Leaders and followers identified by emotional mimicry during collaborative learning: A facial expression recognition study on emotional valence

Muhterem Dindar, Sanna Järvelä, Sara Ahola, Xiaohua Huang, Guoying Zhao, *Senior Member, IEEE*

Abstract—**This study explores the potential of emotional mimicry in identifying the leader and follower students in collaborative learning settings. Our data include video recorded interactions of 24 high school students who worked together in groups of three during a collaborative exam. A facial emotions recognition method was used to capture participants' facial emotions during the collaborative work. Cross-recurrence quantification analysis was applied on the detected facial emotions to see the level and direction of emotional mimicry among the dyads in the same groups. In order to validate the cross-recurrence quantification analysis results, student interactions in terms of leading or following the task were video coded. Our findings showed that the leaders and followers identified by cross-recurrence quantification analysis findings matched the leaders and followers identified by the video coding in 70% of the dyadic interactions across the collaborating groups. The current findings show that video-based facial emotions recognition as a method can add to collaborative learning research, especially explaining some social, and affective dynamics about it. The study further discusses the possible variables that might confound the relationship between emotional mimicry and leader-follower interactions during collaboration.**

Index Terms— **Collaborative learning, educational technology, emotional mimicry, facial expressions recognition, multimodality.**

I. INTRODUCTION

COLLABORATIVE learning is inherently social and interactive [1]. It is known that various socio-emotional conflicts in interactions [2] emerge during collaborative learning, which may be, for example, due to the task content or the relationships among the learners [3]. Consequently, a wide range of emotions are expressed during collaborative learning. This is because emotional expressions serve crucial functions during social interaction such as adjustment to the social environment [4], conveying information [5], facilitating mutual understanding [6], and giving feedback [7]. The emotions that arise during collaboration might change the interactions among the group members and have detrimental effects on team performance if not managed properly [8]. It was found that increased inflammatory and critical expressions slows down the trust building and decreases help seeking among the team members [9]. Positive socioemotional interactions, for example, active listening, and respectful and supportive interactions, were associated with higher social regulation and better collaboration [10]. Drawing on such findings, socio-cognitive view on education asserts that effective learning in collaborative settings is dependent on shared efforts of group members to modify and direct emotions towards task accomplishment [2].

Acknowledging the importance of emotions in learning, a growing body of research has focused on how emotional signals are understood and utilized by the learners. New analytical approaches such as facial recognition analysis offer new venues for studying emotional contagion in collaborative learning settings. In light of this, the current study focuses on an underexplored phenomenon in emotional contagion, emotional mimicry, and presents empirical evidence about its potential to identify the leader and follower students during collaborative learning. The study contributes to the field from several aspects. First, research has shown that success in group work depends on the nature of interactions among the group members [11]. In addition, it is known that group interaction is based on leader and follower roles [12], but their contribution to socioemotional group processes is hard to characterize. Therefore, studying group dynamics from a leader-follower perspective is important in terms of characterizing collaborative processes. Nevertheless, assessing group dynamics in collaborative learning has been a major challenge in educational sciences [13]. Specifically, little progress has been made in developing objective measures of leader and follower interactions in

Muhterem Dindar, Sanna Järvelä, and Sara Ahola are with the Learning and Educational Technology (LET) research group, University of Oulu, Oulu 90014, Finland.

Xiaohua Huang is with the School of Computer Engineering, Nanjing Institute of Technology, Nanjing, China.

Guoying Zhao is with the Center for Machine Vision and Signal Analysis, University of Oulu, Oulu 90014, Finland.

Email:{muhterem.dindar, sanna.jarvela, sara.ahola. guoying.zhao, xiaohua.huang}@oulu.fi

collaboration. The current study address this gap through exploring the potential of emotional mimicry in detecting the leadership and followership patterns during collaborative work objectively. Detecting leaders and followers in collaborative settings might provide valuable insights into collaboration dynamics, and might open new paths to facilitate data-driven support to the group members in terms of participation equality and leadership development. Further, as a response to some methodological limitations about studying emotions in learning settings, this study introduces a novel method, cross-recurrence quantification analysis (CRQA) that can inform about emotional contagion among the collaborating learners.

