, Schell, & Filion, 2017; Kreibig, 2010). Sympathetic arousal (i.e., activation of the sympathetic nervous system) is associated with the so-called fight-or-flight response, as the body anticipates and prepares for action, either mental or physical (Hanoch & Vitouch, 2004; Poh, Swenson, & Picard, 2010). As such, sympathetic arousal has cognitive and affective components (and physical, but that is out of the scope of the paper). For the cognitive part, there is a close interdependence between arousal and the cognitive process of attention (Sharot & Phelps, 2004), which have been suggested as sharing a neural substrate (Critchley, 2002). For the affective part, arousal is the activation dimension of emotion, according to the classical circumplex model of affect (Russell, 1980; Russell & Barrett, 1999). The fact that collaborative work is determined by the neural integration of cognitive, affective, and physiological responses (Damasio, Tranel, & Damasio, 1991) renders arousal applicable to the study of collaboration.

1.1. Collaborative learning

In terms of learning, the process of collaboration is as important as the outcome, since the aim is often not only that students properly carry out a task or reach a correct problem solution, but also that they acquire the skills to carry out such tasks or tackle such problems more efficiently in the future, whether together or alone, in virtue of the appropriation of a co-elaborated, deeper task-domain conceptual understanding (Baker, 2015). Collaboration is said to be conducive to higher order thinking skills through effective communication (Cohen, 1994). Collaborative processes thought to be beneficial for learning include, but are not limited to, knowledge co-construction (Jeong & Hmelo-Silver, 2016), transactivity or building on each other's reasoning (Weinberger, Stegmann, & Fischer, 2007), negotiation (Miyake & Kirschner, 2014), and argumentation (Baker, 2015). Other support processes are, for example, identifying shared knowledge (Roschelle & Teasley, 1995), establishing common ground (Reiter-Palmon, Sinha, Gevers, Odobez, & Volpe, 2017), mutual modelling of the collaborative partner's knowledge state (Dillenbourg, 1999), and coordinating joint efforts (Järvenoja, Järvelä, & Malmberg, 2015). All of these processes implicitly highlight the notion of synchronicity and simultaneity presumed in collaborative learning, as opposed to cooperative learning which is characterized by individually solving sub-tasks through division of labor (Dillenbourg, 1999). "Joint" and/or "together" are keywords that the literature consistently employs to define collaborative learning or associated processes. Such descriptions include (emphasis added in all cases) "working *together*" (Cohen, 1994), "mutual engagement of participants in a coordinated effort to solve the problem *together*" (Roschelle & Teasley, 1995), the notion of "*joint* attention as social cognition" (Tomasello, 1995), partners doing the work *together* (Dillenbourg, 1999), "a *jointly* produced activity" (Enyedy & Stevens, 2014), "an active and *joint* process" (Baker, 2015), and "working *together* toward a shared learning goal" (Jeong & Hmelo-Silver, 2016), among others. Accordingly, shared team states are regarded as critical (Reiter-Palmon et al., 2017).

In their classic definition, Roschelle and Teasley (1995) see collaboration as a coordinated, synchronous activity resulting from a sustained attempt to construct and maintain a shared conception of a problem and the way to its solution. Such coordination and synchronicity are hypothesized to be reflected in physiological commonalities and to influence the physiological responses of the collaborators (Palumbo et al., 2017). Accordingly, there is a need to investigate the characteristics of physiological group responses, such as homogenous activation levels and simultaneous or sequential physiological changes (Knierim, Rissler, Dorner, Maedche, & Weinhardt, 2018). Moreover, physiological metrics of collaboration can be considered universal, as they are objective indices independent of language and culture. This study explores collaborative learning from the physiological perspective of sympathetic arousal. Particularly, it examines arousal direction, arousal levels (i.e., low, medium, high), and influence in terms of arousal contagion (i.e., the spread of arousal between group members) within groups of three (i.e., triads) in a naturalistic collaborative learning setting.

1.2. Interpersonal physiology

Beginning in the 1950s, social scientists started collecting data from two or more people in interpersonal interactions to measure commonalities and interdependence between their physiological states (Thorson et al., 2018), and the body of literature in the field keeps growing (Palumbo et al., 2017). Two reasons for the increasing popularity of physiological measures, apart from their affordances, might be, first, the need for more objective ways to measure cognitive-affective states (DeLozier & Rhodes, 2017), and, second, the rising availability of lightweight, non-cumbersome wearable biosensors, such as wristbands (Swan, 2012). The latter allows for continuous and unobtrusive measures in ecologically valid learning contexts outside the laboratory. The growing interest in interpersonal physiology has been ascribed to its role in early development, language learning, group engagement, social connection, and group membership (Delaherche et al., 2012; Palumbo et al., 2017). One specific practical application to group work, for example, has been the development of intelligent interruption management systems based on a group's arousal levels. It has been found that presenting task-required additional or updated information during arousal acceleration results in a significant increase in both task performance and reported collaboration experience, as compared to randomly interrupting the group to provide such information (Goyal & Fussell, 2017).

Interpersonal physiology is a general term which refers to any observed interdependence or association between two or among more than two persons' physiologies (Palumbo et al., 2017). In dyads, the stability and influence model (Thorson et al., 2018) considers how a person's physiology at a certain point in time is predicted by both his/her own physiology at a prior time point (i.e., the stability effect) and by the other dyad member's physiology at the prior time point (i.e., the influence effect). The occurrence of influence implies the existence of at least an influencer and an influenced in the dyad. This model easily scales up to groups of any size. The influence is not explained by physiology itself, as such, but is associated with cognitive-affective exchanges between or among group members (e.g., one person's heart rate influencing a partner's heart rate via a verbal outburst of anger) (Postmes, Spears, Lee, & Novak, 2005; Thorson et al., 2018).

In this context of physiological influence, our interest in arousal contagion has been inspired by the notion of emotional contagion, given that arousal is the activation dimension of emotions according to the circumplex model of affect (Russell, 1980). Emotional contagion theorists define it as the catching of others' emotions, via feedback from facial and vocal expressions, postures, and behaviors (Hatfield, Cacioppo, & Rapson, 1994). Arousal contagion helps studying the same phenomena in groups (e.g., to assess collaboration or leadership) (Chanel & Muhl, 2015), without the need to distinguish specific emotions, which could result in insufficient discriminant validity, as there are overlapping traits in different emotions (Pekrun, Goetz, Titz, & Perry, 2002). As a degree of physiological activation, arousal is objectively measurable. Moreover, arousal is a concomitant not only of affective processes, but also of cognitive processes. In terms of arousal contagion as a result of interpersonal influence, this paper explores the patterns of its occurrence among triad members in collaborative learning, particularly, the ratio of influencers to those influenced, the delay in the influence effect (i.e., latency), and the number of influencers.

The variety of methodologies and approaches from different disciplines are reflected in the number of conceptualizations used in the study of interpersonal physiology, including but not limited to attunement, concordance, contagion, co-regulation, coupling, covariation, entrainment, influence, linkage, physiological markers of togetherness, and synchrony (Thorson et al., 2018).

Occasionally, these terms imply conceptual differences in what is being measured or are used to refer to specific analytic techniques or theoretical approaches, but a systematic review of the literature shows that this is inconsistent (Palumbo et al., 2017).

Approaches have looked at, for example, the similarities in the amplitudes of the signals, their rate of change, their direction of change, and their linear relationship (i.e., correlation) (Elkins et al., 2009; Pijeira-Díaz, Drachsler, Järvelä, & Kirschner, 2016). Almost all such indices are computed pairwise as a direct consequence of their definition, with the notable exception of directional agreement. Directional agreement refers to the percentage of the time in which the physiological signals of a group of interacting individuals (e.g., in collaborative learning) are going in the same direction, that is, simultaneously rising, falling, or staying constant. Obviously, the physiological signals of at least two individuals are necessary to compute directional agreement, but there is no limit to the number of individuals it can account for, which, in addition to its simplicity, makes it a convenient choice. Previous work has computed directional agreement on the electrodermal activity (EDA) signal as such (Pijeira-Díaz et al., 2016). Here, taking advantage of the flexibility and affordances of such indices as discussed, we use directional agreement applied to arousal (i.e., arousal directional agreement) as a way to look at the physiological commonalities of students learning together.

