Devai, Ricardo Emanuel

Implementing process methods in learning research – Targeting emotional responses in collaborative learning

Master's degree thesis

Learning, Education, and Technology (LET)

Faculty of Education

University of Oulu

2019

University of Oulu
Faculty of Education
Ricardo Emanuel Devai: Implementing process methods in learning research
– Targeting emotional responses in collaborative learning
Master's Thesis in Education – Learning, Education, and Technology (LET)
April 2019

Abstract

Context and aim: While interacting with peers and teachers in a collaborative learning task, students experience socio-emotional challenges and display emotional responses. These responses have two major components: arousal and valence, which influence the learning process and its outcomes. The aim of the study was twofold: first, to explore how group members' arousal levels vary across different phases of a collaborative learning task; and second, to investigate how case students' emotional responses are distributed in the arousal-valence space across the phases of the collaborative task. **Methods**: Twelve 6th graders from a school of Finland participated in a collaborative task, in groups of three students. The task was to build an energy efficient house in three distinct phases: brainstorming, planning, and building. While performing the activity, students wore Empatica E4 wristbands to measure their electrodermal activity (EDA) and were video-recorded with 360° cameras. Arousal levels were calculated in peaks per min (ppm) and classified as low, middle, and high. Emotional valence was classified from video analysis into positive, neutral, and negative. **Results**: The ranges for arousal levels were established between 26 and 88 ppm. Only two students displayed the same arousal level across the three phases of the experiment. Three students displayed higher arousal at first and then fell in to lower levels. Four students had the opposite experience and three students did not display a pattern. As for the case students, the student leading a poorly collaborating group experienced oscillating levels of arousal, from middle to high, and displayed a mix of negative and positive valence most of the time. The student loafing around experienced all arousal levels and positive valence most of the time. **Overall conclusions and relevance**: The study allowed to establish measurement thresholds for arousal as a starting point for future studies in collaborative learning and the arousal-valence space provided a quantifiable picture to help teachers understand the importance of emotional responses in classroom during collaborative learning.

Keywords: arousal; collaborative learning; electrodermal activity; emotional responses; valence

Summary

1. Introduction

In a traditional classroom scenario, a student's routine involves acquiring, managing, and retrieving knowledge from different sources throughout school years. Several approaches can be applied to achieve the desired results, from note-taking to extensive reading, from teacher's explanation to web searches. However, learning processes involve more than knowledge management or cognitive functions. As Bruner puts it (1996, p.146): "learning is not simply a technical business of well managed information processing". Earlier studies, such as Piaget's (1971), were directed at cognitive processes, and logic development. Emotional aspects were not entirely appreciated (Immordino- Yang & Damasio, 2007). However, emotions are as important in learning as cognitive factors, since both are closely related (Isohätälä, Näykki, Järvelä, & Baker, 2017). Fischer and Immordino- Yang (2007, p.195) concluded that "simply having the knowledge does not imply that a student will be able to use it advantageously outside of school". Immordino- Yang & Damasio (2007, p. 9) wrote that "the more educators come to understand the nature of the relationship between emotion and cognition, the better they may be able to leverage this relationship in the design of learning environments". To better understand emotional responses taking place during a learning activity, process-oriented measures are helpful instruments to capture and record these responses during the performance of an activity. A process-oriented measure is taken on-the-fly and complements outcome measures, such as evaluations, test scores, and written pieces of work, providing a bigger picture for teachers and administrators.

Along with cognition and emotion, learning is also a matter of social interactions, learning processes, environmental influence, teaching and many other factors (National Research Council, 2000). In school settings, there is one activity combining students' interactions, learning tasks, and teacher's role: collaborative learning, which can be defined as joint efforts to search for understanding, solve problems, create solutions or products, and grasp meanings (Ma. Laal & Mo. Laal, 2012; Smith & MacGregor, 1992). In collaborative learning processes, students often take the lead in organizing themselves to carry out an activity. During the performance of collaborative tasks, students have an opportunity to plan, defend their ideas, give suggestions, agree, and disagree. Much of the learning comes from these interactions among students (Ma. Laal & Mo. Laal, 2012; Panitz, 1999). These interactions may be exciting at times, but they may be also challenging, both in cognitive and emotional aspects.

As we look at collaborative learning processes from a broader perspective, elements such as social, emotional, and motivational outcomes are relevant because they are present throughout the whole activity and will influence both the processes and the results (Baker, Andriessen & Järvelä, 2013; Järvenoja & Järvelä, 2009; Pekrun, 2016). Thus, it is significant to observe carefully how students react during a collaborative learning session, their excitement level and how these emotional responses are expressed over the different learning phases before the outcome.

These collaborative process measures can be captured through the form of self-reports, videos, and physiological responses, just to mention a few. In this study, a multimodal approach, based on 360º video recording and wristband for physiological responses, has been used to offer distinct perspectives and to calibrate the measurement process. In short, this case study aims at exploring students' emotional responses during a collaborative learning task, through a multimodal process measurement.

The study was conducted with twelve students in a primary school in Finland to explore how emotional states vary during collaborative learning. In this study, the two dimensions of emotions have been considered: arousal, which is a body physiological activation, ranging from calm, relaxed to highly excited; and valence, which characterises emotional responses as positive, neutral, or negative. These two dimensions are studied in the context of the circumplex model of affect (Russell, 1980). During the group activity, students wore biosensors to measure their EDA, an electrical property of the skin, suitable to analyse arousal levels (Dawson, Schell & Filion, 2017). Later, their responses during the activity were analysed through video observation to characterise their valence. Finally, the results were plotted onto an arousal-valence space.

2. Theoretical framework

While students are working together, an observer may notice the expression of emotions among the members of the group. Some students may be more competitive, some students just loaf around, and there are students who might show signs of anger, unhappiness, boredom, excitement, all of these while undertaking a collaborative activity. Therefore, it is important to understand the theoretical background of collaboration and emotions felt and expressed by students.

2.1 Collaborative learning as an emotion-eliciting activity

Collaboration has been emphasised as an important skill since the twentieth century (Laal & Seyed, 2012). Available technology has made communication among people faster and easier and has opened new channels to facilitate access to knowledge and information. Therefore, there has been a growth of communities of knowledge and practice, working collaboratively to develop socially constructed knowledge (Driver, Asoko, Leach, Scott & Mortimer, 1994). In school settings, collaboration can yield benefits to a student's development. When successful, collaborative learning promotes critical thinking, social interaction, intrinsic motivation, and performance enhancement (Gokhale, 2012). However, collaborative learning also poses challenges such as coordination, communication, task-related complexity, and episodic memory recall (Andersson & Rönnberg, 1995; F. Kirschner, Paas, & P. Kirschner, 2009). As students work in groups, some elements are critical to generate a positive learning process. According to D. Johnson & R. Johnson (2009), these elements are: positive interdependence, interaction, personal responsibilities, and social skills.

In a collaborative learning environment, "learners are challenged both socially and emotionally as they listen to different perspectives and are required to articulate and defend their ideas" (Ma. Laal & Mo. Laal, 2012, p.1). While interaction takes place and students are focused on solving a problem collectively (transactional and cognitive challenges), collaborative learning also gives rise to emotional challenges (Järvenoja, Volet, & Järvelä, 2013). Thomson and Fine (1999, p. 21) emphasise that: "The understanding of individual cognition and affect must be analysed in relation to relationship with others". A study conducted by Järvenoja & Järvelä (2009) reveals that students face social challenges such as priorities, communication, teamwork, collaboration, and external constraints. Additional studies help understand the challenges faced by students in a collaborative activity. In a study conducted with 48 sixth-graders (Barron, 2003), organized in triads, participants had to solve problems presented in a movie and math quizzes as well. The primary goal was to investigate the differences in performance to solve problems, considered that all the 48 students were high-achievers. The conversation of four triads were analysed qualitatively. Barron described the results of less successful cases as follows: “relational issues arose that prevented the group from capitalizing on the insights that fellow members had generated. These included competitive interactions, differential efforts to collaborate, and self-focused problem-solving trajectories. Behaviourally, these issues manifest in violation of turn-taking norms, difficulties in gaining the floor, domination of the workbook, and competing claims of competence” (Barron, 2003, p. 42). On the other hand, successful groups displayed a behaviour characterized by low level of ignoring or rejecting ideas, higher attention levels among the group, and mutual eye contact. In another study conducted with four girls and four boys in a US school (Vauras, Iiskala, Kajamies, Kinnunen & Lehtinen, 2003), two high-achieving girls were chosen for case analysis. They had to work out through word and math problems in a mathematics instructional game. The collaboration was of a good quality and the girls showed task-orientation, persistence, attention focus. Both girls recognized how thinking together and telling aloud was helpful during the learning and problem-solving process. They also appreciated their partner´s skills and ability to solve a problem. Negotiation was also a very important skill in the process. From this experiment, those challenges associated with attention, communication, negotiation can be clearly seen, and they are fundamental for a successful collaboration. Another example comes from a study conducted with 107 first year higher education students (pre-service teachers), from Finland. The goal was to find out about cognitive, motivational, and emotional challenges faced by students both in individual and collaborative learning situations (Koivuniemi, Panadero, Malmberg & Järvelä, 2017). The main challenges in collaborative learning regarding motivation were related to time management, cooperation, and concentration. As for cognitive learning challenges, prior knowledge and knowledge construction were reported as the higher ones. Finally, about emotional challenges, the major ones were frustration, fear and excitement, and inferiority.

In collaboration, emotion is a very important dimension since students are interacting with each other constantly and “such activities tend to evoke emotional responses, both positive and negative” (Jones & Issroff, 2005, p. 401). Järvenoja and Järvelä (2009) conducted a study with 63 teachers, from educational programs, in Finland, to find out socio-emotional challenges experienced by students in a collaborative learning situation.

Researchers have selected two collaborative learning tasks from two different groups, made up of four members each. Concerning social challenges, there were different responses. In one of the groups, participants said that emotional responses came from challenges in collaboration, external constraints, and teamwork. In the second group, personal priorities, communication, and teamwork were mentioned as socio-emotional challenges.

From the previous examples, one can see that competition, communication, acceptance, attention, motivation, teamwork can all trigger emotional responses in a collaborative learning setting. As mentioned by Do and Schallert (2004, p. 631), affect is a "critical part of students' experience in the context of a classroom discussion […] a catalyst of the thinking and actions". And these emotions may yield "both functional and dysfunctional behavioural consequences for teams" (Garcia- Prieto, Bellard & Schneider, 2003, p. 8). It can be concluded that emotions play an important role in learning in general and for collaborative learning, in particular. That is why in the next chapters the research on emotions and their role in academic settings as well as the related methodological advancement in this field is elaborated to provide the theoretical ground for the current study.

2.2 Emotions in academic settings

Pekrun (2016, p. 120) recognizes that "few studies have examined students' emotions as they occur in academic settings". Despite that, over the past fifteen years there has been an increase in the number of studies related to emotions in school (Pekrun, 2016). These studies are particularly relevant because emotions will "direct interactions, affect learning and performance, and influence personal growth in both students and teachers" (Phye, Schutz, & Pekrun, 2011, p.13).

Findings of qualitative studies have demonstrated that students experience a very intense and diverse emotional life in academic settings. The expression "academic emotions" has been taken from and it refers to emotions taking place within academic contexts (Pekrun, Goetz, Titz, & Perry, 2002a). For example, a study shows positive correlation in the following context: initially, students had an increase in physiological activation (arousal) while feeling anxiety during an exam; afterwards, there was a reduction of activation when students could properly cope with the situation (Pekrun et al., 2002a).

Academic emotions derive from learning processes and achievement. They bring about emotions related to individuals, tasks, and social environments (Pekrun, Goetz, &

Perry, 2002b). Since academic performance and their corresponding results have to do with status and resources allocation, diverse emotions originate in competitive circumstances, either among individuals or groups. In the end, the sum of individual emotions informs about the group's emotional climate because there is a reciprocal causation "between emotions, learning, and achievement" (Pekrun et al., 2002a). Students working in a group are constantly evaluating or appraising "their relations to their environments in terms of their implications for their well-being" (Cahour, 2013). For example, a student may consider a group work situation as fun, boring, or interesting. Depending on how a student perceives the situation, he or she will be influenced by the emotions in the environment. When other students in a group realize what emotion a particular student is feeling, that episode will also influence the overall emotional climate of the group. This reciprocal influence can be seen in a relationship between teachers and students. If a teacher is enthusiastic about a certain topic, it will trigger students' enthusiasm about the topic as well. Students, in turn, can propel a teacher's positive emotional state by showing interest and enthusiasm (Cahour, 2013).

Considering the individual, a positive state can improve creativity by generating "more associations between ideas, more unusual verbal associations, and produce a more rich context for thinking and … cognitive flexibility" (Baker et al., 2013, p. 61). Although this positive state is seen as efficient in problem solving, Schwarz (2002) points out that negative affective states are preferable for analytical tasks and causal reasoning because they can help the student focus and process more deeply. Since these results are obtained in experimental settings, they should be referred to with caution (Cahour, 2013).

Research has shown that in addition to learning processes, emotions are also related to performance. These emotions could be prospective, such as test anxiety; or retrospective, connected to prior success or failure (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011). While aiming for performance, students feel pressure for achieving desired results and experience expectancies of failure. The pressure and the expectancies are reported as major triggers for emotional arousal (Pekrun et al., 2002b).

In other words, in the best-case scenario, emotions related to learning and performance will bring about beneficial results, even in challenging situations. For example, a student who feels sad about a failure might tap from the negative experience to overcome a future and challenging situation (Pekrun, 1992).

In summary, there are several triggers of emotional responses, due to multiple processes involved. As Mauss & Robinson puts it (2009, p.14), "emotions are constituted by multiple, situationally and individually variable processes". However complex, "both nonverbal behaviour (e.g. facial and vocal expression) and physiological indicators can be used to infer the emotional state of a person, [while] there are no objective methods of measuring the subjective experience of a person during an emotion episode" (Scherer, 2005, p. 712). Thus, physiological indicators can help the process of inferring emotional states.