In the following parts, we first introduce the phenomenon of emotional mimicry and its role in diffusion of emotions in social contexts. Then, we discuss how emotional mimicry is reflected in leader-follower interactions in collaborative settings. Following, we address several challenges in studying emotions in educational sciences and demonstrate how CRQA can be applied to detect leader and follower students in a collaborative task. We conclude by discussing our findings in light of relevant literature and highlight some future research directions on the topic.

II. THEORETICAL FRAMEWORK

Emotions are relatively short-term affective states that arise due to particular physical and social stimulus [14], [15]. They are comprised of individuals' appraisals about their ongoing interaction with the world [16]. Emotions carry two basic features, valence and arousal [17]. Valence refers to the perceived pleasantness of a stimulus (e.g. positive or negative), and arousal indicates the intensity of the emotion elicited by the stimulus [18]. Emotions are subjective experiences. That is, the type and intensity of emotions appear in response to the same situation might differ from one individual to the other [16]. However, emotions can also diffuse among the individuals through social interactions [19]. For example, it was found that emotions such as happiness, anxiety, and depression are disseminated among the connected individuals [20]. Consequently, diffusion of emotions from one individual to the others have implications for collaborative learning, in particular.

A. Emotional contagion, emotional mimicry and collaborative learning

A well-known phenomenon about human interaction is that psychological states are contagious and, specifically people "catch" others' emotions [21]. The contagion of emotions among the interacting individuals functions as a social regulator and facilitates transmission of a particular experience [22]. In the context of collaborative learning, it has been found that emotions spread among the group members through cycle of interactions [10]. Cycle of interactions with positive (e.g. joy, enthusiasm) or negative affect (e.g. frustration, boredom) were found to set the group climate and influence the quality of learning [23].

In social environments, such as collaborative learning, emotional contagion occurs through individuals' copying of each other's emotional expressions [24]. For example, it has been found that one's facial expressions (e.g. angry, sad or happy) in a social setting can trigger corresponding facial expressions in the others [25], [26]. Several studies have also found that people copy virtual agents' emotions in human-computer interaction environments [27], [28]. In the literature, copying or imitation of others' emotional expressions is coined as emotional mimicry. Emotional mimicry is a key mechanism in emotional contagion [29]. It facilitates experiencing of companions' emotions in one's self. The feedback provided to the companions after such an embodied experience develops emotional contagion [30].

Although occurring spontaneously, emotional mimicry is not a meaningless reflexive synchronization between the individuals. Rather, it is intrinsically meaningful, and depends on the relationship between expresser and observer, and the social context [31]. Studies have shown that emotional mimicry is an indicator of initial affiliative bond and empathy among the people [31], [32]. Higher levels of emotional mimicry has been observed among the close friends than strangers [33] and among the group members than out-group members [34]. It has been also found that emotional mimicry is less likely to occur in antagonistic and competitive contexts [35]. Further, congruent emotions among the individuals were found to facilitate successful interaction [36]. Drawing on this, it can be assumed that emotional mimicry increases social coordination and group cohesion through regulating one's relation to the others [31], [37].

One aspect of emotional mimicry is that it can be affected by dominant or submissive characteristics of interaction [38]. Several studies have shown that individuals at a lower social status are more prone to attend to the individuals with high social status and mimic their emotions [39], [40]. During collaboration, dominance is reflected as being active in terms of leading the task or socio-emotions [41]. Studies have shown that individuals who had higher verbal participation to the collaborative tasks are perceived as leaders by the others [42], [43]. Dominant and submissive interactions have been found to emerge gradually in groups that are even comprised of total strangers or individuals with equal status [44]. Thus, scholars have asserted that the social inequality is an essential feature and the driver of all collaborative acts [45], and individuals spontaneously accept dominant or submissive roles as the collaboration process unfolds [46].

As our review shows, most evidence about emotional mimicry comes from social psychology research, so far only few studies explored emotional mimicry in learning contexts. For example, Bakker [47] found a significant association between the self-reported flow experiences of music teachers and their students. In another survey study, Frenzel and colleagues [48] found that teachers' enthusiasm mediated the significant relationship between teachers' and students' enjoyment of mathematics classes. Finally, Becker and colleagues [49] repeatedly measured teachers' and students' emotions during fifteen sessions of learning with an experience sampling method. Confirming past findings, they found significant relationships between teachers' emotions,

instructional methods and students' emotions. Although these studies yielded valuable insights about the phenomenon of emotional mimicry in learning contexts, their findings carry some limitations in terms of capturing emotions in learning environments, especially in collaborative learning context.