1.3. Interpersonal physiology of collaborating students

Different types of interpersonal relationships have been the focus of studies on interpersonal physiology, such as parent-child, therapist-client, individuals in a couple, artist-audience, conductor-choir, friends, teammates, groups of strangers, etc. (Palumbo et al., 2017). Physiological interdependence has been found to be greater, for example, between individuals who are in a romantic relationship (Sbarra & Hazan, 2008), in psychotherapists who are more empathetic to their clients (Delaherche et al., 2012), and in game players who report greater rapport with each other (Noy et al., 2014). However, when it comes to students in a collaborative learning setting, the use of physiological data has received little attention, as evidenced by only two papers found in a recent systematic review of the literature including all publication dates through November 2015 (Palumbo et al., 2017). Early examples are found in the work of Kaplan and associates in the 1960s, which explored sympathetic arousal in relation to the affective orientation between group members in either dyads or four-person groups of medical and nursing students (Kaplan, 1967; Kaplan, Burch, Bedner, & Trenda, 1965; Kaplan, Burch, Bloom, & Edelberg, 1963). Four decades later, sympathetic arousal commonalities were explored in dyads of college students in terms of nonlinear dynamics and linkage effects (Guastello, Pincus, & Gunderson, 2006). Physiological commonalities in emotion management and convergence during collaborative processes have been explored at the dyad level using a variety of physiological responses and measures such as electrocardiography, respiration, EDA, temperature, and eye-tracking (Chanel, Bétrancourt, Pun, Cereghetti, & Molinari, 2013). More recently, physiological commonalities in relation to cognitive and affective states (e.g., mental workload, emotional valence) during a pair-programming task in a classroom environment have been studied using electrocardiography and EDA (Ahonen et al., 2016; Ahonen, Cowley, Hellas, & Puolamäki, 2018). All these previous studies focus on physiological responses from the autonomic nervous system, which can be measured more cheaply, quickly, and unobtrusively than those of the central nervous system, and are more easily interpreted (Novak, Mihelj, & Munih, 2012). Nonetheless, using portable technology in a high school classroom during 11 lessons, commonalities in electroencephalogram signals have been suggested to predict group engagement and social dynamics (Dikker et al., 2017). None of the studies focused on triads as the unit of analysis.

The dyadic nature of some of the relationships studied in general interpersonal physiology (e.g., therapist-client and couples), together with the greater availability of pairwise physiological indices, might be among the reasons why interpersonal physiology has predominantly focused on dyads as units of analysis (Delaherche et al., 2012). However, the general dynamics of triads and larger groups are different to that of dyads. For example, dyads cannot form coalitions, have no newcomers or old-timers, nor can there be majority/minority influence in dyads. Also, negotiation and conflict (i.e., argument) are more complicated in triads and larger groups (Reiter-Palmon et al., 2017). When it comes to learning, dyads are often considered to be peer learning, with real collaborative learning beginning with three or more participants. Therefore, the focus of this article is on triads.

1.4. Aim

The aim of the research reported here is twofold: first, to explore the within-triad commonalities of students in collaborative learning from the physiological perspective of sympathetic arousal (RQ1 and RQ2), and, second, to study arousal contagion among triad members (RQ3 and RQ4). Arousal contagion involves a ratio of influencers to influenced, which could be one-to-one, one-to-two, or two-to-one in a triad. The two aims are expressed through the following research questions:

RQ1. What is the arousal directional agreement of triads in collaborative learning during a lesson?

RQ2. How often are triad members on the same arousal level (i.e., low, medium or high) lesson-wise during collaborative learning?

RQ3. What is the pattern of high arousal contagion among triad members in collaborative learning in terms of the ratio of influencers to those influenced and the contagion latency?

RQ4. How is the distribution of arousal contagion cases in triads during collaborative learning in terms of influencers?

2. Method

2.1. Participants

The participants ($N = 24$) in the study were Finnish high school students, randomly selected from those enrolled in the regular, elective Advanced Physics course, which runs twice during the spring term. Although more students enrolled in the course, the number of participants selected for the study was dictated by the twelve available Empatica E4 wristband sensors (see *2.2*) which enabled the simultaneous tracking of four triads (i.e., 12 participants) during each of the two course runs. The gender distribution was six females (25%) and eighteen males (75%), ranging from 16 to 17 years of age. All participants were of Caucasian ethnicity. They were high achievers in the preceding physics course (*mean* = 9.0/10; *SD* = 0.6).

After approval for the study was obtained from the relevant ethics committee, participants were asked to provide informed written consent. Participation in the study was voluntary, and the students could revoke their consent at any time during the course.
From the 24 participants, eight triads were formed based on the principles of gender balance (Bear & Woolley, 2011) and heterogeneity of learning regulation profiles, for the sake of between-team comparability. Students were asked to answer the Motivated Strategies for

Learning Questionnaire (Pintrich, Smith, Garcia, & Mckeachie, 1993) as a measure of their self-regulation profile. Based on the questionnaire score, students were categorized into three statistically significantly different groups ($F(2, 40) = 41.35$, $p < .001$) of self-regulation, namely, low, middle, and high. The descriptive statistics of the questionnaire score for the different groups are provided in Table *1*. With this information, seven of the eight triads were formed with one student from each of three self-regulation categories, and one triad (triad 1) consisted of one high and two low self-regulating students. As for gender distribution, six triads had one female and two males, and the remaining two triads (triads 3 and 6) were composed of three males.

Table 1

Descriptive Statistics of the Motivated Strategies for Learning Questionnaire Score

Self-regulation profile	Mean	SD
Low	111.7	12.8
Middle	139.5	5.6
High	156.3	18.3

2.2. Materials

Empatica® E4 wristbands (Empatica Inc., Cambridge, MA, USA) were used to record EDA continuously and unobtrusively during the two course runs. In contrast to lower-accuracy, consumer-oriented wristband sensors available on the market—typically for fitness tracking—the E4 is a research quality multi-sensor wristband (Garbarino, Lai, Tognetti, Picard, & Bender, 2014), and is therefore significantly more expensive, which limited the number at our disposal to twelve. The wristband embeds four sensors—EDA, photo-plethysmograph, thermometer, and accelerometer—and has two mutually exclusive modes of operation: Bluetooth streaming and internal memory recording. The latter was chosen as more suitable for offline processing, given our exploratory intentions. The streaming mode, on the other hand, offers potential for real-time applications, such as for biofeedback, learning dashboards, or intervention flags. The E4 wristband EDA sensor uses the exosomatic method, which measures skin conductance in microSiemens (μS) by applying a small external current (Edelberg, 1967). The sampling frequency of the EDA sensor is 4 Hz (i.e., four samples per second).

2.3. Procedure

The Advanced Physics course was selected to explore arousal in the classroom because it is elective and rated by the teachers as difficult, a combination which potentially could raise the level of sympathetic arousal. That the course is elective enables making the assumption that the students taking the course are interested in the topic, since it is their own choice. Interest invites attention and engagement, both positively correlated to arousal. Once the students are engaged in their learning, difficulty increases cognitive and emotional arousal.

The course has a frequency of three 75 min lessons weekly. In all, the course consists of 18 lessons. Lesson topics include, among others, waves, interference, dispersion, and geometrical optics. In general, the lessons begin with a theoretical part—the teacher's explanation of the topic—followed by a collaborative practical consisting of either paper-and-pencil problems or hands-on experiments. For this study, in the practical part, students put the learned physical theories, laws, and principles to the test in triads, and the participants remained in the same triads for the duration of the course.

The classroom temperature was kept constant at 23 °C to avoid thermoregulatory interference with the psychological EDA measurement (Boucsein, 2012). This did not require any intervention, as it was the standard thermostat setting in the school.

Students were instructed to wear the E4 wristband on their non-dominant hand in order to reduce the effect of movement artefacts (Boucsein et al., 2012), and to adjust the wristband strap neither loosely nor tightly so that the electrodes made proper but not excessive contact with the skin which would result in a pressure artefact (Edelberg, 1967). They put the wristband on at the beginning of the lesson and took it off at the end.

The attendance average (i.e., average number of participants present in each course lesson) was 10.9/12 (*SD* = 1.3) students for the first course run, and 10.5/12 (*SD* = 1.3) students for the second.

3. Analysis

Arousal at a certain time instant (with a resolution of 250 ms, the sampling interval of the sensor) was operationalized as the count of skin conductance response (SCR; peak in the EDA signal) onsets during the previous min.

Ledalab (version 3.4.9; http://www.ledalab.de/) software, based on MATLAB (The MathWorks, Inc., Natick, MA, USA.), was used for EDA signal processing as recommended by the wristband manufacturer (Empatica Inc., 2015). The Ledalab algorithm assumes raw, unfiltered data as the input for decomposition analysis and feature extraction of the EDA signal (Benedek & Kaernbach, 2010). Therefore, no signal pre-processing, such as cleaning or filtering, was performed on the EDA signal as obtained from the wristband sensor.

Continuous decomposition analysis (Benedek & Kaernbach, 2010) was used for extracting SCRs (peaks), as it enables separate detection of superimposed responses. This is important to obtain an unbiased measure since, more often than not, a subsequent SCR occurs during the decay of a previous one. In addition, one must decide on the amplitude threshold beyond which a change in conductance will be considered an SCR. A choice was made for 0.05 µS, which is the customary minimum amplitude used in EDA research (Braithwaite, Watson, Jones, & Rowe, 2015; Dawson et al., 2017).