2.3 Academic emotions and collaborative learning

Although emotions are identified at individual levels, they are triggered in social contexts, especially when students interact with teachers and peers (Schutz, Hong, Cross, & Osbon, 2006). In other words, emotional episodes are not taking place independently or in an isolated manner, belonging exclusively to an individual or to the environment. There is a constant interaction among students, environment (other students and teachers), and the socio-historical context in which these relations take place (Schutz, Aultman & Williams-Johnson, 2009). Empirical research on social and emotional learning programs has shown that, at group level, providing emotional support to students helps them set higher expectations, encourages bonding in classroom and engagement in collaborative learning, as revealed by a research project based on 213 studies, conducted with 270,034 students from kindergarten to high school (Durlak, Weissberg, Dymnicki, Taylor & Schellinger, 2011). Emotional support is relevant because when students work in groups, they experience socio-emotional challenges (Thompson & Fine, 1999) and the way these challenges are dealt with will contribute to a richer or poorer collaboration experience, both at individual and group level. In a study conducted by Järvenoja, Malmberg, Järvelä, Näykki and Kontturi (2018), emotional states and motivation fluctuation were investigated in a two-month long science project with 20 students. There were teacher-led, individual, and collaborative learning activities. The experiment investigated students' emotional states across a learning project. These states were measured with EmAtool (Emotion Awareness Tool), at the beginning of a learning session, and the students' evaluation of emotional states were reported as positive, neutral, and negative. The results showed that "emotional states are affected by various situational aspects, such as task, teacher, pedagogical design, and social context" (Järvenoja et al., 2018, p. 15).

Another study, conducted by Andriessen, Pardijs & Baker (2013), focused on the socio-emotional dimension of collaborative group meetings. A group of 13-old-year boys had to design a town area in which they would like to live in. The researchers confirmed that "during collaborative learning interactions, epistemic progress is subject to social/emotional dynamics" (Andriessen et al., 2013, p. 206). The students expressed their ideas, but the group agreement, well-being, and the appearance of being competent were dominant over the activities related to knowledge. Another study also refers to the domination of social dimension over the cognitive one (Andriessen, Baker & van der Puil, 2011), which helps us see the dependable relationship of epistemic and emotional aspects in a given collaborative activity. Andriessen et al. (2013) described the boys experiencing mild tension when disagreeing about project characteristics or lack of clarity about the assignment; however, these tensions did not negatively affect the group. The group did not display any tense emotions either within the group or towards the environment. Conversely, the relaxing actions, such as agreement, ignoring issues, humour were stronger and beneficial to the group. This intense emotional expression is confirmed by Järvenoja and Järvelä (2009, p. 465) as emotions expressed during students' interactions in collaborative settings are usually "higher than in conventional learning situations".

Näykki, Järvelä, Kirschner and Järvenoja (2014) conducted a study that reported challenges experienced by groups in collaborative learning. They found out that the groups most frequently expressed socio-emotional challenges, whereas cognitive challenges were the least reported. As an example of socio-emotional challenges, the authors described that "the case group's collaborative learning was endangered as they (students) employed avoidance-focused strategies to regulate socio-emotional challenges and restore emotional balance within the group. The group members escaped the unpleasant situation by withdrawing from the group activities and by concentrating on secondary activities" (Näykki et al., 2014, p. 12). In the end, they were able to finish their task (met a cognitive goal), but had difficulty dealing with emotional aspects of the task.

There is a greater number of researchers conducting studies on the impact of emotions on learning (Järvenoja & Järvelä, 2009; Schutz et al., 2006), and more recently there are researchers interested in studying emotional experiences in collaborative learning situations (Näykki et al., 2014). There is also an advance in technology-enhanced learning, but there are less successful research results related to socio-emotional engagement (Järvenoja et al., 2018).

2.4 Emotions and its components

To progress in research on emotions in academic settings in general and particularly in collaborative learning, in last few years, learning researchers have started to implement various methods for investigating the role of emotions as a part of the learning process. In order to do that, there is a need to move forward from recognising and specifying different emotions (such as joy, pride, or anxiety) and break down the components of emotional responses that simultaneously can have (meta)cognitive, motivational, and physiological origins. An emotional response is prompted when an individual assesses an event that happened before, at a personal level (Garcia- Prieto et al., 2003; Scherer, Schorr & Johnstone, 2001). The appraisal process takes place at conscious or unconscious levels and responses are triggered by "component systems, such as subjective experience, facial expression, cognitive processing, and physiological changes" (Fredrickson, 2001, p. 2). As an example of subjective experience, emotions may arise from past events and the origin is unclear to the individual. Concerning cognitive processing, it involves processes such as memory, attention, problem-solving. Stored knowledge and interest level will also trigger emotional responses in the individual or while working collaboratively. As for physiological changes, the body may sweat, the pupils become dilated (Scherer et al., 2001).

From another perspective, components of emotion can be organized into three different categories: physical component, which has to do with physiological changes in the body; the cognitive process, which qualifies the emotion felt and it is related to the appraisal of an event; and behavioural signs, which are the outward expression of the emotion. If an individual decides to fight or run away after assessing a situation (either consciously or unconsciously), the actions resulting from race or fight decision are an outward expression of an emotion (Pally, 1998).

Since the terms emotion and affect will be presented throughout this study, it is important to observe the main differences between these concepts, based on Fredrickson's (2001) (Table 1). There is also the difference between affect and feeling. A feeling is a sensation a person has, generated by previous events, and it is personal. Two people may have different sensations when facing the same event and, therefore, name their feelings differently. Concerning affect, this is more challenging to label, it is non-conscious, and it is related to intensity. In short, affect "is the body's way of preparing itself for action in a given circumstance by adding a quantitative dimension of intensity to the quality of an experience"

(Shouse, 2005, p. 5). Therefore, emotion, affect, and feeling can be used to characterise different emotional experiences.

Table 1

Main differences between emotion and affect

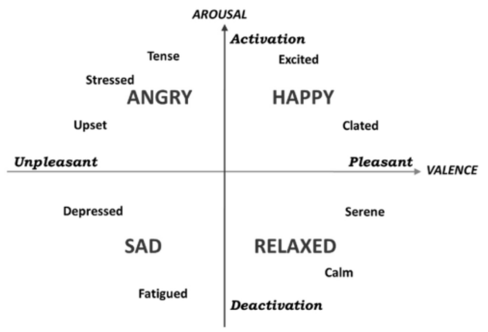
Emotion	**Affect**
Personally meaningful circumstance	Free-floating
(there is an object)	(objectless)
Brief, multicomponent systems	Long-lasting and related to subjective experience
Conceptualized into categories such as joy, anger, fear	Varying along two dimensions: activation (arousal) and valence

2.5 Arousal and valence – components of emotion

Besides looking at the origin and effects of emotion from the standpoint of categories, such as physical, cognitive, and behavioural, there is another possibility to characterise emotion by looking at its dimensions. As described by several authors (Barrett, 2006 a, b; Linnenbrink, 2007; Russell, 1980), the way a person feels at a certain point in time can be characterised by the combination of two dimensions of emotion: arousal and valence. Arousal is related to the body excitement level and ranges from low to high. Valence is related to attractiveness (positive valence) and averseness (negative valence). While arousal is characterized in a continuum between low and high activation levels, valence can be classified from positive to negative, including a neutral state. In other words, valence is how we judge a situation and arousal is the degree of excitement (Chanel, Ansari-Asl, & Pun, 2007).

There are attempts to identify more dimensions other than arousal or valence, but due to difficulties related to consistency, a two-dimensional space consisting of arousal and valence has been used by several authors to map emotions (Feldman, 1995; Nicolaou, Gunes & Pantic, 2011; Nummenmaa, M. & Nummenmaa, L., 2008; Ringeval et al., 2015; Scherer, 2005; Zhang, Tian, Jiang, Huang & Gao, 2008). Barrett and Russell (1999) have proposed that valence and arousal are the two most important dimensions of emotions and these two dimensions can be used to classify student's emotions (Pekrun, 2006). This two-dimensional space is represented by Russell’s circumplex model of affect (Figure 1) (Barrett & Russell, 1999; Feldman Barrett & Russell, 1998; Linnenbrink, 2007; Russell, 1980).

Figure 1. Russell’s circumplex model of affect



Note. Art published in Valenza, Citi, Lanatá, Scilingo & Barbieri, 2014

While students are working collaboratively in groups, they experience different levels of arousal and valence, which in turn influence their learning process and its outcomes. Positive or negative valence can affect the scope of attention, cognitive flexibility, creative problem solving, and perceptual processing (Isen, 1999), whereas arousal is closely related to attention and cognitive processing (Critchley, Eccles & Garfinkel, 2013). Since these two components of emotion characterise emotional responses and directly affect the learning process, a careful study of each component is required to understand its expression and effects in academic settings.

2.5.1 Arousal

Arousal is a state of being awoken, activated. It can be traced down to three main sources: earlier knowledge stimulation, physiological processes, and stimuli from external world (Luria, 1973). These sources are closely interwoven. A student may be stimulated by novelty (external world) and consider performing the task based on earlier experiences. In the meantime, bodily and unconscious processes (physiological/metabolic) are taking place.

In terms of variation, high arousal describes higher levels of activation, stress, excitement, and low arousal refers to a state of relaxation, boredom. There is also a moderate level of arousal mentioned by Schlosberg (1954). He mentions that a person may need a moderate level to play chess and higher levels for sprinting. Thus, optimal levels of arousal may vary according to the person and to the task (Oxendine, 1970; Schlosberg, 1954; Zuckerman, 2014). The Yerkes-Dodson principle (Blair, 2010) suggests that there should be a moderate or optimal arousal level for every person. In Psychology, the Yerkes-Dodson principle describes the relationship between arousal and performance. This law suggests that when you are highly-aroused, your performance can be improved. However, if you are excessively highly-aroused, your performance may be jeopardized. For example, this law would explain the reaction of students taking an important university admission test. Highly-aroused students may be alert and activated to take the test and obtain good results. If this alertness is excessive, it may lead to "blanks" and stressful situations, impairing the performance during the exam. Duffy (as cited in Zubek, 1969, p.410) says that "in the same stimulus situation there are differences between individuals in the degree of arousal". And the differences in arousal are more related to personality traits than a single physiological system (Blachly, 1970).

D'Mello, Hays, Williams, Cade, Brown and Olney (2010) conducted a study with 90 college students from a mid-south university in the United States. The purpose of the study was to create a computer-based tutoring system, resembling human tutors, in which students could have different types of dialog with the system, including collaborative activities. After each lecture, student engagement was measured through an affect grid, based on arousal and valence. The results showed that arousal and not valence was the most important for learning gains; arousal has been correlated with deep learning gains. The results also indicated that arousal was higher in situations in which collaborative lecture was present rather than less interactive options.

In another study, conducted with online classes, students participated in two synchronous and two asynchronous online discussions. The results indicated that the synchronous discussion induced personal participation. This has been considered as a complement to cognitive involvement. There was a feeling of working together, in groups, in the synchronous version of the online discussion. As a result, psychological arousal has been increased as well as motivation and convergence on meaning (Hrastinski, 2008).

In short, arousal is an important component for motivation, engagement, and learning involvement and gains. Arousal characterises the level of enjoyment, excitement, activation of a student regarding a specific task or context. However, it is also important to understand whether the activation level is related to a positive or negative experience. This will be indicated by valence, another dimension of emotion.

2.5.2 Valence

As a dimension of emotion, valence "meets the criteria to be a useful category that will support the scientific study of emotional processing. First, valence is a basic, invariant building block of emotional life that can be observed in self-reports of experience as well as in virtually all instrument-based measures of emotion" (Barrett, 2006b, p. 50). For instance, positive valence may include adjectives such as enjoyment and hope whereas negative valence is related to frustration, boredom, and hopelessness. In a learning context, valence "reflects the attraction to particular learning goals" (Garrison, 1997, p. 27). This attraction is related to personal needs and preferences. Personal needs will inform if a specific goal is relevant to the student and preferences are related to emotional states at a certain point in time, reinforcing the concept that valence is not a fixed characteristic of the individual or the situation per se, but it varies according to the circumstances and it can be best captured as referring to a certain point in time (Barrett, 2006b).

In a study focused on observing university students' emotional responses in a web-based learning environment (Nummenmaa, M. & Nummenmaa, L., 2008), including collaborative activities, valence was measured through the Self-Assessment Manikin (Bradley & Lang, 1994), based on an HTML page. In the study, students had to select a face ranging from smiling happy figure (pleasant) to a frowning unhappy figure (unpleasant) every time they logged off an activity. The results showed that students felt positive emotions online in both situations: when other students were present and posted activities, and when there was no one and students had to work individually. Thus, the online environment itself seemed to be positive to the students. The study also revealed that students not doing much (i.e. lurkers) had higher negative emotional levels.

Another study, about socially shared regulation in collaborative learning, was conducted by Isohätälä, Järvenoja & Järvelä (2017). The participants were 24 student teachers in six small groups. They were video-recorded during collaborative activities in mathematics. One of the findings shows that students expressed a strong socio-emotional valence, indicated by group cohesion, goal externalization, and verbal and non-verbal positive communication.

A study conducted by Laniado, Kaltenbrunner, Castillo and Morell (2012) aimed at finding out about emotions felt in a peer-production environment, the Wikipedia case. The authors selected approximately 12,000 editors who had actively contributed, considering that he or she had written at least 100 comments on Wikipedia. The emotional content was measured by the words used, according to the Affective Norms for English Words (ANEW). There was a narrow range for valence because the articles are formal and there is not much controversy. One of the relevant findings has to do with tone. "in Wikipedia, experienced contributors tend to have a more positive tone (valence)" (Laniado et al., 2012, p. 5). These studies emphasise the importance of valence as an indicator of the quality of emotional states. They are related to the situation being experienced by participants rather than a fixed attribute of a person or a task.

2.5.3 Valence-arousal space

While looking at these dimensions of emotion, arousal and valence, it is possible to characterise a student's learning experience by observing and categorizing their emotional responses into a valence-arousal space. In this space, four areas can be identified: positive and negative activation and positive and negative deactivation, as described earlier in Figure 1. For example, a positive activation (positive valence and high arousal) can be illustrated in the following situation: a student is sure to have achieved a good result in an assessment and the results are about to be announced. There is a high arousal anxiety on the side of the student and his or her family and friends, and a neutral valence, a stand-by moment. If the student is successful indeed, a positive valence will fill in everyone. On the other hand, a positive deactivation (positive valence and low arousal) can be exemplified in the following situation: a humorous ad dealing with a problem, trying to bring a positive focus (positive valence) to a real and negative problem (deactivating). As for negative activation, it means you are pushing to an audience a set of negative messages, creating anxiety and tension. Finally, for negative deactivation, it may be a threat to a student, in which depressing feelings and hopelessness will rise. Therefore, arousal and valence can be used to classify students' emotions.