B. Challenges in researching emotions in learning and video-based facial emotions recognition

Emotions are transient feelings that take place in short durations [50]. Therefore, type and intensity of emotions change rapidly and continually during collaborative learning depending on the task progress and relationships among the learners. Thus, it is challenging to identify the emotions and emotional contagion at a fine-grained level with respect to its temporality. For example, mainstream measures such as self-reports are not capable of capturing the short term fluctuations in emotions during learning [51]. Thus, they mostly inform about the overall emotional climate in a learning setting rather than providing a meticulous view [48]. One way to tackle the limitations of self-reports in emotions research would be to present them repeatedly to the participants during learning (e.g. experience sampling) [49]. However, it also known that asking participants to report about their current emotions might actually alter their emotions [52]. Therefore, it is very difficult to capture emotions reliably as they occur, specifically with self-reports. Considering this, several studies have embraced video analysis method to explore the emotions of learners (e.g. [53]). Although video analysis can provide a more detailed understanding on how emotions rise and emotional contagion take place during learning, it is a laborious and time-taking methodology. Further, a common practice to infer the emotional states of the learners with video analysis has been to code their verbal utterances during collaboration. However, it is known that people can easily express their emotions through facial expressions without any verbalization. Thus, embracing the video analysis method that solely focuses on verbal utterances might not be able to capture the whole variety of emotions in a learning environment.

Another challenge about studying the transitory nature of emotions is that emotions might shift as the learners' appraisals of the task change during learning [52]. For example, a particular learning activity that creates positive emotions at the beginning of a collaborative task might create negative emotions towards the end of the task. Thus, simply looking at the frequency of positive or negative emotions does not provide a comprehensive understanding on the quality of collaboration or learning. For that reason, it is necessary to adopt process-oriented methods that facilitate capturing and analysis of emotional transitions during learning with respect to the social context [54].

Given that, a promising method to identify emotions in collaborative learning settings would be video-based facial emotions recognition. A video is comprised of still images (i.e. frames) that are displayed at a specific rate within a time unit. With the recent technological advancements, it is possible to identify facial emotions of individuals from a video sequence through frame by frame analysis. This allows to identify the changes in facial emotions at very short time intervals (1/3 to 1/25 seconds) which is very difficult to detect with a human eye [55]. The method can provide a detailed view about the transitions of student emotions from beginning to the end of a learning session unobtrusively. So far, only few studies have utilized facial emotions recognition technology to explore a variety of educational phenomena such as conceptual change [56], conceptual conflict [57], communication skills training [58], intelligent education interfaces [5], student engagement [59], and educational game evaluation [60], yet mainly at the individual level. The research on exploring the collaborative learning processes with video-based facial emotions recognition is still nascent.

Overall, the literature states that emotional mimicry in social settings is influenced by the status of the interacting individuals [61]. Specifically, individuals with less status were found to imitate the emotions of others with high status [62]. The past findings also showed that active participation in a collaborative task can be an indicator of high status and leadership within a group. However, the association between emotional mimicry and leader-follower interactions has not been tested in collaborative learning. Drawing on this, the aim of the current study is to investigate whether emotional mimicry can help to identify the leader and follower group members in a collaborative task. Based on this aim the research question and the hypothesis of the study are:

Research question: Can we identify the leader and follower students in a collaborative task by looking at the emotional mimicry between them?

Hypothesis: In a collaborative task, facial emotions of the leader students will be mimicked more by the followers in the same group compared to vice versa.

III. METHODOLOGY

A. Participants and context

Participants of the study were 24 high school students (8 female, 16 male) from a Finnish secondary school. Their ages varied between fifteen and sixteen. The study was conducted in an Advanced Physics course that was optional to register. Participation to the study was voluntary and ethical permissions were granted from both the school administration of the participants and university of the researchers. At the beginning of the course, students were heterogeneously divided into groups of three by the subject teachers and participated in fifteen sessions of collaborative learning activities throughout the school term. The data in the current study only includes the collaborative exam session which students completed with their own group at the end of the term. The session was chosen, since the previous studies have shown that emotional exchanges among the students are more common in challenging situations [63]. Considering this, it was decided that an exam situation would elicit more emotions compared to other collaborative learning sessions. The study was conducted in LeaForum research infrastructure (https://www.oulu.fi/leaf-eng/) at the University of Oulu. Leaforum is equipped with advanced video and audio capture technology that allows to record

collaborative learning activities in a classroom-like setting.