After continuous decomposition analysis, onsets for each individual SCR were obtained in Ledalab. Onsets served as the basis to subsequently calculate the SCR frequency minute-wise, using a moving window approach with a window width of 1 min and a moving step of 250 ms—the sampling interval of the sensor. Thus, the SCR frequency during the lesson was obtained for every time instant (resolution of 250 ms) after the first min of the lesson, as it corresponded to the number of SCR onsets in the previous min. Since no amplitude measure was considered, no correction in the form of standardization was needed for comparison across individuals.

SCR frequency was used to categorize arousal into low, medium, and high levels. Typically, a frequency of 1–3 ppm (peaks/min) occurs at rest (Dawson et al., 2017, p. 225), and, as frequency increases with the arousal level, values higher than 20 ppm are interpreted as high arousal (Boucsein, 2012, p. 222). Accordingly, frequencies of up to 3 ppm were labelled as low arousal, from 20 ppm up were labelled as high arousal, and anything between these two values was labelled as medium arousal.

Using the minimum of 1 ppm as the worst case, we discarded from the analysis all the EDA recording sessions where the total number of SCR detected was less than 1 ppm times the lesson duration in min. In that way, we automatically excluded sessions where improper electrode-skin contact, due to wristband maladjustment, basically led to the measurement of noise rather than actual EDA signal. This devised technique served as a convenient aid to the manual visual inspection, recommended as the starting point for EDA analysis (Braithwaite et al., 2015). As a result, 137/420 (33%) of the EDA recording sessions were discarded as noisy. Additionally, 46/420 (11%) of the recording sessions were missing due to absent students, leaving 237/420 (56%) as valid EDA recording sessions. This is similar to that reported by Arroyo and colleagues (2009) from a study also using biosensors in a high school science classroom. Unfortunately, only 18 cases (corresponding to 18 x 3 = 54 recordings) were found from such valid recordings where valid EDA data were available for all members of a triad in a certain lesson. Therefore, 54 (13% of the 420 nominally expected recordings, and 23% of the 237 valid ones) of the EDA recording sessions were used to answer the research questions, since triads are the unit of analysis of this study. In the final sample for analysis, 7/8 triads (no data for triad 5 in any lesson) and 15/35 lessons were represented.

The computation of arousal directional agreement (for RQ1) was based on the slopes of the trend lines. Similar to how arousal itself was computed, a moving window of 1 min was used to obtain a trend line of arousal every 250 ms. The minute-wise arousal trend line was calculated using the *polyfit* function of Matlab. The sign of the slope of each trend line served as the direction for each instant, namely, upward (+1), constant (0), or downward (-1). Finally, again using a moving window of 1 min with a step of 250 ms, the arousal directional agreement for each triad at every particular instant was computed as the percentage of the previous min in which all triad members' arousal was going in the same direction.

For RQ2, the triad members' arousal levels (different, same: low, same: medium, and same: high) are straightforward from the categorization as low, medium, or high of the minute-wise arousal obtained every 250 ms for each member.

To study high arousal contagion (RQs 3 and 4), the duration of each high arousal interval was marked at the beginning of the interval. For each interval, it was determined if, within it, some other triad member(s) reached the level of activation of high arousal (influencer point of view, potential arousal contagion), and, if so, at which latency (i.e., time elapsed from the interval start to a triad peer's high arousal interval start). It was also determined how many triad peers could have caused it via arousal contagion (influenced point of view), if any.

After the meticulous analysis above described, the results presented in the next section were obtained.

4. Results

4.1. Research question 1

Table *2* lists the descriptive statistics of each triad's arousal directional agreement for the lessons where valid data are available for all triad members. Arousal directional agreement is computed on a min basis, following a moving window approach. In this way, a value is computed for the previous min every 250 ms, the sampling interval of the EDA sensor. Total disagreement (0%), as well as high values of agreement, occurred in all the lessons, and, in half of them, maximum agreement (100%) was reached. On average, arousal directional agreement was below 30%.

Table 2

Descriptive Statistics of Triads' Arousal Directional Agreement

	Min arousal directional agreement (%)

Triad	Lesson	Course run	Min	Max	Mean	SD
1	8	1	0	86	12	20
1	9	1	0	100	26	37
1	11	1	0	100	13	24
2	1	1	0	95	13	24
2	10	1	0	86	2	10
3	9	1	0	86	8	18
3	10	1	0	100	11	23
4	3	1	0	100	28	37
4	7	1	0	89	9	19
4	11	1	0	100	11	23
4	13	1	0	93	11	21
4	18	1	0	100	13	24
6	3	2	0	100	24	28
6	9	2	0	99	7	17
6	14	2	0	100	7	18
6	17	2	0	84	9	20
7	6	2	0	61	4	11
8	2	2	0	100	19	26

Fig. 1 shows the distribution of all the computed min arousal directional agreement. The frequency of total disagreement (0%) is 55 times that of total agreement (100%), which is the next most common value. In between, the figure shows how the full range is covered with a tendency of frequency increasing toward low agreement (right to left).

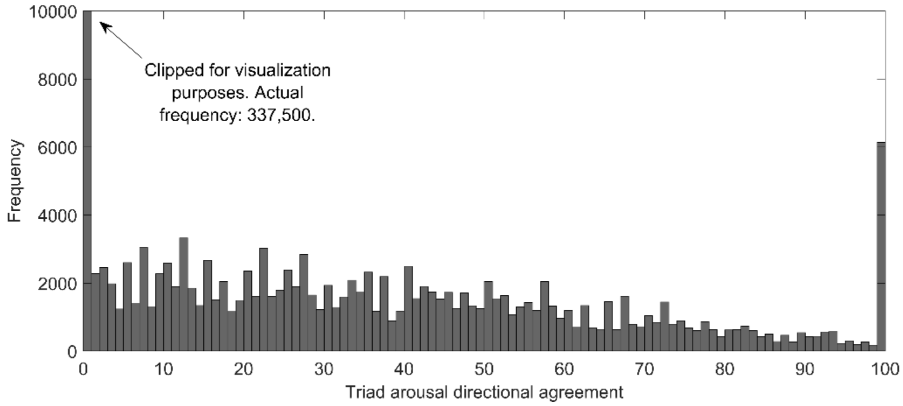

Fig. 1.

Distribution of triads' arousal directional agreement.

Fig. 2 displays, as an example, the evolution of arousal directional agreement for triad 1 during lessons 8, 9, and 11. The figure shows how the visualization of this indicator resembles a succession of peaks. In some particular moments of the lessons, arousal directional agreement increases in the triad, reaching the maximum in two of the three lessons, with a tendency of returning to a total disagreement level (0%). It is also visible that full agreement occurs in some cases instantaneously (e.g., see encircling A), while, in others, it persists for a few min (e.g., see encircling B).

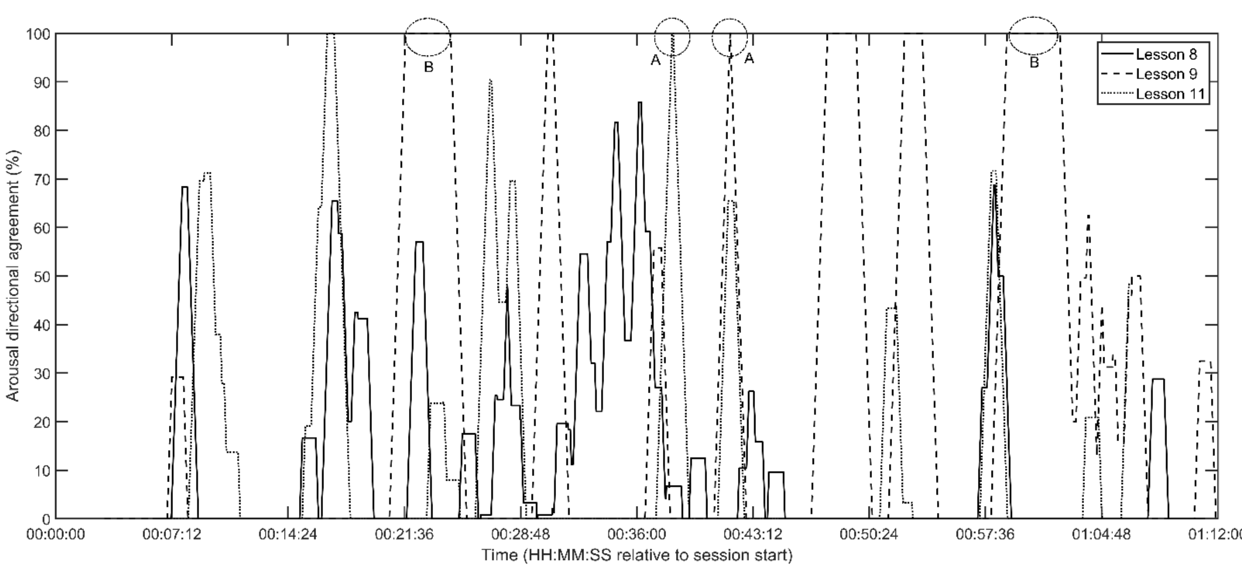

Fig. 2.

Example of a triad's arousal directional agreement throughout three lessons.