One study created an algorithm to analyse faces in a group of static images, based on the arousal-valence space. They first analysed the images and created an annotated database with valence (positive, neutral, and negative) and arousal (high, medium, and low). From the database, they created a framework for facial analysis and obtained a result of 54% for valence recognition and 55% for arousal dimension (Mou, Celiktutan & Gunes, 2015). Another study, also focused on automated processes, has proposed a model to evaluate valence and arousal dimensions in interactive video games. The experiment was conducted with 24 male participants. They had to interact with other players and with the computer as well. Using the Affect Grid, a tool based on the arousal and valence space, the participants rated their emotions after each play condition. As a result, the authors proposed a mean emotion modelled from physiological data to evaluate users' emotions while playing games. They also suggested video analysis to compare the emotional responses at certain points of the game so that the corresponding time can be seen in the video recording (Mandryk, Atkins & Inkpen, 2006).

To conclude, studies have shown the role of emotion in learning (Cahour, 2013; Pekrun, 2016; Pekrun et al., 2002a; Pekrun et al., 2002b; Phye et al., 2011), indicating that emotional responses and states are as important as cognitive aspects. These emotional responses are expressed as a multicomponent system. For example, a student who is worried and anxious about a task may feel negative feelings (affective) and the student may be wondering how to solve it (cognitive). The student may feel an urge to escape the situation (motivational) and, at the same time, there are bodily signs and activation undergoing the body (physiological).

3. Methodological considerations

Since there are different methodologies, instruments, and bio signal devices to capture process-oriented approaches, it is important to summarise the advantages and limitations of each one to narrow down to the methodological choices selected for this study. One aspect to be considered additionally is the fact that the students are in an authentic academic setting, performing a collaborative learning activity. Thus, the least obtrusive methods are suitable to avoid interference as much as possible. Self-reports, think aloud protocols, video analysis, and physiological measurement are described in the following sections. A more detailed explanation about the method selected for this study is explained in Chapter 5.

3.1 Self-reports or process-oriented measures for studying emotions in learning

Self-reports in the form of questionnaires, surveys, smiley sheets have been generally used to measure emotional aspects in academic settings (Panadero, Klug &, Järvelä, 2016; Zimmerman, 2008) as well as areas such as perceptions of self-efficacy beliefs (Bandura, 2012) and those dealing with large amounts of data (Barnes, 2013). Several self-report instruments have been used, such as Academic Emotion Questionnaire, AEQ (Pekrun et al., 2011); Epistemic Emotion Scales, EES (Pekrun, Vogl, Muis, & Sinatra, 2017); Pleasure, Arousal and Dimension (PAD) Scale (Mehrabian, 1996); International Affective Picture System, IAPS (Lang, Bradley, & Cuthbert, 1997). A clear advantage of self-reports is that a standardized test would be beneficial for data collection efficiency. Besides, self-reports capture opinions from the student's standpoint. However, there are some disadvantages. One of them is that certain words may suggest answers that the participant would have not chosen if these words were not on the list. Conversely, the participant may need a word, which is not on the list. Another limitation is that lists of traits may not capture changes in students' strategies as a result of teacher's intervention (Panadero et al., 2016; Scherer, 2005; Veenman, 2005; Winne & Perry, 2000). Also, some self-reports are based on what the student plan to do or did during certain activities and these perceptions do not match with what really happened during learning (Zimmerman, 2008).

In order to capture different moments of a study session and observe what really happened, there are process-oriented measures, which are complementary and allow for greater calibration. Process-oriented instruments collect data while the activity is in progress, capturing actions, facial expressions, and verbal and non-verbal signals from participants.

Some traditional methods to assess a process include think-aloud protocols (participants think aloud while performing a task), and video analysis.

Think aloud protocols have the advantage of capturing thoughts in real time while keeping the learner engaged in a real activity, rather than a fictitious situation. On the other hand, there are limitations regarding the verbalisation process and hearing one's own voice, which might affect the outcome (Fonteyn, Kuipers & Grobe, 1993). Video analysis is a process in which a videotaped session can be used for further analysis, codification, and reports. There are several methods to use video: from the hands of the subject (participatory), video elicitation, and video-based fieldwork. One of the main advantages is that it captures temporal sequential interaction. In a study, "video records enriched the understanding of group interaction as a moment-by-moment process reflected through members' intonation, facial expressions and body language, in addition to conversation" (Näykki et al., 2014, p. 12). A common limitation from videos has to do with collecting large amounts of rich data and the results are too much descriptive and less analytical (Jewitt, 2012).

All these process-oriented methods carry a certain degree of subjectivity either by the subject of the research or the researcher in charge. The advent of new bio signal technologies has enhanced these methods. These technologies allow researchers to capture and measure physiological responses in an unobtrusive manner. "As these measures are biological in nature, they are not subject to social desirability or any other associated concern" (Mauss & Robinson, 2009, p. 3). Thus, these instruments can "capture what may be beyond a respondent's conscious control and understanding" (Ciuk, Troy, & Jones, 2015, p. 4).

One of these instruments are wearable biosensors. They have provided the opportunity to carry out studies in several areas such as sleepiness (Goh, Galster & Marcus, 2000; Loredo, Clausen, Ancoli-Israel, & Dimsdale, 1999; Loredo, Ziegler, Ancoli-Israel, Clausen, & Dimsdale, 1999), psychological disorders (Cattell, 1972; Fan, C. Chen, S. Chen, Decety & Cheng, 2013; Thayer, Takahashi & Pauli, 1988; Woods & White, 2005), sexual arousal (Conrad & Wincze, 1976; Haywood, Grossman, & Cavanaugh, 1990; J. Berman et al., 2001), and music and gambling (Baumgartner, Willi & Jäncke, 2007; Moodie & Finnigan, 2005; Rockloff, Signal, & Dyer, 2007). These technologies generate trace data, which are multiple data points collected over a period of time rather than longer intervals.

3.2 Trace data

Trace data is obtained from transaction logging software, based on processes that are unobtrusive and less subject to bias since they record physiological responses, which can be plotted and analysed afterwards. As examples of trace methods, there are: EDA measurement, heart rate, facial expression recognition. These methods can be used independently or in combination with other types of measures, such as in this study with video data. In the past, it was very time consuming and costly to gather trace data, however with new and less expensive technologies, the collection process has become more popular and economically viable.

Trace data has three relevant components for further study and data analysis; logs, temporality, and granularity. Logs are the electronic record of interactions between a system and its respective users. Temporality is the period chosen for the study. It could be a whole session, measured in hours, or it could last days, depending on the aim of the study. Finally, granularity is the smallest item recorded in the experiment. It shows the level of detail for every recorded unit.

It is recommended that "researchers continue to integrate interdisciplinary methodologies to capture engagement using trace methodologies, such as log-file data, eye-movement data, physiological measures, and self-report measures of engagement processes" (Azevedo, 2015, p. 89).

3.3 Trace methodologies - physiological measures

When a student faces challenges in a social context, for instance, physiological responses are triggered beyond the conscious control, as explained earlier in Chapter 2. These responses originate in the Autonomic Nervous System (ANS), a physiological system made up of nerves and cells, responsible for the innervation of blood vessels, and the airways, heart, intestines, and urogenital organs (Gabella, 2001). There are sympathetic and parasympathetic branches in the ANS. The sympathetic is associated with activation and the parasympathetic with relaxation (Mauss & Robinson, 2009). The usual indices assessed for ANS activation are electrodermal (through sweat glands) and cardiovascular responses (through blood circulatory systems). Electrodermal responses are typically classified into two types: tonic skin conductance level (SCL) and phasic skin conductance response (SCR). The tonic type is the slow-changing levels, quite smooth and in the background, while the phasic one is the rapidchanging peaks, the changes in electrical conductivity of the skin (Braithwaite, Watson, Jones, & Rowe, 2013; Mauss & Robinson, 2009).

Cardiovascular and electrodermal systems have been used as data source for the study of effects of emotion on ANS activity, as described in an extensive study by Kreibig (2010). Mauss and Robinson (2009, p.5) suggests that "measures of emotional responding appear to be structured along dimensions (e.g., valence, arousal) rather than discrete emotional states (e.g., sadness, fear, anger)". Among the emotional dimensions presented in several studies, arousal and valence have been considered to be the major ones (Boucsein, 2012; Ciuk et al., 2015; Pijeira-Díaz, Drachsler, Kirschner, & Järvelä, 2018; Spangler, Pekrun, Kramer, & Hofmann, 2002; Taylor & Epstein, 1967).

3.4 Electrodermal activity and arousal measurement

Since arousal originates in the ANS, a physiological measure has been considered to explore students' arousal levels during the task. This is EDA, the electrical variation in the skin conductance. It is widely considered a reliable measure for arousal (Boucsein, 2012). It works like this: as the sympathetic system reacts to a stimulus, arousal levels are heightened and sweat glands of the skin are innervated (activated). This arousal is indicated by SCRs or peaks. Boucsein (2012, p. 377) mentions an investigation conducted by Greenwald, Cook, and Land in which "EDA appeared as a sensitive and specific measure of arousal", when compared to similar results in studies conducted by Winton, Putnam and Krauss (1984), and Johnsen, Thayer, and Hugdal (1995). EDA has been consolidated as a valid measure for arousal and it is considered a "clean measure of sympathetic arousal" (Pijeira-Díaz et al., 2018, p. 4). The present case study is focused on sympathetic arousal, hereinafter simply called arousal.

EDA is a signal related to the intensity of arousal. Eiser (1986, p. 197) wrote that arousal "determines the intensity, but not the quality, of an emotional state." It does not indicate if the state is positive or negative (i.e., valence). A person might feel highly aroused in a positive way (e.g., excited, happy, alert) or in a negative way (e.g., upset, nervous, tense) as described in the schematic map of core affect (Russell & Barrett, 1999). For characterizing emotions, it is important to analyse contextual data with EDA (Immordino-Yang, Christodoulou, Pekrun, & Linnenbrink-Garcia, 2014). Since arousal is a complex and multidimensional phenomenon (Boucsein, 2012), data triangulation can be used to generate more contextualized measures (Panadero et al., 2016). Boucsein (2012, p. 370) also states that the "in classic psychophysiological emotion research, emotional quality is determined via subjective variables while their quantitative properties are measured by ANS parameters."

In terms of length measured, studies with EDA are generally related to discrete stimuli, continuous stimuli, or long-term setting. Discrete stimuli mean that "a subject may be presented with a visual cue, tone, image, mechanical, electrical, olfactory, light stimulation, etc." (BIOPAC Systems, Inc., 2018, para. 1). Continuous stimuli are for longer periods, such as physical and mental performance, different forms of interaction. Long-term settings are related to traits and can be seen in studies with epileptic individuals, sleeping patterns, psychopathology (Henriques, Paiva & Antunes, 2013).

A single method for measuring and analysing EDA in relation to arousal levels, in a collaborative learning situation, has not been identified in the literature. Currently, different methods have been developed, resulting in a situated and context-based approach. A study conducted with 30 participants, for instance, aiming at finding out to which extent EDA can be used to recognize emotions and the variation of EDA between human-human and human-robot interaction applied a method called "FineAlly" for emotion recognition. Individuals were also classified according to the Myers-Briggs type indicator (a temperament test). One of the results showed that emotion recognition improves significantly when using more flexible methods to mine the electrodermal signal (Henriques et al., 2013). Another study, conducted by Christopoulos, Uy & Yap (2016), has developed a table to interpret SCR through phases and relevant metrics. The authors describe two major categories for analysis: acquisition and extinction. Acquisition is a process through which a stimulus acquires value and extinction is a process through which a conditioned stimulus loses its value. Depending on the responses, a logical path shows how to interpret the signal. The signals are associated with emotions, decisions, and eventually behaviour. It is a method presented as a way of analysing EDA related to emotions. Another method has been described in a paper called "EDA Positive Change". Leiner, Fahr, and Früh (2012) carried out an experiment of continuous stimuli during media exposure. In the study, an EDA Positive Change parameter is proposed as a simple algorithm to calculate from raw skin conductance data and is similar to the frequency of SCRs. In short, a method may put together different approaches, models of emotion, data processing, and EDA mining.

The combination of different approaches and measurement instruments is suitable for performing cross analysis and fine tuning the experiment. In a study (Lei, Sala, & Jasra, 2017), participants watched videos associated with different emotions. EDA and heart rate (HR) data were collected, and the facial expressions were recorded. The outcomes demonstrated a higher connection between emotions and SCR as compared to emotion and HR. Another study (Kettunen, Ravaja, Näätänen, Keskivaara, & Keltikangas-Järvinen, 1998) conducted with 37 middle-aged men, examined the synchronization between indices of ANS, EDA, HR, and subjective and behavioural arousal. The results were: Phasic EDA and HR accelerations were synchronized, whereas there was no association between Tonic EDA and HR in-between subject analysis. Finally, a study conducted in Japan showed the positive relation between EDA and eye blink to indicate visual attention (Sakai et al., 2017).

Concerning emotional responses in academic situations, a triangulation of approaches may be beneficial to capture multiple facets from the experiment. These facets arise from participants' behavioural changes during the session, verbal and non-verbal features registered through video analysis, and physiological signs, which are collected without a subject's interference. This is particularly relevant as there are multicomponent systems involved in academic emotions. Pekrun defines emotions as "sets of interrelated psychological processes, whereby affective, cognitive, motivational, and physiological components are of primary importance" (Pekrun et al., 2011, p. 37).

Currently, there is a lack of an accurate understanding of how emotional experiences of group members, participating in a collaborative activity in an authentic learning setting, unfold along different learning phases of a project, in which students are encouraged to work in groups and reach a goal, managing both cognitive skills and emotional states. Specifically, no research has investigated temporal variations in a collaborative task and emotional progress both in terms of arousal and valence. Progress data and robust measurement are crucial so that theory can continue to grow (Renninger & Hidi, 2015). Therefore, this study combines video analysis and physiological responses to better capture and understand emotional responses in a collaborative learning setting.