B. Procedure

Prior to the collaborative exam, wireless microphones were attached to all participants. This allowed to record each participant's voice with minimal interference from other voices in the environment. Each group of students was then seated at a separate half-hexagon shaped table located in front of a video camera. The video camera recorded the groups from the point of view facing directly to the student at the center of the group. The Ladybug3 Firewire camera was used for video recording. The recording was done at 16 frame per second with a resolution of 1600X1200 pixels. The seating plan and video camera location is presented in Fig. 1. After the seating arrangement, collaborative groups were asked to conduct a hands-on physics experiment and report their findings as a
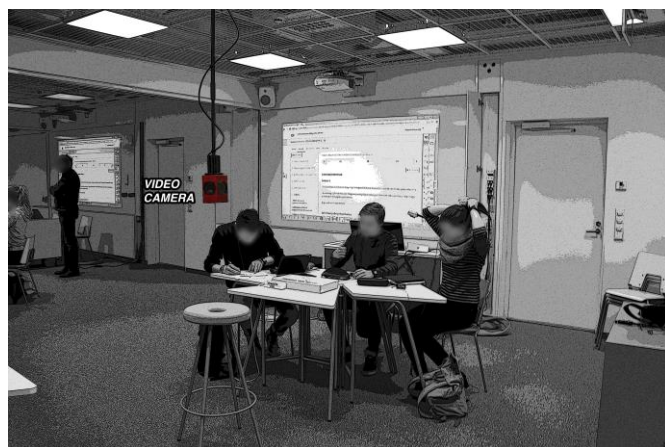

Fig. 1. Data collection setting.

group. The task was to determine the refractive index of light for water by using the necessary equipment present at the table and write down the setting of their experiment, all the calculations and argumentations related to the experiment. No restrictions about moving head or body were announced to the participants. The instructions about the exam task was explained to the classroom by the teacher. Participants completed the exam task without any direct help from the teacher. However, teachers made some public announcement to the classroom during the exam to elaborate some issues with the task.

IV. DATA AND ANALYSIS

The data in the current study includes the video recordings of the participants during the aforementioned collaborative exam session. Participant interactions were examined with both video coding and video-based facial emotions recognition analysis. The aim of the video coding was to identify how long each member in a group lead or followed others during the collaboration. Video-based facial emotions recognition analysis allowed us to classify the participants' facial emotions valence as positive, negative or neutral in mutually exclusive categories. Then a CRQA [64] was run on classified facial emotions to detect the leader and follower students in terms of emotional

mimicry during dyadic interactions within the group. The CRQA results were then matched with the video coding findings to check whether it is possible to detect the leader and follower students during collaborative work from the level and direction of emotional mimicry between them.

A. Video coding

Based on the Nonverbal Indexes of Students' Physical Interactivity (NISPI) framework [13], a coding scheme was developed to describe participants' task leading and following behaviors based on their verbal utterances and physical actions during collaboration. The scheme allows assigning a separate code to each group member during a collaborative task for each 1-second clip. That is, for each second of collaboration, the off-task code (0) is assigned to a participant when she/he is not doing any task related activity (e.g. using a mobile phone and placing books in the backpack). The individual task behavior code (1) is used when the participant works on the task alone without interacting with anyone (e.g. taking notes, reading a textbook). The task follow code (2) is used when the participants' eye gaze follows another group member during the task (e.g. listening to a group member, watching a group member doing the experiment, and approving a group member with nodding the head or saying "yes"). The task lead code (3) is assigned when the participant actively leads the collaborative task with his verbal utterances or physical actions (e.g. talking about how to progress with the task, leading an experiment or hands-on activity, explaining a task related issue to the other(s).