Note. Highlighted examples: encircling A for instantaneous full agreement and encircling B for longer lasting full agreement.

4.2. Research question 2

After looking at the directional agreement of arousal in the triad members, the commonalities in the level of arousal of the triad were inspected. These are summarized in *Table 3*. Most of the time (from about 60–95%) triad members were at different arousal levels, and when they were at the same level, it was mainly low arousal (or deactivated), followed by medium and high simultaneous arousal. The results show that, in a very small percentage of the lessons, the triad is simultaneously in a state of activation (whether medium or high arousal). Moreover, simultaneous high arousal never occurred for half of the triads.

Table 3

Percent of Lesson Time the Triad are in the same or a Different Arousal State

	Triad members' arousal level (% of lesson time)				
Triad	Different	Same: low	Same: medium	Same: high	Lessons aggregated (valid EDA)
1	59.2	39.1	1.7	0.0	3
2	95.0	4.8	0.3	0.0	2
3	71.6	26.7	1.7	0.0	2
4	59.2	39.8	0.8	0.1	5
6	88.5	6.0	3.9	1.6	4

7	64.8	33.7	1.5	0.0	1
8	91.7	3.0	1.4	3.8	1

4.3. Research question 3

This question examines the occurrence and latency of possible arousal contagion, particularly of high arousal as the most intense activation level. Overall, 635 high arousal intervals were found individually, of which 263 (41%) might have caused or were caused by arousal contagion. Table *4* lists the results. Possible arousal contagion cases took place mostly (71.3%) on a 1:1 basis, and 1:2 contagion cases accounted for about 15%, while cases where two triad members in high arousal could have brought the third member to that activation level (i.e., 2:1) through arousal contagion represented almost 14%.

Table 4

Possible High Arousal Contagion Cases

Number of influencers	Number of influenced	Number of cases	Percent of cases
1	1	82	71.3
1	2	17	14.8
2	1	16	13.9

As for latency, cases of possible arousal contagion were found to occur from within a few seconds up to close to 10 min. As displayed in the histogram of Fig. 3, the chances of contagion tend to decrease as time goes by, with most of it happening within one min.

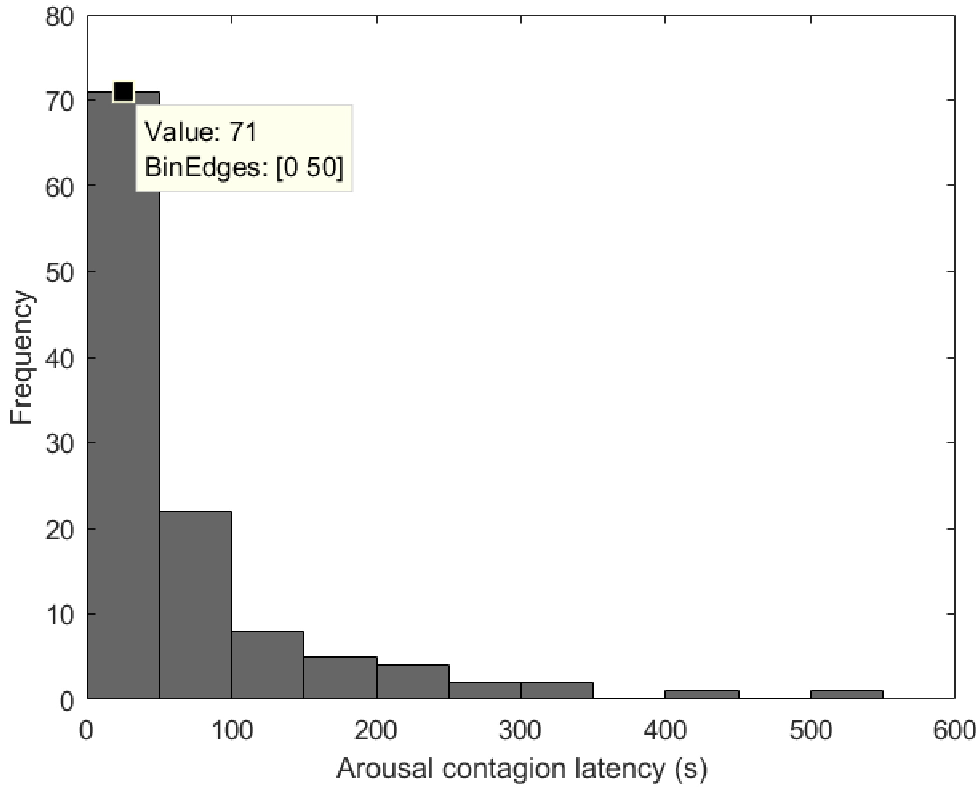

Fig. 3.

Distribution of arousal contagion latency.

4.4. Research question 4

Based on the data of possible arousal contagion cases, the triad members acting as influencers were examined, as summarized in Table *5*. The results show how, in some cases, there is a predominant influencer in the triad from whom the arousal contagion arises (e.g., student 3 in triad 4), or two influencers (e.g., students 1 and 2 in triad 6), while in other cases the arousal contagion is quite evenly originated from all three members (e.g., triad 8).

Table 5

Possible Influencers among Triad Members

Triad	Count of possible arousal contagion cases caused by the different triad members			Lessons aggregated (valid EDA)
	Student 1	Student 2	Student 3	
1	0	0	3	3
2	2	1	8	2
3	1	1	0	2
4	3	2	10	5
6	22	22	7	4
7	6	0	1	1
8	4	3	3	1

5. Discussion

RQ1. Arousal directional agreement is a simple, yet insightful, index to measure the extent to which team members are moving in the same direction within the activation dimension. Even though the team is sitting together in the same classroom situation, sharing the task and environment, the results reveal that their arousal direction is mostly not aligned. In general, they move toward activation or relaxation at distinct moments, probably driven by different outcomes of the cognitive appraisal process (Critchley et al., 2013). However, although low agreement seems to be the norm, full agreement (100%) was found to be the second most common value. The visualization example of Fig. 2. illustrates that, for some particular intervals, one of the triads sustained a state of full arousal directional agreement for several continuous min (encircled and labelled B in the figure). Thus, they were completely attuned, in terms of activation, in those moments. It seems reasonable to claim that they were truly collaborating, given that collaboration is a jointly produced activity (Enyedy & Stevens, 2014), an active process (Baker, 2015), where joint attention is considered paramount (Tomasello, 1995), and given the close relation between arousal and attention (Sharot & Phelps, 2004). The same triad, however, did not make it to full agreement at any other time during an entire lesson. Similarly, this occurred for half of the cases, with some triads (e.g., two, seven) never reaching full agreement during the course. These results are aligned with the review by Kreijns, Kirschner, and Jochems (2003), showing that, despite the affordances of the classroom for collaborative learning, including technological support, collaboration does not always happen, which they ascribed to the social aspects of collaboration often being taken for granted.

RQ2. Not only were triad members most commonly moving in different arousal directions, but also predominantly functioning at different arousal levels, as the results for RQ2 showed. Students seem to share moments of relaxation or deactivation when none of the team members are in a state of activation, whether medium or high. Other than that, the low percentage of simultaneous medium or high arousal suggests that, when the team is doing something, the work is being led or conducted by one or two of the triad members who have been termed "task doers" or the people who are leading the problem-solving at the moment (Miyake & Kirschner, 2014). Meanwhile, the other member(s) is/are less involved or engaged. Collaborative learning, it is known, is much more than putting students into groups to learn (Johnson & Johnson, 2002). Since collaboration presupposes "a mostly synchronous interaction," the results of RQ1 and RQ2 align with the view that, in general, only certain specific phases of group work are truly collaborative (Baker, 2015, p. 455).

RQ3. According to the results, 4 in 10 intervals of high arousal might result in arousal contagion, meaning that a triad member in high arousal (i.e., the influencer) causes a teammate or two (i.e., the influenced) to reach that level of activation (Hatfield et al., 1994). One has to keep in mind that this ratio serves as an upper boundary, since the other team member(s) might have come to the high arousal level on his or her own through mechanisms other than arousal contagion. Notwithstanding, it seems reasonable to assume that most cases are driven by contagion, given that the team members are closely working together, and, although not impossible, it is difficult to isolate oneself in a collaborative learning setting (Miyake & Kirschner, 2014). Also, the importance of coordinated and cohesive participation has been repeatedly emphasized by collaborative learning research (e.g., Barron, 2000; Kreijns et al., 2003). Thus, most likely, team members might be influencing one another according to the stability and influence model (Thorson et al., 2018). The fact that the majority of arousal contagion cases happened on a one-to-one basis suggests that, although there are three students working (learning in this case) together, the interaction seems to be mostly between two of the members, not necessarily always the same two, but between two rather than among the three of them. Although *a priori* it would seem more probable that two triad members in high arousal provoke the third to be highly active, compared to one provoking the other two to reach a high arousal state, the results showed that arousal contagion happened quite similarly on a 1:2 basis (14.8%) as on a 2:1 basis (13.9%).