4. Aim of the study and research questions

The aim of this case study is twofold. The first sub-aim is to explore how group members' arousal levels vary across different phases of a collaborative learning task. The second one investigates how case students' emotional responses are distributed in the arousal-valence space across the phases of the collaborative task. The research questions are:

RQ1: How group members' arousal levels vary across different phases of a collaborative learning task?

RQ2: How are case students' emotional responses distributed in the arousal-valence space across the phases of a collaborative learning task?

5. Method

In Chapter 3, different methodological approaches have been considered for this study: self-reports, process-oriented, trace data, and physiological measures. In the following chapters, the method chosen for this study is described in detail.

5.1 Participants and activity

The participants of this study were sixth graders (approximately twelve years old) of a school in Finland. Twelve students participated in the study, six boys and six girls. The students were assigned to groups of three heterogeneously, based on their interest towards science learning (they previously answered a questionnaire to determine their interest level). The task was to design and construct a model of an energy efficient house that makes use of solar energy. Students could use different materials (e.g. cardboard, tape, aluminium foil, cotton wool, Styrofoam sheets) to build the house. Students had to follow specific and written rules to build the house. The groups were working in a classroom-like learning and research laboratory. Before the building task, some basic information about heat energy was presented to the students by the teacher. After an introduction, in which students learned about solar energy in a jigsaw method (each participant read a different and complementary part of the information) (Aronson & Patnoe, 2011). The task consisted of three different phases: 1) brainstorming 2) planning 3) building. Every phase was timed, and remaining time was also visible all the time to generate time pressure.

In the brainstorming phase, students had to share the expertise gained earlier and brainstorm a list of things to keep in mind while sketching and building the house. After 10 min of brainstorming, students moved on to the planning phase. In the planning phase, students made a sketch and a floor plan of the house. The sketch should display clearly how the house would be built, and which materials would be used. Time for sketching was 20 min. After 5 min working in the planning phase, students received additional information about the dominant direction of the wind and angle of the sun during summer and wintertime. Later, in the building phase, students had 60 min to build the model house and finish the task.

Each phase allowed students to work in a different way during the collaborative task. The brainstorming phase emphasized the understanding and the definition of a shared goal. In the planning phase, the objectives were to create a work plan to build the house, in which negotiation and adaptation skills were required to carry out this phase successfully. In the building phase, cognitive and motor skills were required. It was the practical side of the project in which the plan requirements needed to be worked out to build an energy-efficient house model.

5.2 Data recording

The whole activity was recorded with 360° video cameras, which offered a full view on students, allowing for precision and detailed video analysis. Students' EDA was recorded with Empatica E4 wristbands (Garbarino, Lai, Bender, Picard, & Tognetti, 2014). The wristbands were placed on the non-dominant hand to minimize motion artefact. Room temperature was controlled at 23°C to ensure proper measurement conditions, as recommended by Boucsein (2012).

5.3 Ethical issues

The experiment has been conducted according to the best ethics practice of the University of Oulu, following the responsible conduct of research guidelines (RCR guidelines) and the process of the Finnish Advisory Board on Research Integrity (TENK) with extensive education. Information letters and consent forms to students and their guardians were accepted by the Ethics Committee of Human Sciences of University of Oulu.

The students participating in the experiment were selected on a voluntary basis. A written consent was asked from the students and guardians and both students and guardians signed it. Both forms (student's and guardian's) were required for participation. Also, an information letter (written in a "simple" language, suitable for the age group) about the study was given to both, students and guardians, prior to signing the consent forms. The consent forms included the consent to participate in the study and the permission to use pictures and videos (from which individual student could be recognized) in public presentations. The research team went to the school to inform students about the study. The forms were then given to the students by their own classroom teachers. After the consent forms had been collected, the information was inserted in Excel file and the original forms were stored behind locked doors.

Students were not exposed to any harm whatsoever and the respect for students' dignity was carefully followed by the research team. Names of students and school have been kept under privacy during research meetings, presentations, and the elaboration of the manuscript. Although data belongs to a broader project conducted by the LET Research Unit team, there is no affiliation involved and no potential conflict of interests.

5.4 Validity and reliability

To ensure validity and reliability, the following conditions were met: the experiment was conducted with four groups of 3 students each; the room temperature, the atmosphere, and the instructions given resembled of a regular classroom; the Empatica E-4 wristband was used to measure EDA. It is a validated tool in psychophysiological studies; 360° videos were recorded to ensure multimodal data analysis and compliance with adequate teaching practices; the total number of video files recorded and data collected during the session is available for analysis; all data files and tables generated are available under request, without disclosing participant's names and school names.

6. Methodological and analytical steps for electrodermal activity data processing (Aim 1)

As mentioned earlier in Chapter 3, the objective of this study is twofold. The first sub-aim is to explore how group members' arousal levels vary across different phases of a collaborative learning task. There are three crucial task or context characteristics that influence on the nature of the physiological data (namely EDA collected with Empatica E4 wristbands) collected in this study. First, the data was collected in a real learning situation, in which students performed an authentic learning task within a science domain, contrasted to studies held in laboratory settings. Second, the learning task required students to work in groups. Hence, the task was collaborative in nature and relied on interaction among group members. Consequently, the final characteristic was related to emotional responses from group members, situating the activity in a socio-emotional context. These three factors are reflected in the empirical aim of the study and have underlined the search of relevant data processing and analysis methods.

In addition to these task and context characteristics, there are different aspects of the EDA data to be identified and considered while physiological data is being analysed. These aspects are visual inspection, physiological compliance (PC), directional agreement (DA), and, finally, SCR. They are explained in the following sections and the grounds for selecting SCR to be used as trace data in the empirical case study.

6.1 Visual inspection

Visual inspection of the EDA signal is recommended for measurement fidelity (Shaffer, Combatalade, Peper, & Meehan, 2016) and to identify possible artefacts (unimportant drift factors) (Braithwaite et al., 2013). All data from Empatica wristband was downloaded and the files were imported into an application called Ledalab (ledalab.de), running on MathLab software. The application generated a chart, but there was no immediately visible trend, similarity, or mutual influence between the students. This is line with Boucsein's reference about charts, in which "evaluation of phasic changes focuses mainly on irregularly appearing single events, rather than on patterns that may be characterized by changes in frequency and/or amplitude" (Boucsein, 2012. p. 150).

The advances of visual inspection are related to the comparison of specific events, analysis of graph slopes, and the identification of artefacts, which can be noisy and daunting to the researcher (Blain, Mihailidis & Chau, 2008; Mönttinen, Koskimäki, Siirtola & Röning, 2017; Taylor et al., 2015). However, for this study, visual inspection was not selected as a method for further analysis due to the amount of data traces (over 28.000 logs per student) and the fact that preliminary analysis showed no identifiable patterns among students during the collaborative learning activity. Next, the possibilities of physiological synchrony among students were explored as a method for analysis based on two approaches: PC and DA.

6.2 Physiological compliance and directional agreement

The second characteristic is PC, "the correlation between physiological measures of team members over time" (Elkins et al., 2009, p. 1). If you measure EDA signal from two or more students and they have a close correspondence or mutual influence, you have cues of PC.

As you analyse PC in a group of students, for example, DA is the simplest of PC indicators and is the most sensitive for differences in team performance (Pijeira-Díaz, Drachsler, Järvelä & Kirschner, 2016). It shows if students are going towards the same direction between two data points. For instance, if the value at data point 1 were lower than the value at data point 2, the DA is considered "increasing". If the value at data point 1 were higher, then data point 2 would be considered "decreasing" (Elkins et al., 2009).

The EDA data of this study was explored from the DA point of view in the following way: EDA data points from three students (ID 02, ID 08, ID 04) were compared, at the same point in time, to search for DA (Table 2). If there was an increase in the EDA data point compared to the previous data point in a student's log, it was considered as "1" (meaning the current EDA data point is higher than the data point before). If there was a decrease, it is considered as "-1" (meaning the current EDA data point is lower than the data point before), and if there was no change, it was classified as "0". If the students had the same value at the same time, the result is "TRUE", meaning that there was a DA at that point in time. In other words, students experienced higher arousal, maintained it, or decreased it. EDA samples were taken every 250ms. The longest synchronous period detected among all students in a group was one occurrence of 1,5sec (6 logs in a row), which does not appear to be a sound and sufficient evidence to make an affirmation about an existing synchrony between the students in a group.

PC and Social PC have been studied in the context of team performance and it is a well-established concept to predict performance in groups with higher levels of synchrony

(Elkins et al., 2009; Henning, Boucsein & Gil, 2001; Pijeira-Díaz et al., 2016). However, for this study, no PC has been identified longer than 1,5s, which might be an indicator of low PC or might not be applicable to the goals of this particular study. Therefore, a fourth characteristic has been considered: SCR.

Table 2

Sample of a directional agreement calculation

ID 02	ID 08	ID 04	dir_02	dir_08	dir_04	Result
0.55	0.13	0.27	0	1	1	FALSE
0.55	0.13	0.28	-1	-1	-1	TRUE
0.54	0.13	0.27	-1	-1	1	FALSE
0.54	0.13	0.27	0	1	-1	FALSE
0.54	0.13	0.27	1	1	1	TRUE
0.54	0.13	0.27	-1	0	1	FALSE
0.54	0.13	0.28	1	0	-1	FALSE
0.54	0.13	0.27	0	0	0	TRUE
0.54	0.13	0.27	0	0	0	TRUE
0.54	0.13	0.27	1	1	0	FALSE
0.54	0.14	0.27	0	-1	1	FALSE
0.54	0.13	0.27	-1	-1	0	FALSE
0.54	0.12	0.27	0	-1	-1	FALSE
0.54	0.08	0.27	1	1	1	TRUE
0.54	0.10	0.27	0	1	0	FALSE
0.54	0.12	0.27	0	1	-1	FALSE

6.3 Skin conductance responses

The SCR represent the fast-changing elements in the skin conductance signal and can be identified as “peaks” in the signal. It is a result of the sympathetic neuronal activity. The SCR was the methodological choice for this study, since the number of ppm is an indicator of SCRs, a unit to measure arousal episodes (Boucsein, 2012; Braithwaite et al., 2013; Pijeira et al., 2018).

7. Analysis procedure for the empirical case analysis on group members' emotional arousal and valence during the collaborative learning process (Aim 2)

In the twofold aim, the second one was to implement the physiological data in exploring if the physiological arousal could indicate emotional arousal. This was done by relating the EDA data that indicates emotional arousal with the valence that is visible through case group's collaborative working and interaction. In practical terms, students' physiological data was combined with their valence, observed from video recorded collaborative working session. Emotional episodes were first depicted from the video data and then were coded according to their valence: positive, neutral, negative (See Table 3 in the Video data section – 7.2). Some studies consider valence as a function of arousal, whereas other studies do not find relationships between them (Kuppens, Tuerlinckx, Russell & Barrett, 2013), At the end of data analysis, it was decided to use an orthogonal model as proposed by Russell (1980), the circumplex model of affect. Emotional valence was plotted in the arousal-valence space, creating a table based on the frequency of peaks combined with emotional valence episodes.

7.1 Electrodermal activity data

Data from wristband devices were downloaded using the software provided by the manufacturer (Empatica). The results were plotted onto an Excel sheet into two columns: the first column is the exact log time, recorded every 250ms, and the second one is the electrodermal signal value. Data obtained from EDA wristband is raw and has been transformed by the Continuous Decomposition Analysis (Benedek & Kaernbach, 2010). According to Boucsein (2012, p. 150), "the question of whether statistical analysis should be carried out with transformed data or with raw data still cannot be answered in general". The purpose of data transformation is to improve the validity of the EDA measurement. The Continuous Decomposition Analysis (CDA) method accounts for SCRs superposition in EDA signal and implies a more robust analysis, which is relevant concerning artefacts (Benedek & Kaernbach, 2010).

This dataset was imported into Ledalab software, based on Matlab (MathWorks, Inc.). Once imported, datasets were transformed by the CDA software. The outcome is a transformed dataset containing the onset time and the signal amplitude, which was used for calculations of arousal peaks during the collaborative learning session.

7.2 Video data

The 360° video was imported into Observer XT (Noldus - a software for observational data collection, analysis, and presentation). The video has been analysed by a person and it was manually codified to categorize valence. Based on the mean duration (24,6 seconds) of episode coding in preliminary analysis, the video data was first sequenced in 30 second-segments. Those segments were then coded into two categories in terms of whether they included emotional responses or not. These responses included verbal (e.g. “We are so good”) or other signs (e.g. laughing, sighing, lack of focus) of positive or negative emotions and negatively or positively charged interactions (e.g. joking, encouraging, criticizing, arguing).

Students’ individual emotional states were coded into four categories (positive, negative, neutral, unclear) (Table 3). Only one code per student was given for each 30-second segment. Valence was coded as positive, when a student expressed clear signs of positive emotions or made a positively charged comment, and negative in opposite cases. If there was no emotional expression, valence was coded as neutral and if valence was not clear (e.g. multiple indicators in one segment), valence was coded as unclear.

Table 3

Types and examples of valence coding

Valence	Indicators	Example emotions (Kreibig, 2010; Linnenbrink-Garcia, Rogat, & Koskey, 2011; Pekrun et al., 2002a; Russell & Barrett, 1999)
Positive	Verbal signs (e.g. “We are so good”) Bodily signs (e.g. laughing, giggling) Positively charged interaction (e.g. joking, praising, encouraging)	Excitement Happiness Enjoyment Hope Pride Relief

Negative	Verbal signs (e.g. "We are idiots") Bodily signs (e.g. sighing) Lack of focus (e.g. playing with equipment, wondering around) Negatively charged interaction (e.g. arguing, criticizing others) Tensioned silence	Anger Anxiety Frustration Annoyance Shame Disgust Fear Hopelessness Boredom
Neutral	No visible emotional response	-
Unclear	Both, positive and negative, emotional indicators (e.g. positive bodily sign + negative verbal sign)	

The following examples describe one 30-second segment.

Example 1. Mainly negative valence. Student 7 is joking during the brainstorming phase and the other students are getting frustrated.

Student	Utterance (indicator)	Valence
Student 7	*"We have only two mins left to finish this phase (laughing)."*	Positive
Student 3	*"Ok, so let's hurry up! (emphatically, frustrated)"*	Negative
Student 7	*"The door. Let's write about the door. Hey, let's*	Positive

	make jokes about the door! (smiling, joking)"	
Student 1	*"(sighs) We are doing research and not making jokes. (frustrated)"*	Negative
Student 3	*"Ok, now really, what important information you have in your notes? (emphatically, frustrated)"*	Negative

Final codes: Student 1: *Negative*; Student 3: *Negative*; Student 7: *Positive*

Example 2. Positive valence. Students are joking together about a toilet while planning the house.