TABLE I
SAMPLE VIDEO CODING OF LEADER AND FOLLOWER INTERACTIONS WITHIN A COLLABORATING GROUP

Time(s)	M1	M2	M3	Group code
1	0	0	0	All group members are off-task.
2	0	0	1	M1 and M2 are off-task. M3 is doing individual task activities.
3	1	0	3	M1 is doing individual task activites. M2 is off-task. M3 is leading the task.
4	2	3	1	M1 is following M2. M2 is leading the task. M3 is doing an individual task-related activity.
5	3	2	2	M1 is leading the task. M2 and M3 are following M1.
N	3	3	2	M1 and M2 are leading the task together. M3 is following them.

M=Member

The teacher follow code (4) is assigned when a participant's eye gaze follows the teacher when the teacher gives instructions or explanations. Each student was coded separately on the same timeline. The individual audio recordings were combined with the video data when coding each participants' behaviors during collaboration. Then, the individual codes were synchronously combined as group codes. The group codes allowed us to quantify the leader and follower interactivity among the group members for each second. Table I summarizes the overall coding process. All group interactions were coded by two independent researchers. The researchers had sufficient previous experiences in video coding. They were trained on a

small sub-sample of video data prior to the actual coding with the current coding scheme.

The Cohen's Kappa (κ) values, an indicator of the consistency of the coding by the researchers, varied between 0.62 and 0.89 among the groups (Group 1=0.84, Group 2=0.76, Group 3=0.79 Group 4=0.62, Group 6=0.75, Group 7=0.87, Group 8=0.79) except than Group 5 (κ=0.37). It was observed that the low consistency in Group 5 was due to a participant turning his/her back to the video camera often during the collaborative task. Consequently, his/her interactions with the other group members could not be coded reliably. Thus, findings about the Group 5 should be interpreted cautiously.

B. Video-based facial emotions recognition

For facial emotion recognition, it is important to use robust and efficient face detector. In our implementation, we use the face detection and alignment [65] to detect and then align the faces, because it has been demonstrated to be robust to the challenging environment and different pose. The parameter of face model and landmark detection is the same to [65].
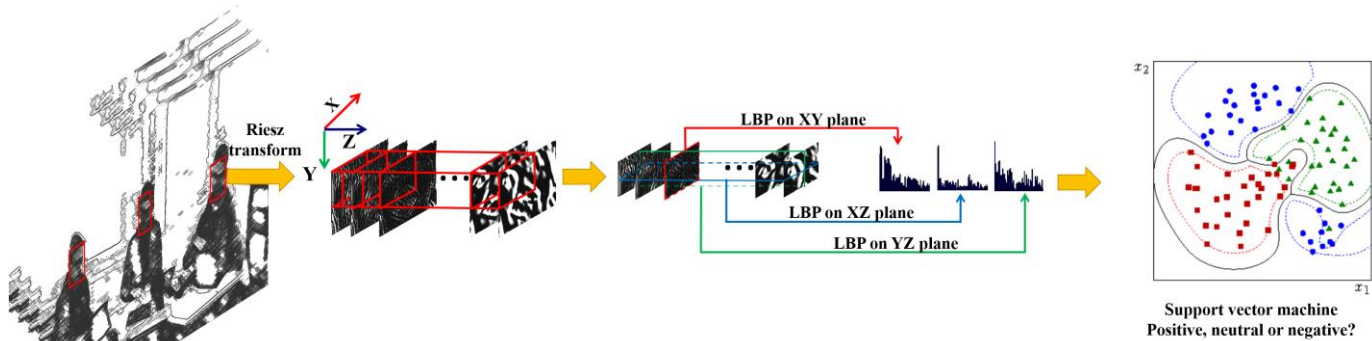
outperformed two baseline algorithms, namely LBP (43.49%) and VGG-face pretrained model (58.21%). The whole results on four commonly used databases solidly shows our method is reliable in valence prediction.