Provided that arousal contagion occurred, its latency was found to be relatively short, mainly within the first min. This might suggest that arousal contagion is most likely to happen if the high arousal cues are made available early by the triad member(s) at that activation level (i.e., the influencer). Longer latency values (e.g., from 5 to 10 min) might be explained in a number of ways. On one hand, no visible cue might be indicating that another team member is at a high arousal level (Thorson et al., 2018). On the other hand, the cue might be visible, but the appraisal differs at first (e.g., the to be influenced team member does not see a need to become active at that moment according to the situation or his/her own motivation to participate), and the situation might be reappraised a few min later (e.g., due to a persistent cue or a situation change) (Koole, 2009).

RQ4. When looking at who in the triad is causing the potential arousal contagion (i.e., the group influencer(s)), evidence of all profiles which could be expected was found, namely, arousal contagion being led by one or two in the triad, and a distributed leadership with all three members quite evenly responsible for arousal contagion. These results highlight the distinct dynamics of collaborative learning (cf. Enyedy & Stevens, 2014; Miyake & Kirschner, 2014) unfolding in different triads.

Triads with a predominant influencer are aligned with reports of systematic inequalities in participation during collaborative learning, which have been ascribed, among other factors, to differences in self-efficacy (Cohen, 1994). Self-efficacy influences thought patterns, actions, and arousal (Bandura, 1982). The predictive power of self-efficacy beliefs on students' academic functioning has been extensively verified (Zimmerman & Schunk, 2003). However, given that the sample is composed of high achievers, it is reasonable to presume high self-efficacy in all participants. Therefore, self-efficacy alone does not seem to explain the cases of a clear triad influencer. Although students working together are assumed to be equals in terms of their statuses and rights to intervene in the interaction, they are rarely equals in terms of other relevant social and cognitive characteristics (Baker, 2015).

Triads with two predominant influencers might be indicating that the third student has been left out, as sometimes happens in collaborative learning (Miyake & Kirschner, 2014), or that he/she has disengaged from the group work (Azevedo, 2015). It is known that low achievers progressively become passive when collaborating with high achievers (Dillenbourg et al., 1996). The fact that in the study all the students were high achievers may suggest that what is important for the group dynamics in this sense is not the absolute level of achievement, but that relative to the other group members.

Balanced triads in terms of influencing members seem to be the ideal theoretical case where all members have comparable roles and are more likely to benefit from the collaboration in a similar proportion (Kuhn, 2015; Resnick, 1991). The results suggest that such triads have reached equity in interaction, a state which is desirable to ensure (Cohen, 1994). In addition, by similarly influencing one another among the triad members, they seem to have reached positive interdependence, one of the basic elements of productive and truly collaborative group work (Johnson & Johnson, 2002).

5.1. Implications

If not properly implemented, collaborative learning has been shown to be ineffective and even counterproductive (Barron, 2003). Collaborative skills require training, as research shows that they are not an automatic consequence of group work (Cohen, 1994). Consequently, an interest in developing metrics to help distinguishing effective from ineffective collaboration, sometimes called collaborative learning analytics, has risen in educational psychology, so that they could be used in training (Kreijns et al., 2003; Wise & Schwarz, 2017). Few instruments in educational psychology capture joint action, and there is a need to provide accounts of interactions that capture the dynamic interplay in collaborative learning (Barron, 2000). The results of this study provide an insight into interaction indicators related to sympathetic arousal, such as direction, level, and contagion, at the triad level in the naturalistic setting of a high school physics course. Such variables and methods here studied, with potential for real time actionable metrics, provide an alternative or complement to expensive microgenetic methods characterizing verbal interactions, which are time consuming and do not scale up well (Wise & Schwarz, 2017). The instantaneous characterization of interactions according to parameters such as commonalities and influence among peers could help to detect fluctuations in attention and coordination in the classroom, features of interaction that are not often described in research on collaborative learning processes (Barron, 2000), as well as providing criteria for optimization of and interventions on group formation (Ahonen et al., 2018), and assisting in the iterative design of tasks which lead to actual group work rather than to alternating task-doers or division of labor.

The theoretical constructs and methods employed in collaborative learning research often neglect time and sequence factors, which is particularly critical for groups working together over weeks and months, as is often the case in academic settings (Reimann, 2009). Computational methods such as physiological computing (Fairclough, 2009), given their high granularity, time resolution, and process orientation, offer particular promise for the theorization of the ways in which collaboration unfolds over time, by facilitating the identification of common and consequential sequences of events (Wise & Schwarz, 2017). This study represents a step in such direction, by providing evidence from the physiological perspective of sympathetic arousal.

It has been argued that groups should become aware of their interpersonal and collaborative processes as they work and take time to discuss how they are doing as a group (Cohen, 1994). To that extent, arousal commonalities and contagion indicators can be provided as feedback on online learning environment dashboards to students and/or teachers, to support awareness of how attuned the group is (cf. Sanna Järvelä et al., 2015), and use it as a trigger for reflection on whether adaptations are needed. To date, such learning-supporting dashboards have largely focused on log data, although the affordances of physiological measures have been acknowledged (Schwendimann et al., 2017). In fact, some recent developments demonstrate how physiological measures can be collected, processed, and visualized on learning analytics dashboards (Schneider, Di Mitri, Limbu, & Drachsler, 2018). However, most of these physiological data supported dashboards purely focus on individual learner feedback (Di Mitri et al., 2018). Dashboards for collaborative learning supported with physiological data are only available on a conceptual level (Praharaj, Scheffel, Drachsler, & Specht, 2018). But the findings from this article can contribute to designing such collaborative dashboard by suggesting relevant indicators such as arousal commonalities and contagion.

5.2. Limitations

The number of E4 wristbands available limited the number of triads ($N = 8$) we were able to track. Furthermore, the triad data available for analysis were reduced due to absence of students or the limited quality of the EDA sensor recordings caused by too tight or too loose wristband contact with the skin. When using triads as the unit of analysis, those factors are more significant since problems in the data of a triad member translate into the inability to use the data of the other two members, as the triad data would be incomplete. As discussed in 3, in practice, the study was able to use 23% of the valid EDA recording sessions. Although the nature of this paper is exploratory, conclusions would have been stronger had data from all triads in all lessons been valid. Further technological developments are needed to detect in real-time the poor quality of the wearable biosensor recordings due to maladjustment so that the amount of data discarded is minimized (Di Mitri et al., 2017).

Nonetheless, the method used is strong in terms of its theoretical grounding, the ecological validity and the duration of the data collection. Naturalistic designs are still rare in studies of interpersonal physiology (Ahonen et al., 2018; Dikker et al., 2017). To answer the research questions, data was collected during two runs of an 18-lesson physics course, each lesson lasting 75 min. Physiological data collection for previous related studies has lasted, for example, five sessions of 45 min each (Kaplan, 1967; Kaplan et al., 1963), two consecutive 20-min sessions (Kaplan et al., 1965), 20 min (Guastello et al., 2006), 45 min (Chanel et al., 2013), 60 min (Ahonen et al., 2016), 90 min (Ahonen et al., 2018), and in the longest duration to the best of our knowledge, 11 classes of 50-min each (Dikker et al., 2017).

It is known that gender plays a role in collaboration (Bear & Woolley, 2011), and that same gender structure is desirable when comparing group measures based on EDA (Boucsein et al., 2012). Unfortunately, due to the predominantly male gender of the students enrolled in the course, it was not possible to compose all triads with the same gender composition, having six triads with one female and two males, and two triads composed of three males. Nonetheless, clear effects of gender in psychophysiological studies have been elusive (Cacioppo & Tassinary, 1990). Research using electrocardiography and/or EDA measures has reported largely comparable responses in men and women resulting in gender having only a small effect (Neumann & Waldstein, 2001), and in other cases has deliberately excluded gender from analysis after initial examination of results, due to inconsistent effect on the variables (Simo Järvelä, Kivikangas, Katsyri, & Ravaja, 2014). Therefore, the reportedly small and inconsistent effect of gender mitigates this limitation.

Finally, the fact that the study was conducted with a student population of high achievers must be considered when it comes to external validity or generalizability of the results to a more heterogeneous sample.

6. Conclusions

This paper explored both symmetrical (i.e., the commonalities in the direction and level of arousal in a triad) and asymmetrical (i.e., the phenomenon of arousal contagion inspired by the notion of emotional contagion) physiological aspects in terms of sympathetic arousal, during collaborative learning in the naturalistic setting of the classroom. The study benefited from the affordances of a biosensor wristband for unobtrusive and continuous measurement of EDA to index sympathetic arousal as a degree of physiological activation, which is one of the dimensions of emotion in the circumplex model of affect and is closely interdependent with the cognitive process of attention. Given the synchronicity and coordination presupposed in collaborative learning, including joint attention, the results provide evidence for the claim that only certain specific phases of group work are truly collaborative. The predominance of arousal contagion on a one-to-one basis (over 70%) suggests that the strongest interactions, including influence, occur mostly at a pair level, even if the team is bigger, a triad in this case. In addition, the study illustrates the potential of physiological measures to enrich and strengthen research on collaborative learning, by exploring a new form of data using a computational approach with high granularity to characterize the process, which is significantly less labor-intensive than traditional approaches, such as conversation and video analyses, and offers affordances which are all welcome and needed in the field (Wise & Schwarz, 2017). Furthermore, the article provides evidence of promising indicators that could be used in learning analytics dashboard to gain new insights into and reflect upon collaborative learning processes.