Student	Utterance (indicator)	Valence
Student 7	*"Remember to make a toilet there. (laughing)"*	Positive
Student 1	*"Yeah, it is quite useful."*	
Student 3	*"Well, I don't know. Toilets are a bit overrated. (joking)"*	Positive
Student 1	*"Yeah, we don't need those. They can just go under the tree. (joking)"*	Positive
Student 5	*"Let's make a golden toilet seat. (laughing, joking)"*	Positive
Student 3	*"Yeah, let's just but the tree in the front yard. (joking)"*	Positive
Student 1	*(laughing)*	Positive
Student 7	*"Let's make an outhouse. (laughing, joking)"*	
Student 3	*"An outhouse. Let's make it here... (laughing, joking)"*	Positive
Student 1	*(laughing)*	Positive

Final codes: Student 1: *Positive*; Student 3: *Positive*; Student 7: *Positive*

Example 3. Unclear and neutral valence. Student 7 is playing with a cotton wad and trying to hide aluminium foil in it. Student 1 does not seem very excited about it. Later, when Student 3 drops more foil on the table, Student 7 and Student 1 are very eager to grab it and start to play with it. The valence of emotional responses of Students 7 and 1 are not clear since playing with equipment (boredom) and whining would be considered as negative indicators but enthusiasm and joking as positive indicators. Student 3 is expressing no emotions during this 30-second segment, so his valence is coded as neutral.

Student	Utterance (indicator)	Valence
Student 7	*"(playing with a cotton wad, joking) Can you see it now?"*	Unclear
Student 1	*"It's inside the wad."*	
Student 7	*2Can you see it?"*	
Student 1	*"It's inside it!"*	
Student 7	*"No, it's not...! (whining)"*	Negative
Student 1	*"Yeah, it is... (whining)"*	Negative
Student 3	*"Here is some foil for you... (drops some foil on the table, no emotional indicator)"*	Neutral
Student 1	*(grabs the foil with enthusiasm, starts to play with the foil)*	Unclear
Student 7	*(grabs the foil with enthusiasm, starts to play with the foil)*	Unclear

Final codes: Student 1. *Unclear*; Student 3: *Neutral*; Student 7: *Unclear*

Törmänen, Järvenoja, Kurki, Järvelä & Devai (2018)

To classify the students playing a leading role in their groups, the following code has been considered and applied during the video recording analysis:

- Giving directions to peers.
- Demanding work to be done.
- Giving feedback to peers.
- Providing encouragement or putting pressure on peers to finish the task.

The data from EDA and video was synchronised to allow further calculation and analysis. Based on UTC (Coordinated Universal Time), EDA processed logs were plotted onto a column and the coded video was placed next to it, matching exact times. Thus, a comparison between the exact signal and the corresponding video event time could be performed. In order to compute the number of peaks per minute, a moving window approach was employed to account for the activity continuous flow rather than abrupt changes at specific times. This moving average calculated continuous periods of 60 seconds, starting at every 0,25 sec. For EDA data and video data synchronization, there is an error margin < 0.25sec (250ms), corresponding to the sampling frequency of the EDA signal, 4hz, that is to say one sample every 250 ms.

7.3 Thresholds for results

To be more specific on tracking arousal changes, thresholds for arousal levels have been proposed so that individual differences could be considered (Figure 2). Since arousal levels are different for every individual, it was necessary to calculate means so that arousal changes during the collaborative learning task could be mapped and analysed. Also, to date there is no general measurement reference for EDA in the context of collaborative learning.

This reference was created by adding up one standard deviation up and one down from the arousal mean of every student. The mean used in the experiment was calculated based on the whole session. Based on the reference, three arousal levels have been established: high, middle, and low.

Figure 2. Thresholds for arousal levels

8. Case study results

The results have been considered under two different perspectives. For the first research question, arousal levels from a group of twelve students were measured and analysed. For the second research question, the focus has been placed on two students, who demonstrated different behaviours during the collaborative learning task. One of them lead the activities and the other one loafed around most of the time. Both students have been compared in terms of arousal and valence. The results are described in the following sections.

8.1 Research question 1

"How group members' arousal levels vary across different phases of a collaborative learning task?"

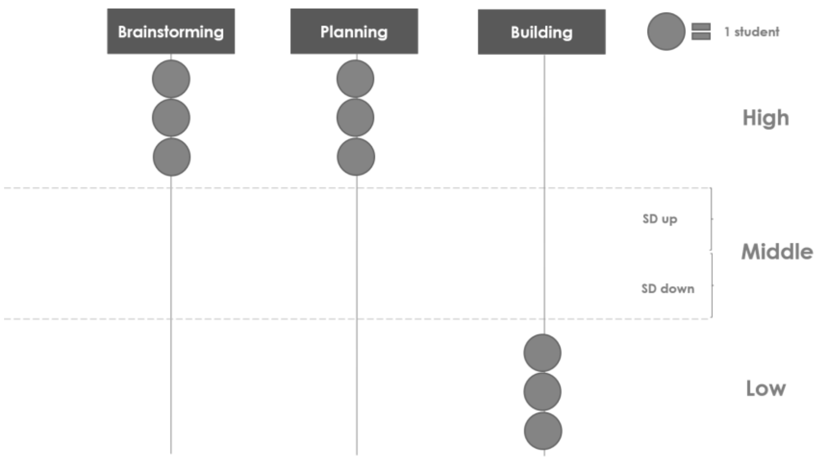
This question targets variation in students' arousal levels across three collaborative learning phases, namely brainstorming, planning, and building. Four groups were selected for the case analysis. In general, the results show that all case group members' arousal level varied, during the whole experiment, within the range of 0 and 121 peaks per min. A threshold based on the average of minimum peaks and the average of maximum peaks was set: respectively, 26 and 88 ppm. Considering this averaged range, the concentration of peaks within it is presented as follows: 9 out of 12 students (75%) had 95.3% or a higher number of peaks within this range. 2 out of 12 students (16%) had between 80% and 90% of their peaks within the range, and only one student had 43,2% of his peaks within the same range (26 and 88 ppm).

Next, an arousal profile was created for each case group member based on the fluctuation in arousal between the three collaborative learning phases. As a result, three different types of arousal profiles emerged. The first profile showed a decrease in the arousal level towards the end of the experiment. The second profile was the opposite, showing increasing arousal as the collaborative task progressed. Finally, the third profile revealed two students who displayed steady arousal levels throughout the three phases.

In the first profile, the results showed that 3 out of 12 students fell in to the decreasing arousal profile. Arousal development of the students in this profile displayed higher arousal levels throughout the two initial phases: brainstorming and planning (Figure 3). Based on the threshold for arousal levels, previously introduced in this case study, the three students performed their activities during the first two phases ranging between middle and high arousal levels. In the Brainstorming phase, the first student spent 89.3% of the time in middle and high arousal levels; the second student, 99.8%; and the third one 98.2%. In the Planning phase, the first student spent 97.5% of the phase time varying between middle and high arousal levels; the second student, 100%; and the third one, 87.5%. During the third phase (building), their arousal went to low levels. The first student spent 96.5% of the phase time ranging between middle and low arousal levels; the second student, 96.2%; and the third one, 97.4%.

These three students were all from different groups. When contrasted to the video recording, the observation results showed that two of the students had a leading role in their group, telling the team members what they should do. For these two leaders, the work in their group was conflictive because team members spent most of the time playing around and carrying out off-topic conversation. The third student was not leading any activity, and she was in a group where activities flowed smoothly, and everyone seemed interested in accomplishing the task. The atmosphere was calm, and the group members shared the goal of building a house and worked systematically towards it.

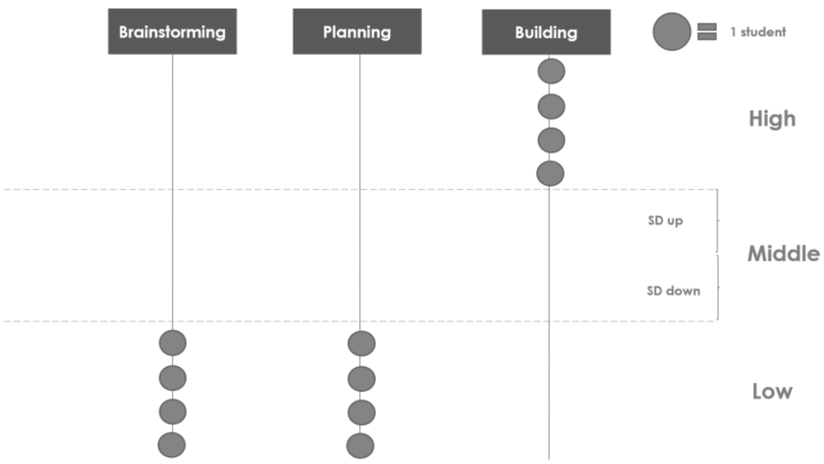
Figure 3. Changes from higher arousal levels in brainstorming and planning phases to lower arousal level in building



The second profile was defined by an increase in arousal towards the end. Altogether, four out of twelve students experienced lower arousal levels during the brainstorming and planning phases. At the building phase, their arousal went up to higher levels (Figure 4). The four students performed their activities during the first two phases ranging between middle and low arousal levels. In the Brainstorming phase, the first student spent 94.9% of the phase time in middle and low arousal levels; the second student, 89.2%; the third one, 100%, and the fourth one, 87.4%. In the Planning phase, the first student spent 91.8% of the phase time varying between middle and low arousal levels; the second student, 91%; the third one, 99.6%, and the fourth one, 88.6%. During the third phase (building), their arousal went to higher levels (Fig. 5). The first student spent 86.4% of the phase time ranging between middle and high arousal levels; the second student, 96.2%; the third one, 92.5%, and the fourth one, 88.6%.

When contrasted to the video observation results, none of these four students had leading roles. Three of them were not concerned about accomplishing the task, while the fourth student belonged to the group which had a shared goal and the atmosphere was calm and supportive.

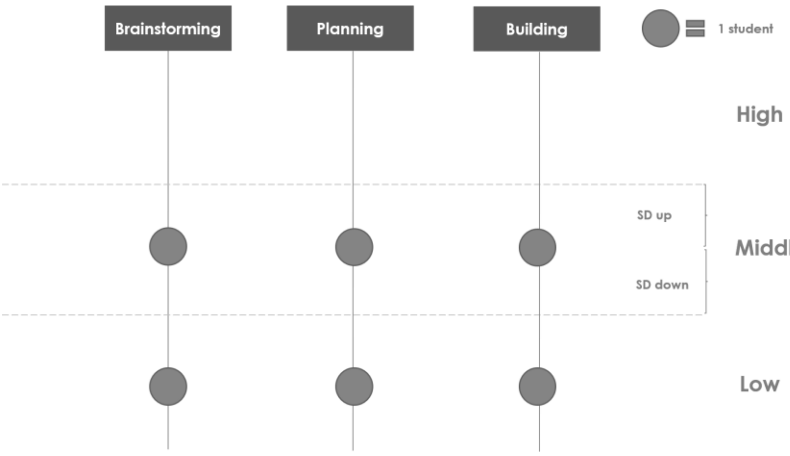
Figure 4. Changes from lower arousal levels in brainstorming and planning to higher arousal levels in building



Finally, the third profile represented a steady arousal level development. The results showed that two out of twelve students maintained their arousal levels steady across the three phases (Figure 5). The first student ranged between middle and low arousal levels. In every phase of the experiment, this student remained 100% of the time in middle and lower levels, never experiencing high arousal levels. In the Brainstorming phase, it was 55.8% of the time in middle levels and 44.2% in low levels. In the Planning phase, it was 50.4% of the time in middle levels and 49.6% in low levels. In the Building phase, it was 86% of the time in middle levels and 14% in low levels. The second student displayed arousal levels in the middle range most of the time across the phases. In the Brainstorming phase, it was 91% of the time in middle level; in the Planning phase, it was 98.2% of the time in middle level; and in the Building phase, it was 92.2% of the time in middle level.

These two students were in the same group. The first student played the leading role and the second one was most of the time joking and having off-topic conversations. The one who led the group gave several commands to the other students, but they did not seem to care. The leader seemed frustrated at the building phase.

Figure 5. Students who displayed similar arousal levels across the three phases.



The remaining three students did not display a clear pattern in terms of arousal level changes, it was rather a variation between the arousal levels. The first student spent 55.8% of the time in high arousal levels during the Brainstorming phase, 98.9% in middle and low levels in the Planning phase, and 93.8% in middle levels during the Building phase. The second student spent most of the time in the Brainstorming and Planning phases at middle arousal levels (99.9% in Brainstorming; 95.7% in Planning) and arousal went up at the Building phase, 94.3% of the time ranging between middle and high arousal levels. Finally, the third student spent 50.1% of the time at middle levels in the Brainstorming phase and split the other half between high and low levels. In the Planning phase, it was 95.7% of the time varying between middle and high; in the Building phase, it was 90.7 of the time ranging between middle and low.

8.2 Research question 2

"How are case students' emotional responses distributed in the arousal-valence space across the phases of a collaborative learning task?".

The purpose of the second research question was to explore the possible connection between two types of process data, namely physiological data (EDA) and behavioural data (video). Particularly, the aim was to illustrate through visualization how arousal, determined from EDA data, and valence, determined from video data, are combined with each other during each phase of the collaborative learning task. Episodes identified as 'unclear valence', which may be interpreted ambiguously, such as surprise, were not considered for the analysis. Student 1 (further represented by the pseudonym "Anna") and student 3 (further represented by the pseudonym "Tapio"), from the same group, were selected for case analysis, because they displayed very different behaviours during the task. Tapio was the one leading the group and experiencing pressure to finish the task. Anna spent most of the time playing around and having off-topic conversation with a third student in the group. The purpose of picturing valence and arousal levels was to characterize emotional responses in a broader way, since arousal indicates the activation level of an emotional response, but not the quality of it.