For RVLBP, we followed the parameter setup of [70]. Specifically, we firstly divided a face into 4×4 overlapped blocks, and design the commonly used log-Gabor filter with 5 scales and 8 orientations. Based on filter sets, we could obtain the first-order and second-order Riesz faces by convolving images with filters. Due to the efficient ability of LBP, we could easily obtain the discriminative and informative facial expression features. Finally, we used support vector machine (SVM) with linear kernel as classification. The penalty parameter of SVM is optimally selected by using three-fold cross validation protocol. The framework of individual's emotion recognition is shown in Fig. 2. For obtaining facial expression classification model, we chose the recently new facial expression database [71]. This database is collected from Google and Flickr images according to keywords. All images were labeled with three emotion categories (positive, neutral

Fig. 2. The framework of individual's emotion recognition, where the red, blue and green colors in the right image mean the feature distribution of the positive, neutral and negative emotion samples in support vector machine, respectively.

For evaluating the performance of a collaborating group, we design a robust and efficient pipeline for analyzing facial emotion recognition. Currently, there are quite many methods for facial expression analysis [66]–[69]. In this paper, we primarily focused on the feature extraction for extracting feature from each student in collaborative learning environment. For facial feature extractor, we implemented our recently proposed approach, namely, Riesz-based volume local binary pattern (RVLBP) [70]. This method has been extensively evaluated in different tasks of emotion recognition, and achieves the considerable results in the real-world databases. For example, our RVLBP method has been evaluated in our previous study [70]. It has been conducted in three facial expression tasks (spontaneous facial expression recognition, smile detection and happiness intensity estimation). Our previously experimental results had demonstrated that our RVLBP can achieve the promising results. Additionally, we conducted another experiments on the challenging GAFF database for valence prediction (positive, neutral and negative). RVLBP obtained the recognition rate of 67.02%, which

and negative).

A commonly used method [72] for face detection and landmark localization has been used in the current study. As in [72], the reliability of the current model for face detection and landmark localization have been extensively evaluated. The method utilized in the current study has been compared with OpenCV face detection, Deformable part model and Google Picasa's face detector. An average precision of 88.7% has been obtained on AFW dataset for face detection. It is believed that

Fig. 3. Face detection and landmark localization method.

it will handle mostly movement in our dataset. Additionally, the current method can handle the head rotation from -90 degree to 90 degree, as shown in the following Fig. 3.

C. Cross-recurrence quantification analysis

CRQA allows quantifying to what extend two streams of data co-visit the same state over time [73]. One advantage of CRQA is that it is a "non-linear" technique that is not restricted by the assumptions of the linear statistics [74]. CRQA has gained attention as a promising method in human sciences to investigate the dynamics between the interacting individuals. For example, it has been used to investigate the eye gaze alignment [73], facial gestures mimicry [27], emotional contagion [28], psycho-motor control [75], and turn-taking [76] among the interlocutors.

In this study, CRQA R package [64] was used to calculate the level of emotional mimicry among the members of the same collaborative group. The parameters of the analysis were ws = 80, datatype = continuous, radius = 0.001, delay = 1, embed = 1, rescale = 0, normalize = 0, mindiagline = 2, minvertline = 2, tw = 0. In CRQA, the alignment between two time series is represented as recurrence rate (RR). High RR means that two time series show highly similar patterns over time [64]. A recurrence in this research context is defined as two participants from the same group displaying same facial emotional expressions (i.e. positive, negative or neutral) at a specific time point. Since CRQA can only calculate the alignment between two time series, we calculated three separate RRs for each group to determine the dyadic emotional mimicry between the team members. In order to observe how and to what extend facial emotions of a specific participant is mimicked by others in the group, we calculated RR for time lags between [-5, +5] seconds which is considered as the maximum window length for conscious response of facial expression muscles to a stimulus [77]. Considering that facial emotions data is captured at 16 frames per second, [-5, +5] seconds refer to [-80, +80] lag interval in the dataset. In the current study, lag refers to the time difference between two time series. Analyzing lags in time series allows observing cause and effect relationships between the series (e.g. How the current behavior of a specific signal affects another signal's behavior one sec later?).