This exploratory study has provided an ecologically valid picture of group dynamics in terms of arousal direction, level, and contagion among triad members during collaborative learning. However, qualitative studies are needed to further characterize and understand the variations here described, in terms of why a triad reaches full agreement in one lesson but not in another and what situations cause triad members to sustain full arousal directional agreement during several min.

References

Ahonen, L., Cowley, B., Torniainen, J., Ukkonen, A., Vihavainen, A., & Puolamäki, K. (2016). Cognitive collaboration found in cardiac physiology: study in classroom environment. *PLoS One*, *11*, 1–16. https://doi.org/10.1371/journal.pone.0159178

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*, 1–16. https://doi.org/10.1038/s41598-018-21518-3

Arroyo, I., Cooper, D. G., Burleson, W., Woolf, B. P., Muldner, K., & Christopherson, R. (2009). Emotion sensors go to school. In *Proceedings of the Conference on Artificial Intelligence in Education* (pp. 17–24). Amsterdam: IOS Press.

Azevedo, R. (2015). Defining and measuring engagement and learning in science: conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist*, *50*, 84–94. https://doi.org/10.1080/00461520.2015.1004069

Baker, M. J. (2015). Collaboration in collaborative learning. *Interaction Studies*, *16*, 451–473. https://doi.org/10.1075/is.16.3.05bak

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, *37*, 122–147. https://doi.org/10.1037/0003-066X.37.2.122

Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *Journal of the Learning Sciences*, *9*, 403–436. https://doi.org/10.1207/S15327809JLS0904

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, *12*, 307–359. https://doi.org/10.1207/S15327809JLS1203_1

Bear, J. B., & Woolley, A. W. (2011). The role of gender in team collaboration and performance. *Interdisciplinary Science Reviews*, *36*, 146–153.

https://doi.org/10.1179/030801811X13013181961473

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*, 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Boucsein, W. (2012). *Electrodermal activity* (2nd ed.). New York, NY: Springer.

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology*, *49*, 1017–1034. https://doi.org/10.1111/j.1469-8986.2012.01384.x

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2015). *A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments*. Birmingham. Retrieved from https://www.birmingham.ac.uk/Documents/college-les/psych/saal/guide-electrodermal-activity.pdf

Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, *45*, 16–28. https://doi.org/10.1037/0003-066X.45.1.16

Cattell, R. B. (1972). The nature and genesis of mood states: A theoretical model with experimental measurement concerning anxiety, depression, arousal, and other mood states. In C. D. Spielberger (Ed.), *Anxiety: current trends in theory and research* (pp. 115–183). New York, NY: Academic Press.

Chanel, G., Bétrancourt, M., Pun, T., Cereghetti, D., & Molinari, G. (2013). Assessment of computer-supported collaborative processes using interpersonal physiological and eye-movement coupling. In *Humaine Association Conference on Affective Computing and*

Intelligent Interaction (pp. 116–122). Geneva: IEEE. https://doi.org/10.1109/ACII.2013.26

Chanel, G., & Muhl, C. (2015). Connecting brains and bodies: applying physiological computing to support social interaction. *Interacting with Computers*, *27*, 534–550. https://doi.org/10.1093/iwc/iwv013

Cohen, E. G. (1994). Restructuring the classroom: conditions for productive small groups. *Review of Educational Research*, *64*, 1–35. https://doi.org/10.3102/00346543064001001

Critchley, H. D. (2002). Electrodermal responses: what happens in the brain. *Neuroscientist*, *8*, 132–142. https://doi.org/10.1177/107385840200800209

Critchley, H. D., Eccles, J., & Garfinkel, S. N. (2013). Interaction between cognition, emotion, and the autonomic nervous system. In R.M. Buijs & D. F. Swaab (Eds.), *Handbook of Clinical Neurology* (Vol. 117, pp. 59–77). Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-444-53491-0.00006-7

Damasio, A. R., Tranel, D., & Damasio, H. C. (1991). Somatic markers and the guidance of behavior: theory and preliminary testing. In H. S. Levin, H. M. Eisenberg, & A. L. Benton (Eds.), *Frontal Lobe Function and Dysfunction*. New York, NY: Oxford University Press.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2017). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (4th ed., pp. 217–243). Cambridge, UK: Cambridge University Press.

De Lecea, L., Carter, M. E., & Adamantidis, A. (2012). Shining light on wakefulness and arousal. *Biological Psychiatry*, *71*(12), 1046–1052. https://doi.org/10.1016/j.biopsych.2012.01.032

Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., & Cohen, D. (2012). Interpersonal synchrony: a survey of evaluation methods across disciplines. *IEEE*

Transactions on Affective Computing, *3*, 349–365. https://doi.org/10.1109/T-AFFC.2012.12

DeLozier, S. J., & Rhodes, M. G. (2017). Flipped classrooms: a review of key ideas and recommendations for practice. *Educational Psychology Review*, *29*, 141–151. https://doi.org/10.1007/s10648-015-9356-9

Di Mitri, D., Scheffel, M., Drachsler, H., Börner, D., Ternier, S., & Specht, M. (2017). Learning Pulse: a machine learning approach for predicting performance in self-regulated learning using multimodal data. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 188–197). New York, NY: ACM. https://doi.org/10.1145/3027385.3027447

Di Mitri, D., Schneider, J., Specht, M., & Drachsler, H. (2018). From signals to knowledge: a conceptual model for multimodal learning analytics. *Journal of Computer Assisted Learning*, *34*, 338–349. https://doi.org/10.1111/jcal.12288

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., … Poeppel, D. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current Biology*, *27*, 1375–1380. https://doi.org/10.1016/j.cub.2017.04.002

Dillenbourg, P. (1999). What do you mean by “collaborative learning”? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1–19). Oxford: Elsevier.

Dillenbourg, P., Baker, M. J., Blaye, A., & O’Malley, C. (1996). The evolution of research on collaborative learning. In P. Reimann & H. Spada (Eds.), *Learning in humans and machine: towards an interdisciplinary learning science* (pp. 189–211). Oxford: Pergamon. https://doi.org/10.1007/978-1-4020-9827-7_1

Edelberg, R. (1967). Electrical properties of the skin. In C. C. Brown (Ed.), *Methods in* *psychophysiology* (pp. 1–53). Baltimore, MD: Williams & Wilkins.

Elkins, A. N., Muth, E. R., Hoover, A. W., Walker, A. D., Carpenter, T. L., & Switzer, F. S. (2009). Physiological compliance and team performance. *Applied Ergonomics*, *40*, 997–1003. https://doi.org/10.1016/j.apergo.2009.02.002

Empatica Inc. (2015). Recommended tools for signal processing and data analysis. Retrieved March 1, 2018, from https://support.empatica.com/hc/en-us/articles/202872739-Recommended-tools-for-signal-processing-and-data-analysis

Enyedy, N., & Stevens, R. (2014). Analyzing collaboration. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 191–212). Cambridge, UK: Cambridge University Press.

Fairclough, S. H. (2009). Fundamentals of physiological computing. *Interacting with Computers*, *21*, 133–145. https://doi.org/10.1016/j.intcom.2008.10.011

Frijda, N. H. (1986). *The emotions*. Cambridge, MA: Cambridge University Press.

Garbarino, M., Lai, M., Tognetti, S., Picard, R. W., & Bender, D. (2014). Empatica E3 - A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. In *Wireless Mobile Communication and Healthcare (Mobihealth), 2014 EAI 4th International Conference on* (pp. 39–42). Athens, Greece: IEEE. https://doi.org/10.4108/icst.mobihealth.2014.257418

Goyal, N., & Fussell, S. R. (2017). Intelligent interruption management using electro dermal activity based physiological sensor for collaborative sensemaking. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. New York, NY: ACM. https://doi.org/10.1145/3130917

Guastello, S. J., Pincus, D., & Gunderson, P. R. (2006). Electrodermal arousal between participants in a conversation: nonlinear dynamics and linkage effects. *Nonlinear Dynamics, Psychology, and Life Sciences*, *10*, 365–399.

Hanoch, Y., & Vitouch, O. (2004). When less is more: information, emotional arousal and the ecological reframing of the Yerkes–Dodson law. *Theory & Psychology*, *14*, 427–452. https://doi.org/10.1177/0959354304044918

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion*. New York, NY: Cambridge University Press.