8.2.1 Brainstorming phase

The aim of the brainstorming phase was to activate the group members to share with each other their individual prior knowledge about the prerequisites to build the model house. In the brainstorming phase (10 mins), the situation in the case group was the following: Tapio was leading the group work. He tried to ask questions from Anna about the topic so that she could write something down. She did not really help. In the end of the brainstorming phase, Anna started to count down the time left. Tapio asked her to stop. She did. Then, Anna engaged in off-topic discussion with a third student in the group. As a result, Tapio had to finish the task alone.

In terms of arousal and valence, Anna predominantly experienced middle arousal level (58% of the time) combined with positive valence (49,5% of the time), while Tapio displayed middle arousal level only (100% of the time), constantly switching between positive (43% of the time), negative (36%), and neutral valence (21% of the time). In chronological terms, Anna was characterized by low arousal and positive valence in the beginning of the brainstorming phase. In the end, she experienced middle arousal and alternating valence, between positive and negative. Tapio displayed a middle arousal level along the brainstorming, however his changes in mood were frequent in terms of valence (Table 4).

Table 4

Frequencies of arousal-valence episodes in the brainstorming phase

Valence \ Arousal		*High*		*Middle*		*Low*		*SUM*	*SUM*
		Anna	Tapio	Anna	Tapio	Anna	Tapio	Anna	Tapio
Positive	Anna	1 (8,5%)		3 (25%)		2 (16%)		6 (49,5%)	
	Tapio				6 (43%)				6 (43%)
Neutral	Anna			1 (8,5%)		1 (8,5%)		2 (17%)	
	Tapio				3 (21%)				3 (21%)
Negative	Anna			3 (25%)		1 (8,5%)		4 (33,5%)	
	Tapio				5 (36%)				5 (36%)
SUM	Anna	1 (8,5%)		7 (58%)		4 (33,5%)		**12** (100%)	
SUM	Tapio		0 (0%)		14 (100%)		0 (0%)		**14** (100%)

8.2.2 Planning phase

In the planning phase (20 min), the third student in the group suggested that Anna did all the sketching because she was good at it. However, Tapio prevented Anna from doing it. And then Tapio, the one leading the group, started to draw the sketch himself. Anna started to joke about whether to build a toilet inside or outside the house. Later she said she did it because she was bored. Tapio kept working alone on the sketch and Anna was mostly joking or having off-topic discussions. Anna participated for a few moments and answered some of the questions raised by Tapio. In the end, Tapio did all the sketch by himself.

Concerning arousal and valence, Anna predominantly experienced middle arousal level (68% of the time) combined with positive valence (77,5% of the time), whereas Tapio displayed predominantly middle arousal level (60% of the time), constantly switching between positive (50% of the time), negative (33%), and neutral valence (17% of the time). In chronological terms, Anna was predominantly characterized by middle arousal and positive valence along the planning phase, while Tapio had higher levels of arousal at the first half of the phase and middle level for the second half, while valence was continuously switching back and forth from positive to negative (Table 5).

Table 5

Frequencies of arousal-valence episodes, in the planning phase

Valence \ Arousal		*High*		*Middle*		*Low*		*SUM*	*SUM*
		Anna	Tapio	Anna	Tapio	Anna	Tapio	Anna	Tapio
Positive	Anna	4 (13%)		15 (54%)		3 (10,5%)		22 (77,5%)	
	Tapio	4 (13%)			11 (37%)				15 (50%)
Neutral	Anna	1 (4%)		1 (4%)				2 (8%)	
	Tapio	2 (7%)			3 (10%)				5 (17%)
Negative	Anna			3 (10,5%)		1 (4%)		4 (14,5%)	
	Tapio	6 (20%)			4 (13%)				10 (33%)
SUM	Anna	5 (18%)		19 (68%)		4 (14%)		**28** (100%)	
SUM	Tapio	12 (40%)			18 (60%)		0 (0%)		**30** (100%)

8.2.3 Building phase

In the building phase (60 min), Tapio started building the model house as soon as the phase began. He gave instructions to the third student in the group to sort handouts and sheets of paper to the same pile. Tapio started to cut cardboard. The third student asked if she could cut, but Anna and Tapio said “no” to her. Tapio also said he did not trust the third student. The girls were just watching while Tapio was cutting the cardboard. Tapio said that the girls could participate by decorating the house, but he did not agree when Anna suggested pink walls to the house. The girls kept fooling around, joking, and observing while Tapio built the house. Tapio did not ask for help until the end of the experiment. If the girls tried to participate, Tapio asked them not to touch anything. At some point, Anna started to help Tapio with taping the walls together. Still, she did not do much to help. Tapio said that Anna had done something to help, but the third student had only been disturbing the work. Tapio finished the house in time, but he had no time to do extra decoration but put the cardboard walls together.

Concerning arousal and valence, Tapio predominantly experienced middle arousal level (75,5% of the time) combined with positive valence (79,5% of the time), while Anna displayed predominantly middle arousal level (90,5% of the time), and negative valence (73,5% of the time. In chronological terms, Tapio had middle arousal levels and positive valence in the first half, and a bit more of high arousal episodes in the second half, but valence was primarily positive along the second half. Anna experienced middle arousal and negative valence in the first half, and towards the end of the activity, the arousal levels remained the same, but negative valence increased (Table 6).

Table 6

Frequencies of arousal-valence episodes in the building phase

Valence / Arousal		*High*		*Middle*		*Low*		*SUM*	*SUM*
		Tapio	Anna	Tapio	Anna	Tapio	Anna	Tapio	Anna
Positive	Tapio	8 (16,5%)		28 (57%)		3 (6%)		39 (79,5%)	
	Anna				11 (17%)				11 (17%)
Neutral	Tapio			2 (4%)				2 (4%)	
	Anna				6 (9,5%)				6 (9,5%)
Negative	Tapio	1 (2%)		7 (14,5%)				8 (16,5%)	
	Anna		2 (3%)		41 (64%)		4 (6,5%)		47 (73,5%)
SUM	Tapio	9 (18,5%)		37 (75,5%)		3 (6%)		**49** (100%)	
SUM	Anna		2 (3%)		58 (90,5%)		4 (6,5%)		**64** (100%)

In sum, the results for RQ1 showed that 84% of the students experienced changes in arousal levels during the three phases of the experiment. These changes were either from low to high or high to low in most students (58%) and other students experienced changes taking place around middle levels and moving to higher or lower levels (26%). Only 16% of the students remained at the same arousal level across the three phases. Concerning RQ2, as the contrast between arousal levels and emotional valence were observed in two case students, the results were: in the brainstorming phase, Anna demonstrated low to middle arousal level, characterized by neutral to positive valence; while Tapio displayed middle level arousal and frequent changes in valence, switching between positive and negative. In the planning phase, Anna experienced a middle level arousal and positive valence, whereas Tapio displayed high and middle arousal levels, again with frequent changes in valence, from positive to negative. In the building phase, Tapio displayed middle and high arousal levels and positive valence while Anna demonstrated middle arousal level and increasing negative valence towards the end of the experiment. Thus, Anna experienced all arousal levels and mainly positive emotional valence in the first two phases, while Tapio ranged between middle and high arousal levels, and experienced a mix ranging between positive and negative emotional valence, moving to more positive towards the end of the collaborative activity.

9. Discussion

The aim of this study was twofold: the first sub-aim explored how group members' arousal levels vary across different phases of a collaborative learning task. The second sub-aim investigated how case students' emotional responses are distributed in the arousal-valence space across the phases of the collaborative task. For triangulation purposes, the empirical results of the physiological arousal were further contrasted with an emotional valence analysis, from another process-oriented data source, namely video-observation of two case students. The arousal variation and emotional valence observation of students took place across three collaborative learning phases: brainstorming, planning, and building. Since there were no clear analytical approaches that would reliably suit this type of data analysis, an extensive exploration and trial of different methodological approaches, particularly focused on PC and arousal measurement, was first conducted.

Classic studies in emotion research have considered arousal and valence to be major dimensions of emotional responses (Russell, 1980; Russell & Barrett, 1999). Concerning methodological aspects, the first procedure was to compare students' arousal during the activities to see if there was a synchronous physiological response. Since the longest synchronous period detected was 1,5 s, the DA was not considered relevant for this study. Also, it should be noted that this study focused particularly on investigating individual group members' as a unit of the analysis, not the group or person-in-the-group, even though the learning setting was collaborative in nature. However, in a former study about physiological coupling indices (PCI) (Pijeira-Díaz et al, 2016), DA has been considered the most relevant PCI to compute differences at a group level. One possibility for not finding synchrony among students might be related to the fact that students may be "experiencing similar patterns of physiological activity", but these responses might not be related to the interactions among them (Palumbo et al., 2017). Other studies, however, reveal that physiological synchrony has taken place in educational settings, during collaborative learning, at statistically significant level. One study (Haataja, Malmberg & Järvelä, 2018) found out that when students were monitoring themselves while performing the collaborative activity, their values of physiological synchrony were higher. In a study conducted about PC and team performance, a positive relationship has been found. "The correlation measure showed the strongest predictive relationship with performance" (Elkins et al., 2009, p. 1002). In a comparative study of physiological measures and PC, social-physiological compliance has an important "role in determining proficiency in social performance" (Henning et al., 2001, p. 230), however "the strong associations between social physiological compliance and improved team performance that were identified … cannot be readily attributed to matched task behaviors resulting in matched physiological responses" (Henning et al., 2001, p. 230). There was another study, conducted in a science class with 6th graders students (Gillies et al., 2016), in which cooperation was encouraged in group activity and EDA was measured. The results revealed that "the overall level of synchrony between students across the whole class was reduced during the small group cooperative learning activities" (Gillies et al., 2016). The authors said that students were working in groups, but in a more independent way, thus reducing synchrony levels among them. Additional studies have shown evidence of physiological synchrony (Knoblich, Butterfill & Sebanz, 2011; Marsh, Richardson & Schmidt, 2009; Mønster, Håkonsson, Eskildsen & Wallot, 2016).

With reference to arousal measurement, the number of peaks from all of the twelve students in the experiment were concentrated in a range between 26 and 88 ppm. This range allowed us to establish a starting point for the empirical investigation. Although arousal was later measured individually, it was important to have an overall range to consider as a framework. In addition to peaks per min, the experiment was measured by phases, considering the duration of each phase as a framework for students' collaborative activity. This measurement was in line with Revelle & Loftus (1992), who emphasized that the timing and the temporal resolution of the measures should be determined in studies with arousal measurement. To the present, no standard range has been identified to measure arousal in collaborative learning activities. The current work contributes with a preliminary range, allowing further studies to be developed and compared to these ranges.

Regarding the variation of arousal levels during the three phases of the collaborative task, most of the students experienced a change in arousal levels, either from high to low or low to high (ten out of twelve students have changed their arousal levels along the phases). These changes reinforce the concept of arousal levels experienced individually, according to personality styles and task requirements (Revelle & Loftus, 1992). In a study focused on EDA and affect, the Myers-Briggs Temperament Indicator© (a temperament assessment) was applied and the results showed that psychological traits can be used as a reference for studies focused on measuring emotional interactions (Henriques et al., 2013). Results of the current study showed that while most of the students experienced some variation in physiological arousal levels, these variations differed between the students and within the groups. Considering the twelve students in the experiment, three of them displayed higher arousal levels at brainstorming and planning, and then fell to lower levels of arousal; two of these students were leading conflictive groups and one student was a member of a calm and collaborative group. The next set of four students were just members of groups (did not play a leading role) and arousal remained at lower levels during the phases of brainstorming and planning and was higher at the building phase. Although three of these members were in conflictive groups, they did not display higher levels of arousals during the first two phases. The task itself and the pressure from the group leader apparently did not generate higher arousal levels. In the next set of two students, who were in the same conflictive group, their arousal levels remained at middle arousal level throughout the whole experiment. The remaining three students in the experiment did not display any pattern in their arousal levels. Changes in arousal levels may also function as a regulating process, in which low arousal levels may require higher levels for pleasurable states or high arousal levels may demand for lower arousal levels, which will produce a more rewarding state (Zuckerman, 2014). While this study indicates changes in arousal levels and takes into account arousal variation, emotional atmosphere, and degrees of commitment, more research is needed, however, to investigate if and how the variation in arousal level is related to individuals' personal learning commitment on the one hand, and to the collaborative group work and group members' roles and interaction on the other hand. Further studies can focus more specifically on commitment levels, emotional atmosphere, and roles and interactions in a collaborative learning activity. The arousal ranges established as well as activity type can be both used as a reference for further studies in the area.

Concerning the second research question, this study contributes to place students' emotions experienced during a collaborative learning activity into Russell's circumplex model of affect, which combines arousal (from deactivated to activated, or low to high) and valence (from unpleasant to pleasant, or negative to positive). The relationship between valence and arousal has been comparatively described in a study (Kuppens et al., 2013), in which six types of relations have been identified in the literature, based on a moment to moment basis. Although there was no conclusive evidence of an asymmetric V-relationship between arousal and valence, most of studies but one point to this relationship. At group level, there were some interesting results. Overall, "as people feel more positive or negative, they tend to experience higher levels of arousal" (Kuppens et al., 2013, p.16). In the arousal-valence combination, individual differences have also been related to broad personality aspects (Kuppens et al., 2013).

In the present study, Anna displayed an increasing arousal level from the brainstorming to the building phase, with valence moving from positive to negative. Since Tapio was leading the activities and pushing the two other students to help, he displayed from middle to high arousal level and a shift from negative to positive valence. When we look at Tapio, in the building phase, there was a shift in valence. Tapio seemed more positive and Anna displayed negative valence. This is in line with the fact that tension and disagreement may lead to negative emotional arousal (Järvenoja et al., 2013).

In a videotaped study, conducted with sixth graders, affect was assessed in a 5-week math class and qualitative results suggested that negative affect (tiredness or tension) has been associated with higher levels of social loafing. Another interesting result showed that in lower arousal levels (neutral to deactivated) in conjunction with positive affect seemed to support positive group interactions (Linnenbrink et al., 2011). Looking at the arousal level, it may increase "when students perceive incongruity of goals or obstacles on their way to the goal during the learning process" (Phye et al., 2011). The perception of students may generate an emotional response or not. When students realize there is some obstacle and they can manage it without difficulties, they may not be aware of emotional changes and move ahead with the activity. However, students who consider the level of challenge negatively may interpret the obstacles as harm, loss, or potential threat. Thus, students have a primary appraisal process when faced with obstacles and a second assessment, depending on their capacity to deal with the situation (Boekaerts, 2007). Clearly, personality traits and social challenges influence arousal and valence combination, which may change on a moment to moment.

In this study, triangulation has been a valuable approach to study learning processes since EDA allowed for precise arousal measures, which has been contrasted to video observation to characterise valence. In a study called RECOLA (Ringeval, Sonderegger, Sauer & Lalanne, 2013), conducted with 46 users in France, different approaches have been employed: EDA and ECG measurement, video, audio, self-reports, and annotators. Students worked in a teamwork task, in dyads, and later these different sources were contrasted to create a corpus of remote collaborative and emotional interactions. Emotions have been measured on the dimensions of arousal and valence. As for the results, there was a "good inter-annotator agreement rate for the affective dimensions, and a fairly good one for the social dimensions" (Ringeval et al., 2013, p.7), but the relevance for emotion recognition is questionable. Another study, contrasting self-reports and physiological data has considered that the optimal way to deal with self-reports is to complement them with other measures

(Semmer, Grebner & Elfering, 2003). A multimodal study has revealed that the combination of video, audio, and physiological signal is key to better predict emotional arousal and valence from spontaneous recordings (Ringeval et al., 2015).

10. Conclusion

In a broad field of study comprising arousal, EDA, and valence, this case study makes a twofold contribution: on methodology and on empirical aspects. As Boucsein (2012, p.350) mentioned, “the focus of psychophysiological research moved away from a study type based on the assumption of a generalized unitary arousal concept to investigations of such conditions which were supposed to elicit certain arousal states and alterations.” Working in groups and being exposed to social interactions are external conditions affecting a student. As wearable devices become popular and interconnected with different measurement systems, they are of great help to analyse students’ emotional responses to advance studies in the field of EDA and arousal in authentic learning situations.

Therefore, in group interactions, physiological data can help teachers and researchers observe students’ emotional states and responses. There is a “clear interplay between affect and quality of group interactions” (Linnenbrink-Garcia et al., 2011). Having said that, when working with small groups, it is crucial to maintain positive group interactions, since negative affect can start a cycle of discouraging the performance of participants in a collaborative activity and reduce their engagement (Linnenbrink-Garcia et al., 2011). These accumulated negative experiences may contribute to social loafing or tuning out as described by Do and Schallert (2004, p. 632): “the most frequent antecedent to tuning out is an experience of a strong negative emotion or an accumulation of smaller aggravations.” It is interesting to observe that the student can simply expect something negative to happen and that prediction is enough to detach from the group, not because they are uninterested in the topic, but because they want to protect themselves from negative emotional experiences. Thus, tuning out also can mean a self-regulation attitude to protect and to avoid negative experiences, either during the event or just by predicting such an event (Do & Schallert, 2004). These emotional responses detected by arousal levels and classified in terms of valence can further be helpful to teachers´ intervention.

The current work helps to combine arousal and valence, in a single space, based on arousal data and video observation to make the process more precise and help see the evolution of arousal levels to other aspects affecting learning. In this study, the physiological arousal was related with valence through the connection of every emotional episode mapped along each phase of the experiment. Russell and Barrett (1999) say that a person can feel any combination of valence and arousal. From the experiment, we can see that students experience a variety of emotions related to their personality, to the task and to the social interactions in a group. It is recommended that the full range of students' emotional experiences should be considered and studied (Cahour, 2013). In conclusion, by understanding and being aware of arousal and valence changes throughout class time, teachers can improve the quality of group interactions and teaching material, helping students "manage with affective experiences, so that they can be resources rather than constraints" (Järvenoja & Järvelä, 2013, p.177).

10.1 Limitations

There are two important limitations with a potential impact: the number of participants (N=12) and the researcher's limited knowledge on psychophysiological measurement. Since this is an exploratory study, a larger number of participants would have required an extensive research team and resources for proper analysis and statistical attempts as described in the method research section of this study. The analyses have been performed based on descriptive quantitative analysis. This analysis limited the possibilities for interpretation and generalization. As for the limited knowledge, there is an impact on broader comparisons and in-depth signal analysis.

10.2 Practical implications

Understanding arousal effects can help teachers direct their efforts to create more favourable conditions to improve learning processes and outcomes. By using biosensors and proper dashboards, teachers can provide support by changing the environment (adjusting challenge levels, for instance) or interacting with students to help them find their optimal arousal level for a learning activity (Pijeira-Díaz et al., 2018).

Since arousal affects cognition, memory, and performance (Cattell & Child, 1975; Eysenck, 2012), it is important to realize – as precise as possible – how a student responds emotionally speaking about tasks, roles in groups, social challenges to focus on concrete actions to create a balance between cognition and emotion. The measurement process provides data and insights to follow a student's physiological arousal over a period of time and find ways to intervene and provide emotional support.

The arousal-valence space can contribute to understand how students facing the same challenge can react differently, depending on their ability to handle a challenge (Nakamura & Csikszentmihalyi, 2009). Measurement data from arousal and valence can provide a picture of a student from an emotional perspective, based on Russell's circumplex model of affect (Russell, 1980).

10.3 Future directions

In further studies, larger samples of students will make it possible to try inferential statistics to identify trends in arousal levels. It may also be considered a teacher's intervention as discrete stimuli so that specific events in the study can be measured for arousal.

It is recommendable to combine other physiological measures such as facial expression recognition and heart rate to run more detailed analysis on student's performance during the collaborative activities (cf. Henning et al., 2001; Henriques et al., 2013; Lei et al., 2017).

References

Andersson, J., & Rönnberg, J. (1995). Recall suffers from collaboration: Joint recall effects of friendship and task complexity. *Applied cognitive psychology*, *9*(3), 199-211.

Andriessen, J., Baker, M., & van der Puil, C. (2011). Socio-cognitive tension in collaborative working relations. *Learning across sites: new tools, infrastructures and practices*, 222-242.

Andriessen, J., Pardijs, M., & Baker, M. (2013). Getting on and getting along. *Affective learning together: Social and emotional dimensions of collaborative learning*, 205-230.

Aronson, E., & Patnoe, S. (2011). *Cooperation in the classroom: The jigsaw method* (3rd ed.). London, UK: Pinter & Martin.

Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist*, *50*(1), 84-94.

Baker, M., Andriessen, J., & Järvelä, S. (Eds.). (2013). *Affective learning together: Social and emotional dimensions of collaborative learning*. Routledge.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited.

Barnes, Barry. *Scientific knowledge and sociological theory*. Routledge, 2013.

Barrett, L. F. (2006a). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and social psychology review*, *10*(1), 20-46.

Barrett, L. F. (2006b). Valence is a basic building block of emotional life. *Journal of Research in Personality*, *40*(1), 35-55.

Barrett, L. F., & Russell, J. A. (1999). The structure of current affect: Controversies and emerging consensus. *Current directions in psychological science*, *8*(1), 10-14.

Barron, B. (2003). When smart groups fail. *The journal of the learning sciences*, *12*(3), 307-359.

Baumgartner, T., Willi, M., & Jäncke, L. (2007). Modulation of corticospinal activity by strong emotions evoked by pictures and classical music: a transcranial magnetic stimulation study. *Neuroreport*, *18*(3), 261-265.

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. Journal of neuroscience methods, 190(1), 80-91.

Berman, J. R., Berman, L. A., Lin, H., Flaherty, E., Lahey, N., Goldstein, I., & Cantey-Kiser, J. (2001). Effect of sildenafil on subjective and physiologic parameters of the female sexual response in women with sexual arousal disorder. *Journal of Sex &Marital Therapy*, *27*(5), 411-420.

BIOPAC Systems, Inc. (2018). *EDA: Electrodermal Activity | Event-Related EDA Analysis | Research | BIOPAC*. [online] Available at: https://www.biopac.com/application/electrodermal-activity/advanced-feature/event-related-eda-analysis/ [Accessed 19 Aug. 2018].

Blachly, P. H. (1970). Sensory Deprivation: Fifteen Years of Research. *Archives of General Psychiatry*, *22*(1), 90-91.

Blain, S., Mihailidis, A., & Chau, T. (2008). Assessing the potential of electrodermal activity as an alternative access pathway. *Medical engineering & physics*, *30*(4), 498-505.

Blair, C. (2010). Stress and the development of self- regulation in context. *Child development perspectives*, *4*(3), 181-188.

Boekaerts, M. (2007). Understanding students' affective processes in the classroom. In *Emotion in education* (pp. 37-56).

Boucsein, W. (2012). *Electrodermal activity*. Springer Science & Business Media.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, *25*(1), 49-59.

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

Bruner, J. S. (1996). *The culture of education*. Harvard University Press.

Cahour, B. (2013). Characteristics, emergence and circulation in interactional learning. *Affective learning together*, 52-70.

Cattell, R. B. (1972). The nature and genesis of mood states: A theoretical model with experimental measurements concerning anxiety, depression, arousal, and other mood states. In *Anxiety: Current Trends in Theory and Research, Volume 1* (pp. 115-183).

Cattell, R. B., & Child, D. (1975). *Motivation and dynamic structure*. John Wiley & Sons.

Chanel, G., Ansari-Asl, K., & Pun, T. (2007, October). Valence-arousal evaluation using physiological signals in an emotion recall paradigm. In *Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on* (pp. 2662-2667). IEEE.

Christopoulos, G. I., Uy, M. A., & Yap, W. J. (2016). The body and the brain: measuring skin conductance responses to understand the emotional experience. *Organizational Research Methods*, 1094428116681073.

Ciuk, D., Troy, A., & Jones, M. (2015). Measuring emotion: Self-reports vs. physiological indicators.

Conrad, S. R., & Wincze, J. P. (1976). Orgasmic reconditioning: A controlled study of its effects upon the sexual arousal and behavior of adult male homosexuals. *Behavior Therapy*, *7*(2), 155-166.

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). Elsevier.

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, United Kingdom: Cambridge University Press.

D'Mello, S., Hays, P., Williams, C., Cade, W., Brown, J., & Olney, A. (2010, June). Collaborative lecturing by human and computer tutors. In *International Conference on Intelligent Tutoring Systems* (pp. 178-187). Springer, Berlin, Heidelberg.

Do, S. L., & Schallert, D. L. (2004). Emotions and Classroom Talk: Toward a Model of the Role of Affect in Students' Experiences of Classroom Discussions. *Journal of educational psychology*, *96*(4), 619.

Driver, R., Asoko, H., Leach, J., Scott, P., & Mortimer, E. (1994). Constructing scientific knowledge in the classroom. *Educational researcher*, *23*(7), 5-12.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta- analysis of school-based universal interventions. *Child development*, *82*(1), 405-432.

Eiser, J. R. (1986). *Social psychology: Attitudes, cognition and social behaviour*. Cambridge University Press.

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*(6), 997-1003.

Eysenck, M. (2012). *Attention and arousal: Cognition and performance*. Springer Science & Business Media.

Fan, Y. T., Chen, C., Chen, S. C., Decety, J., & Cheng, Y. (2013). Empathic arousal and social understanding in individuals with autism: evidence from fMRI and ERP measurements. *Social cognitive and affective neuroscience*, *9*(8), 1203-1213.

Feldman, L. A. (1995). Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of personality and social psychology*, *69*(1), 153.

Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of personality and social psychology*, *74*(4), 967.

Fischer, K. W., & Immordino-Yang, M. H. (2007). *The Jossey-Bass reader on the brain and learning*. John Wiley & Sons.

Fonteyn, M. E., Kuipers, B., & Grobe, S. J. (1993). A description of think aloud method and protocol analysis. *Qualitative health research*, *3*(4), 430-441.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American psychologist*, *56*(3), 218.

Gabella, G. (2001). Autonomic nervous system. *e LS*.

Garbarino, M., Lai, M., Bender, D., Picard, R. W., & Tognetti, S. (2014, November). 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). IEEE.

Garcia- Prieto, P., Bellard, E., & Schneider, S. C. (2003). Experiencing diversity, conflict, and emotions in teams. *Applied psychology*, *52*(3), 413-440.

Garrison, D. R. (1997). Self-directed learning: Toward a comprehensive model. *Adult education quarterly*, *48*(1), 18-33.

Gillies, R. M., Carroll, A., Cunnington, R., Rafter, M., Palghat, K., Bednark, J., & Bourgeois, A. (2016). Multimodal representations during an inquiry problem-solving activity in a Year 6 science class: A case study investigating cooperation, physiological arousal and belief states. *Australian Journal of Education*, *60*(2), 111-127.

Goh, D. Y., Galster, P., & Marcus, C. L. (2000). Sleep architecture and respiratory disturbances in children with obstructive sleep apnea. *American journal of respiratory and critical care medicine*, *162*(2), 682-686.

Gokhale, A. A. (2012). Collaborative learning and critical thinking. In *Encyclopedia of the sciences of learning* (pp. 634-636). Springer US.

Greenwald, M. K., Cook, E. W., & Lang, P. J. (1989). Affective judgment and psychophysiological response: Dimensional covariation in the evaluation of pictorial stimuli. Journal of Psychophysiology, 3, 51–64.

Haataja, E., Malmberg, J., & Järvelä, S. (2018). Monitoring in collaborative learning: Co-occurrence of observed behavior and physiological synchrony explored. *Computers in Human Behavior*.

Haywood, T. W., Grossman, L. S., & Cavanaugh, J. L. (1990). Subjective versus objective measurements of deviant sexual arousal in clinical evaluations of alleged child molesters. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *2*(3), 269.

Henning, R. A., Boucsein, W., & Gil, M. C. (2001). Social–physiological compliance as a determinant of team performance. *International Journal of Psychophysiology*, *40*(3), 221-232.

Henriques, R., Paiva, A., & Antunes, C. (2013, September). Accessing emotion patterns from affective interactions using electrodermal activity. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference* (pp. 43-48). IEEE.

Hrastinski, S. (2008). The potential of synchronous communication to enhance participation in online discussions: A case study of two e-learning courses. *Information & Management*, *45*(7), 499-506.

Immordino□Yang, M. H., & Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, brain, and education*, *1*(1), 3-10.

Immordino-Yang, M. H., Christodoulou, J. A., Pekrun, R., & Linnenbrink-Garcia, L. (2014). Neuroscientific contributions to emotion measurement in educational contexts. *Handbook of Emotions in Education. New York: Routledge*, 607-624.

Isen, A. M. (1999). On the relationship between affect and creative problem solving. *Affect, creative experience, and psychological adjustment*, *3*, 17.