Järvelä, S., Kirschner, P. A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, J., … Järvenoja, H. (2015). Enhancing socially shared regulation in collaborative learning groups: designing for CSCL regulation tools. *Educational Technology Research and Development*, *63*, 125–142. https://doi.org/10.1007/s11423-014-9358-1

Järvelä, S., Kivikangas, J. M., Katsyri, J., & Ravaja, N. (2014). Physiological linkage of dyadic gaming experience. *Simulation & Gaming*, *45*, 24–40. https://doi.org/10.1177/1046878113513080

Järvenoja, H., Järvelä, S., & Malmberg, J. (2015). Understanding regulated learning in situative and contextual frameworks. *Educational Psychologist*, *50*, 204–219. https://doi.org/10.1080/00461520.2015.1075400

Jeong, H., & Hmelo-Silver, C. E. (2016). Seven affordances of computer-supported collaborative learning: how to support collaborative learning? How can technologies help? *Educational Psychologist*, *51*, 247–265. https://doi.org/10.1080/00461520.2016.1158654

Johnson, D. W., & Johnson, R. T. (2002). Learning together and alone: overview and meta- analysis. *Asia Pacific Journal of Education*, *22*, 95–105. https://doi.org/10.1080/0218879020220110

Juvina, I., Larue, O., & Hough, A. (2017). Modeling valuation and core affect in a cognitive architecture: The impact of valence and arousal on memory and decision-making. *Cognitive Systems Research*, *48*, 4–24. https://doi.org/10.1016/j.cogsys.2017.06.002

Kaplan, H. B. (1967). Physiological correlates (GSR) of affect in small groups. *Journal of Psychosomatic Research*, *11*, 173–179. https://doi.org/10.1016/0022-3999(67)90004-9

Kaplan, H. B., Burch, N. R., Bedner, T. D., & Trenda, J. D. (1965). Physiologic (GSR) activity and perceptions of social behavior in positive, negative and neutral pairs. *The Journal of Nervous and Mental Disease*, *140*, 457–463.

Kaplan, H. B., Burch, N. R., Bloom, S. W., & Edelberg, R. (1963). Affective orientation and physiological activity (GSR) in small peer groups. *Psychosomatic Medicine*, *25*, 245–252.

Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano, J. (2018). From cognitive load theory to collaborative cognitive load theory. *International Journal of Computer-Supported Collaborative Learning*, *13*, 213–233. https://doi.org/10.1007/s11412-018-9277-y

Knierim, M. T., Rissler, R., Dorner, V., Maedche, A., & Weinhardt, C. (2018). The psychophysiology of flow: a systematic review of peripheral nervous system features. In F. Davis, R. Riedl, J. vom Brocke, P.-M. Léger, & A. B. Randolph (Eds.), *Information Systems and Neuroscience: Gmunden Retreat on NeuroIS 2017*. Springer. https://doi.org/10.1007/978-3-319-67431-5_13

Koole, S. L. (2009). The psychology of emotion regulation: an integrative review. *Cognition & Emotion*, *23*, 4–41. https://doi.org/10.1080/02699930802619031

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: a review. *Biological Psychology*, *84*, 394–421. https://doi.org/10.1016/j.biopsycho.2010.03.010

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in Human Behavior*, *19*, 335–353. https://doi.org/10.1016/S0747-5632(02)00057-2

Kuhn, D. (2015). Thinking together and alone. *Educational Researcher*, *44*, 46–53. https://doi.org/10.3102/0013189X15569530

Lacey, J. I. (1967). Somatic response patterning and stress: some revisions of activation theory. In M. H. Appley & R. Trumbull (Eds.), *Psychological stress: issues in research* (pp. 14–42). New York, NY: Appleton-Century-Crofts.

Miyake, N., & Kirschner, P. A. (2014). The social and interactive dimensions of collaborative learning. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 418–438). Cambridge, UK: Cambridge University Press.

Neumann, S. A., & Waldstein, S. R. (2001). Similar patterns of cardiovascular response during emotional activation as a function of affective valence and arousal and gender. *Journal of Psychosomatic Research*, *50*, 245–253. https://doi.org/10.1016/S0022-3999(01)00198-2

Novak, D., Mihelj, M., & Munih, M. (2012). A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. *Interacting with Computers*, *24*, 154–172. https://doi.org/10.1016/j.intcom.2012.04.003

OECD. (2017). PISA 2015 collaborative problem-solving framework. In *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving* (pp. 131–188). Paris: OECD Publishing. https://doi.org/http://dx.doi.org/10.1787/9789264281820-8-en

Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., & Goodwin, M. S. (2017). Interpersonal autonomic physiology: a systematic review of the literature. *Personality and Social Psychology Review*, *21*, 99–141. https://doi.org/10.1177/1088868316628405

Pecchinenda, A. (1996). The affective significance of skin conductance activity during a difficult problem-solving task. *Cognition & Emotion*, *10*, 481–504. https://doi.org/10.1080/026999396380123

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: a program of qualitative and quantitative research. *Educational Psychologist*, *37*, 91–105. https://doi.org/10.1207/S15326985EP3702

Pijeira-Díaz, H. J., Drachsler, H., Järvelä, S., & Kirschner, P. A. (2016). Investigating collaborative learning success with physiological coupling indices based on electrodermal activity. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16* (pp. 64–73). New York, NY: ACM. https://doi.org/10.1145/2883851.2883897

Pintrich, P. R., Smith, D. A. F., Garcia, T., & Mckeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational and Psychological Measurement*, *53*, 801–813. https://doi.org/10.1177/0013164493053003024

Poh, M. Z., Swenson, N. C., & Picard, R. W. (2010). A wearable sensor for unobtrusive, long-term assesment of electrodermal activity. *IEEE Transactions on Biomedical Engineering*, *57*, 1243–1252. https://doi.org/10.1109/tbme.2009.2038487

Postmes, T., Spears, R., Lee, A. T., & Novak, R. J. (2005). Individuality and social influence in groups: inductive and deductive routes to group identity. *Journal of Personality and Social Psychology*, *89*, 747–763. https://doi.org/10.1037/0022-3514.89.5.747

Praharaj, S., Scheffel, M., Drachsler, H., & Specht, M. (2018). Multimodal analytics for real-time feedback in co-located collaboration. In V. Pammer-Schindler, M. Pérez-Sanagustín, H. Drachsler, R. Elferink, & M. Scheffel (Eds.), *European Conference on Technology Enhanced Learning* (pp. 187–201). Cham: Springer. https://doi.org/10.1007/978-3-319-98572-5_15

Reimann, P. (2009). Time is precious: variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, *4*, 239–257. https://doi.org/10.1007/s11412-009-9070-z

Reiter-Palmon, R., Sinha, T., Gevers, J., Odobez, J.-M., & Volpe, G. (2017). Theories and models of teams and groups. *Small Group Research*, *48*, 1–24. https://doi.org/10.1177/1046496417722841

Resnick, L. B. (1991). Shared cognition: thinking as social practice. In L. & S. T. L., Resnick, J. (Ed.), *Perspectives on Socially Shared Cognition* (pp. 1–23).

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer supported collaborative learning* (pp. 69–97). Berlin: Springer. https://doi.org/10.1007/978-3-642-85098-1_5

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/h0077714

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of Personality and Social Psychology*, *76*, 805–819. https://doi.org/10.1037/0022-3514.76.5.805

Sawyer, K. (2013). Qualitative methodologies for studying small groups. In A. M. O. Cindy E. Hmelo-Silver, Clark A. Chinn, Carol Chan (Ed.), *The international handbook of collaborative learning* (pp. 126–148). Routledge.

Sbarra, D. A., & Hazan, C. (2008). Coregulation, dysregulation, self-regulation: an integrative analysis and empirical agenda for understanding adult attachment, separation, loss, and recovery. *Personality and Social Psychology Review*, *12*, 141–167. https://doi.org/10.1177/1088868308315702

Schneider, J., Di Mitri, D., Limbu, B., & Drachsler, H. (2018). Multimodal learning hub: a tool for capturing customizable multimodal learning experiences. In V. Pammer-Schindler, M. Pérez-Sanagustín, H. Drachsler, R. Elferink, & M. Scheffel (Eds.), *European Conference on Technology Enhanced Learning* (pp. 45–58). Cham: Springer. https://doi.org/10.1007/978-3-319-98572-5_4

Schwartz, M. S., Collura, T. F., Kamiya, J., & Schwartz, N. M. (2017). The history and definitions of biofeedback and applied psychophysiology. In M. S. Schwartz & F. Andrasik (Eds.), *Biofeedback: a practitioner's guide* (4th ed., pp. 3–23). New York, NY: The Guilford Press.