Isohätälä, J., Järvenoja, H., & Järvelä, S. (2017). Socially shared regulation of learning and participation in social interaction in collaborative learning. *International Journal of Educational Research*, *81*, 11-24.

Isohätälä, J., Näykki, P., Järvelä, S., & Baker, M. J. (2017). Striking a balance: Socio-emotional processes during argumentation in collaborative learning interaction. Learning, Culture and Social Interaction.

Järvenoja, H., & Järvelä, S. (2009). Emotion control in collaborative learning situations: Do students regulate emotions evoked by social challenges. *British Journal of Educational Psychology*, *79*(3), 463-481.

Järvenoja, H., & Järvelä, S. (2013). Regulating emotions together for motivated collaboration. *Affective learning together. Social and emotional dimensions of collaborative learning*, 162-182.

Järvenoja, H., Volet, S., & Järvelä, S. (2013). Regulation of emotions in socially challenging learning situations: An instrument to measure the adaptive and social nature of the regulation process. Educational Psychology, 33(1), 31-58.

Järvenoja, H., Malmberg, J., Järvelä, S., Näykki, P., & Kontturi, H. (2018). Investigating students' situation-specific emotional state and motivational goals during a learning project within one primary school classroom. *Learning: Research and Practice*, 1-20.

Jewitt, C. (2012). An introduction to using video for research.

Johnsen, B. H., Thayer, J. F., & Hugdahl, K. (1995). Affective judgment of the Ekman faces: A dimensional approach. *Journal of Psychophysiology*.

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational researcher*, *38*(5), 365-379.

Jones, A., & Issroff, K. (2005). Learning technologies: Affective and social issues in computer-supported collaborative learning. *Computers & Education*, *44*(4), 395-408.

Kettunen, J., Ravaja, N., Näätänen, P., Keskivaara, P., & Keltikangas-Järvinen, L. (1998). The synchronization of electrodermal activity and heart rate and its relationship to energetic arousal: A time series approach. *Biological Psychology*, *48*(3), 209-225.

Kirschner, F., Paas, F., & Kirschner, P. A. (2009). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior*, *25*(2), 306-314.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: theory and data. In *Psychology of learning and motivation* (Vol. 54, pp. 59-101). Academic Press.

Koivuniemi, M., Panadero, E., Malmberg, J., & Järvelä, S. (2017). Higher education students' learning challenges and regulatory skills in different learning situations/Desafíos de aprendizaje y habilidades de regulación en distintas situaciones de aprendizaje en estudiantes de educación superior. *Infancia y Aprendizaje*, *40*(1), 19-55.

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological psychology*, *84*(3), 394-421.

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. Psychological Bulletin, 139(4), 917.

Laal, M., & Laal, M. (2012). Collaborative learning: what is it?. *Procedia-Social and Behavioral Sciences*, *31*, 491-495.

Laal, Marjan, and Seyed Mohammad Ghodsi. "Benefits of collaborative learning." *Procedia-social and behavioral sciences* 31 (2012): 486-490.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*, 39-58.

Laniado, D., Kaltenbrunner, A., Castillo, C., & Morell, M. F. (2012, August). Emotions and dialogue in a peer-production community: the case of Wikipedia. In *WikiSym* (p. 9).

Lei, J., Sala, J., & Jasra, S. K. (2017). Identifying correlation between facial expression and heart rate and skin conductance with iMotions biometric platform. *Journal of Emerging Forensic Sciences Research*, *2*(2), 53-83.

Leiner, D., Fahr, A., & Früh, H. (2012). EDA positive change: A simple algorithm for electrodermal activity to measure general audience arousal during media exposure. *Communication Methods and Measures*, *6*(4), 237-250.

Linnenbrink, E. A. (2007). The role of affect in student learning: A multi-dimensional approach to considering the interaction of affect, motivation, and engagement. In *Emotion in education* (pp. 107-124). Academic Press.

Linnenbrink-Garcia, L., Rogat, T. K., & Koskey, K. L. (2011). Affect and engagement during small group instruction. *Contemporary Educational Psychology*, *36*(1), 13-24.

Loredo, J. S., Clausen, J. L., Ancoli-Israel, S., & Dimsdale, J. E. (1999). Night-to-night arousal variability and interscorer reliability of arousal measurements. *Sleep*, *22*(7), 916-920.

Loredo, J. S., Ziegler, M. G., Ancoli-Israel, S., Clausen, J. L., & Dimsdale, J. E. (1999). Relationship of arousals from sleep to sympathetic nervous system activity and BP in obstructive sleep apnea. *Chest*, *116*(3), 655-659.

Luria, A. R. (1973). An introduction to neuro-psychology. In *The working brain*. Penguin Books Harmondsworth.

Mandryk, R. L., Atkins, M. S., & Inkpen, K. M. (2006, April). A continuous and objective evaluation of emotional experience with interactive play environments. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 1027-1036). ACM.

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, *1*(2), 320-339.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and emotion*, *23*(2), 209-237.

Mehrabian, A. (1996). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, *14*(4), 261-292.

Moodie, C., & Finnigan, F. (2005). A comparison of the autonomic arousal of frequent, infrequent and non- gamblers while playing fruit machines. *Addiction*, *100*(1), 51-59.

Mønster, D., Håkonsson, D. D., Eskildsen, J. K., & Wallot, S. (2016). Physiological evidence of interpersonal dynamics in a cooperative production task. *Physiology & behavior*, *156*, 24-34.

Mönttinen, H., Koskimäki, H., Siirtola, P., & Röning, J. (2017). Electrodermal activity asymmetry in sleep-a case study for migraine detection. In *EMBEC & NBC 2017* (pp. 835-838). Springer, Singapore.

Mou, W., Celiktutan, O., & Gunes, H. (2015, May). Group-level arousal and valence recognition in static images: Face, body and context. In *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)* (Vol. 5, pp. 1-6). IEEE.

Nakamura, J., & Csikszentmihalyi, M. (2009). Flow theory and research. *Handbook of positive psychology*, 195-206.

National Research Council. (2000). *How people learn: Brain, mind, experience, and school: Expanded edition*. National Academies Press.

Näykki, P., Järvelä, S., Kirschner, P. A., & Järvenoja, H. (2014). Socio-emotional conflict in collaborative learning—A process-oriented case study in a higher education context. *International Journal of Educational Research*, *68*, 1-14.

Nicolaou, M. A., Gunes, H., & Pantic, M. (2011). Continuous prediction of spontaneous affect from multiple cues and modalities in valence-arousal space. *IEEE Transactions on Affective Computing*, *2*(2), 92-105.

Nummenmaa, M., & Nummenmaa, L. (2008). University students' emotions, interest and activities in a web- based learning environment. *British Journal of Educational Psychology*, *78*(1), 163-178.

Oxendine, J. B. (1970). Emotional arousal and motor performance. *Quest*, *13*(1), 23-32.

Pally, R. (1998). Emotional processing: The mind-body connection. *International Journal of Psycho-Analysis*, *79*, 349-362.

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*(2), 99-141.

Panadero, E., Klug, J., & Järvelä, S. (2016). Third wave of measurement in the self-regulated learning field: when measurement and intervention come hand in hand. *Scandinavian Journal of Educational Research*, *60*(6), 723-735.

Panitz, T. (1999). Collaborative versus Cooperative Learning: A Comparison of the Two Concepts Which Will Help Us Understand the Underlying Nature of Interactive Learning.

Pekrun, R. (1992). The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology*, *41*(4), 359-376.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002a). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational psychologist*, *37*(2), 91-105.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002b). Positive emotions in education.

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., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary educational psychology*, *36*(1), 36-48.

Pekrun, R. (2016). Academic emotions. *Handbook of motivation at school*, *120*.

Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: the Epistemically-Related Emotion Scales. *Cognition and Emotion*, *31*(6), 1268-1276.

Phye, G. D., Schutz, P., & Pekrun, R. (2011). *Emotion in education*. Elsevier.

Piaget, J. (1971). Biology and knowledge: An essay on the relations between organic regulations and cognitive processes.

Pijeira-Díaz, H. J., Drachsler, H., Järvelä, S., & Kirschner, P. A. (2016, April). Investigating collaborative learning success with physiological coupling indices based on electrodermal activity. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 64-73). ACM.

Pijeira-Díaz, H. J., Drachsler, H., Kirschner, P. A., & Järvelä, S. (2018). Profiling sympathetic arousal in a physics course: How active are students? *Journal of Computer Assisted Learning*.

Renninger, K. A., & Hidi, S. (2015). *The power of interest for motivation and engagement*. Routledge.

Revelle, W., & Loftus, D. A. (1992). The implications of arousal effects for the study of affect and memory. The handbook of emotion and memory: Research and theory, 113-149.

Ringeval, F., Schuller, B., Valstar, M., Jaiswal, S., Marchi, E., Lalanne, D., ... & Pantic, M. (2015, October). Av+ ec 2015: The first affect recognition challenge bridging across audio, video, and physiological data. In *Proceedings of the 5th International Workshop on Audio/Visual Emotion Challenge* (pp. 3-8). ACM.

Ringeval, F., Sonderegger, A., Sauer, J., & Lalanne, D. (2013, April). Introducing the RECOLA multimodal corpus of remote collaborative and affective interactions. In *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)* (pp. 1-8). IEEE.

Rockloff, M. J., Signal, T., & Dyer, V. (2007). Full of sound and fury, signifying something: The impact of autonomic arousal on EGM gambling. *Journal of Gambling Studies*, *23*(4), 457-46.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

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*(5), 805.

Sakai, T., Tamaki, H., Ota, Y., Egusa, R., Inagaki, S., Kusunoki, F., ... & Mizoguchi, H. (2017). Eda-Based Estimation of Visual Attention by Observation of Eye Blink Frequency. *International Journal on Smart Sensing and Intelligent Systems*, *10*(2), 296-307.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social science information*, *44*(4), 695-729.

Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.). (2001). *Appraisal processes in emotion: Theory, methods, research*. Oxford University Press.

Schlosberg, H. (1954). Three dimensions of emotion. Psychological review, 61(2), 81.

Schutz, P. A., Hong, J. Y., Cross, D. I., & Osbon, J. N. (2006). Reflections on investigating emotion in educational activity settings. *Educational Psychology Review*, *18*(4), 343-360.

Schutz, P. A., Aultman, L. P., & Williams-Johnson, M. R. (2009). Educational psychology perspectives on teachers' emotions. In *Advances in teacher emotion research* (pp. 195-212). Springer, Boston, MA.

Schwarz, N. (2002). Situated cognition and the wisdom of feelings: Cognitive tuning. *The wisdom in feelings*, 144-166.

Semmer, N. K., Grebner, S., & Elfering, A. (2003). Beyond self-report: Using observational, physiological, and situation-based measures in research on occupational stress. In *Emotional and physiological processes and positive intervention strategies* (pp. 205-263). Emerald Group Publishing Limited.

Shaffer, F., Combatalade, D., Peper, E., & Meehan, Z. M. (2016). A Guide to Cleaner Electrodermal Activity Measurements. *Biofeedback*, *44*(2), 90-100.

Shouse, E. (Dec. 2005) "Feeling, Emotion, Affect," *M/C Journal*, 8(6). Retrieved from <http://journal.media-culture.org.au/0512/03-shouse.php>.

Smith, B. L., & MacGregor, J. T. (1992). What is collaborative learning.

Spangler, G., Pekrun, R., Kramer, K., & Hofmann, H. (2002). Students' emotions, physiological reactions, and coping in academic exams. *Anxiety, Stress & Coping*, *15*(4), 413-432.

Taylor, S., Jaques, N., Chen, W., Fedor, S., Sano, A., & Picard, R. (2015, August). Automatic identification of artifacts in electrodermal activity data. In *2015 37th Annual International*

Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 1934-1937). IEEE.

Taylor, S. P., & Epstein, S. (1967). The measurement of autonomic arousal: Some basic issues illustrated by the covariation of heart rate and skin conductance. *Psychosomatic Medicine*, *29*(5), 514-525.

Thayer, R. E., Takahashi, P. J., & Pauli, J. A. (1988). Multidimensional arousal states, diurnal rhythms, cognitive and social processes, and extraversion. *Personality and Individual Differences*, *9*(1), 15-24.

Thomson, L., & Fine, G. A. (1999). Socially shared cognition, affect, and behavior: A review and integration. *Personality and Social Psychology Review*, *3*(4), 278-302.

Törmänen, T., Järvenoja, H., Kurki, K., Järvelä, S., & Devai, R. (2018, August). *Students' emotional valence and physiological arousal during collaborative learning*. Poster presented at 16th International Conference on Motivation 2018, Aarhus, Denmark.

Valenza, G., Citi, L., Lanatá, A., Scilingo, E. P., & Barbieri, R. (2014). Revealing real-time emotional responses: a personalized assessment based on heartbeat dynamics. *Scientific reports*, *4*, 4998.

Vauras, M., Iiskala, T., Kajamies, A., Kinnunen, R., & Lehtinen, E. (2003). Shared-regulation and motivation of collaborating peers: A case analysis. *Psychologia*, *46*(1), 19-37.

Veenman, M. V. J. (2005). The assessment of metacognitive skills: What can be learned from multi-method designs? In C. Artelt & B. Moschner (Eds.), *Lernstrategien und Metakognition: Implikationen für Forschung und Praxis* (pp. 75–97). Berlin: Waxmann.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In *Handbook of self-regulation* (pp. 531-566). Academic Press.

Winton, W. M., Putnam, L. E., & Krauss, R. M. (1984). Facial and autonomic manifestations of the dimensional structure of emotion. *Journal of Experimental Social Psychology*, *20*(3), 195-216.

Woods, S., & White, E. (2005). The association between bullying behaviour, arousal levels and behaviour problems. *Journal of adolescence*, *28*(3), 381-395.

Zhang, S., Tian, Q., Jiang, S., Huang, Q., & Gao, W. (2008, June). Affective MTV analysis based on arousal and valence features. In *2008 IEEE International Conference on Multimedia and Expo* (pp. 1369-1372). IEEE.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American educational research journal*, *45*(1), 166-183.

Zubek, J. P. (1969). *Sensory deprivation: Fifteen years of research*. Appleton-Century-Crofts and Fleschner Publishing Company.

Zuckerman, M. (2014). *Sensation seeking (psychology revivals): beyond the optimal level of arousal*. Psychology Press.