Schwendimann, B. A., Rodriguez-Triana, M., Vozniuk, A., Prieto, L., Boroujeni, M., Holzer, A., … Dillenbourg, P. (2017). Perceiving learning at a glance: a systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*, *10*, 30–41. https://doi.org/10.1109/TLT.2016.2599522

Sharot, T., & Phelps, E. A. (2004). How arousal modulates memory: disentangling the effects of attention and retention. *Cognitive, Affective, & Behavioral Neuroscience*, *4*, 294–306. https://doi.org/10.3758/CABN.4.3.294

Spangler, G., Pekrun, R., Kramer, K., & Hofmann, H. (2002). Students' emotions, physiological reactions, and coping in academic exams. *Anxiety, Stress and Coping*, *15*(4), 413–432. https://doi.org/10.1080/1061580021000056555

Swan, M. (2012). Sensor mania! The internet of things, wearable computing, objective metrics, and the quantified self 2.0. *Journal of Sensor and Actuator Networks*, *1*, 217–253. https://doi.org/10.3390/jsan1030217

Thorson, K. R., West, T. V., & Mendes, W. B. (2018). Measuring physiological influence in dyads: a guide to designing, implementing, and analyzing dyadic physiological studies. *Psychological Methods*. https://doi.org/10.1037/met0000166

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint attention: its origins and role in development* (pp. 103–130). Hillsdale: Erlbaum.

Vogel, F., & Weinberger, A. (2018). Quantifying qualities of collaborative learning processes. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International Handbook of the Learning Sciences* (pp. 500–510). New York, NY: Routledge.

Weinberger, A., Stegmann, K., & Fischer, F. (2007). Knowledge convergence in collaborative learning: concepts and assessment. *Learning and Instruction*, *17*, 416–426. https://doi.org/10.1016/j.learninstruc.2007.03.007

Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning*, *12*, 423–467. https://doi.org/10.1007/s11412-017-9267-5

Zimmerman, B. J., & Schunk, D. H. (2003). Albert Bandura: the scholar and his contributions to educational psychology. In B. J. Zimmerman & D. H. Schunk (Eds.), *Educational psychology: a century of contributions* (p. 504). New York, NY: Routledge.

Acknowledgements

This work was supported by the Academy of Finland [grant number 275440]. The experiment could not have been carried out without the collaboration of teachers, high school students, and administrative personnel from the Oulu University Teacher Training School (Norssi).

Figure1
Click here to download high resolution image

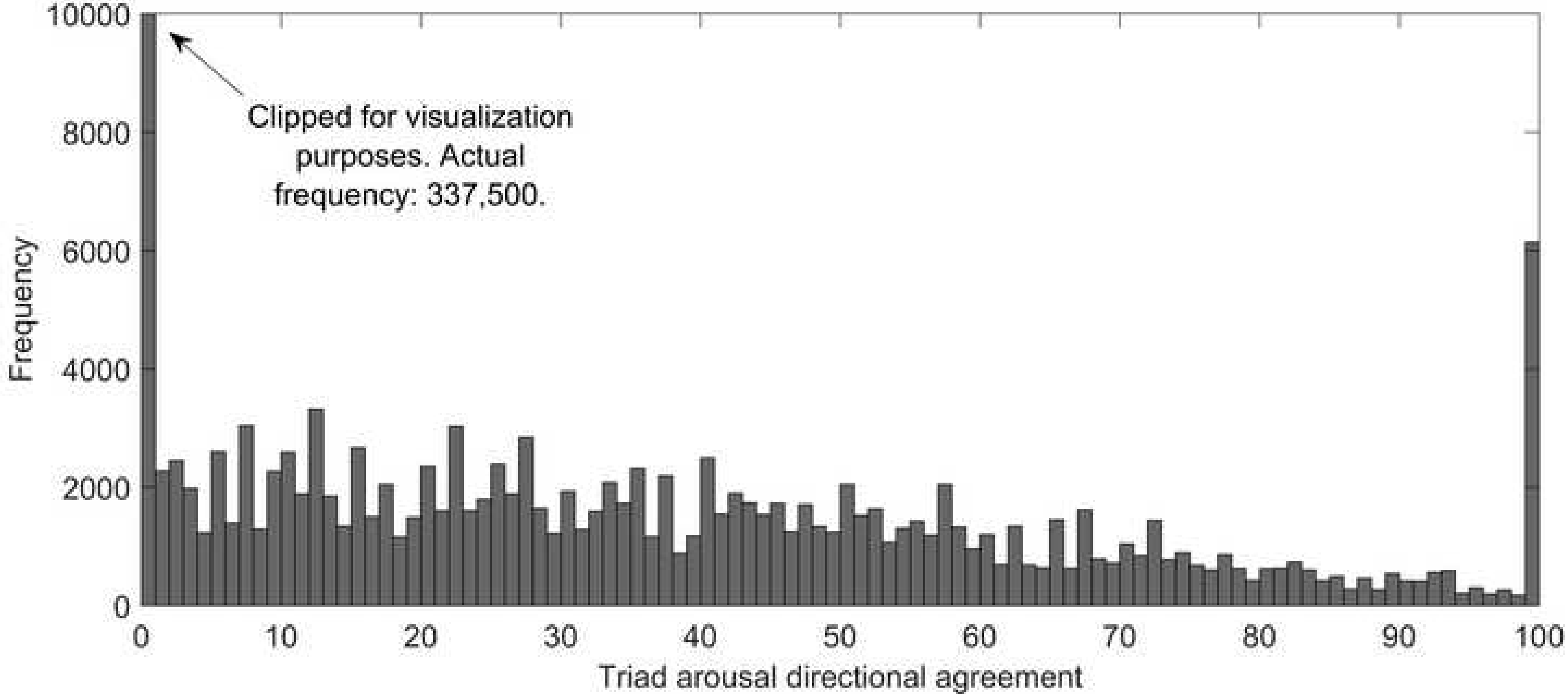

Figure2

Click here to download high resolution image

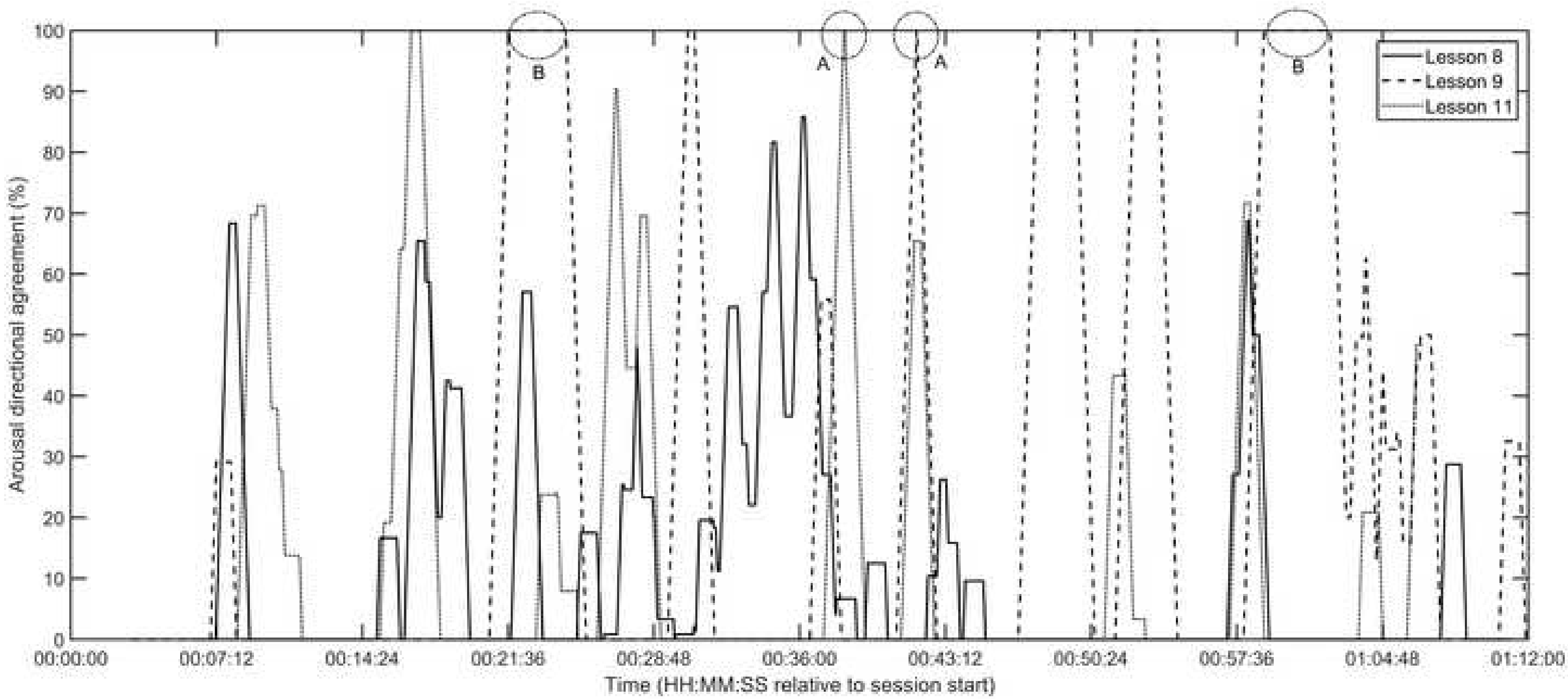

Figure3
Click here to download high resolution image