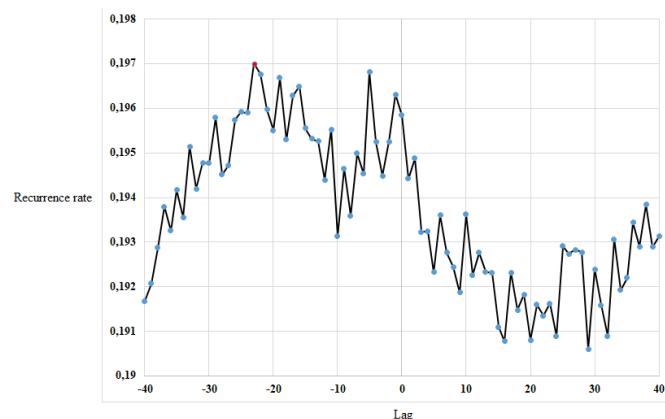Comparing the highest RR values at positive or negative lags

Fig. 4. Highest recurrence rates the collaborating participants at different time lags. The red joint indicates the highest recurrence rate.

can indicate the leader and follower participants [76]. The higher RR at a positive lag indicates that facial emotions of the student whose data is put ahead in the analysis is mimicked more by the student whose data is put behind. For example, In Fig. 4 it can be seen that the highest RR for Participant A and Participant B is observed at a negative lag which means that Participant A mimics Participant B more than Participant B mimics Participant A.

V. RESULTS

A. Descriptive statistics on facial emotional expressions

Our facial emotions detection algorithm revealed that positive facial expressions varied between 0,2% and 39,8%, negative facial expressions varied between 28% and 90,3%, and neutral expressions varied between 8,5% and 51% among the participants during the whole collaborative exam. The most prominent facial expression was negative emotions among most of the participants (n=21), followed by neutral (n=2) and positive (n=1) emotions. Missing values across the participants ranged from 13% to 78% of the total collaboration duration. This is understandable because participants were moving their body freely during the collaborative task. Thus, it was not possible to detect the facial expressions from the video when the participants' faces were diverted too much from the video camera. Nevertheless, the average valid frames captured per second ranged from 3 to 14. These values can be considered as sufficient to understand the alterations of facial emotional expressions during the collaborative task and conduct further analyses.

B. Leaders and followers in the groups, based on emotional mimicry

The highest RR values and their representative lags that were calculated with CRQA analysis are presented in Table II. The highest RR indicates the maximum observed emotional mimicry between the participants during collaboration. The lag column in Table II indicates the time lag at which the maximum emotional mimicry occurred. The ">" and "<" symbols indicate the direction of emotional mimicry (i.e. Who mimicked who?) For example, by looking at the Table II, we can say that Member 1 mimicked facial emotions of Member 2 more compared to vice versa. The highest emotional mimicry among the participants (i.e. RR=19,7%) was observed at a lag of 2 seconds. This means that Member1 was mimicking Member 2's facial emotions with a delay of 2 seconds.

In order to check that the CRQA results were not randomly obtained, a surrogate analysis CRQA was conducted by randomly shuffling data of each participant. In surrogate analysis, it is crucial that the actual data and the surrogate data have the same characteristics such as sampling frequency, and data length [78]. In the current study, the total task completion time was different for each group. Therefore, surrogate data was created by shuffling the data within the participants in order to preserve time distribution [79]. Then, a separate CRQA was conducted for each dyad in the groups with the surrogate data. Following, an independent samples t-test was run to see whether the CRQA results on actual data (n=24 dyads) differed

TABLE II
HIGHEST EMOTIONAL MIMICRY RECURRENCE RATES AMONG THE
COLLABORATING PARTICIPANTS

Group	Members	Highest RR value(%)	Lag (seconds)
1	M1<M2	19,7	2
1	M1>M3	23,5	5
1	M2<M3	25,6	5
2	M1< M 2	12,6	1
2	M 1> M 3	20,3	4
2	M 2> M 3	14,6	1
3	M 1> M 2	17,0	1
3	M 1> M 3	13,2	3
3	M 2> M 3	19,8	2
4	M 1> M 2	26,4	5
4	M 1< M 3	13,9	3
4	M 2<M 3	16,2	5
5	M 1> M 2	04,9	4
5	M 1> M 3	11,1	1
5	M 2>M 3	22,1	4
6	M 1>M 2	06,3	2
6	M 1< M 3	16,7	4
6	M 2< M 3	12,2	1
7	M 1> M 2	21,3	5
7	M 1< M 3	10,6	1
7	M 2< M 3	16,5	4
8	M 1> M 2	28,2	3
8	M 1> M 3	11,2	2
8	M 2> M 3	15,8	4

M=Member; A >B indicates that facial emotions of participant A is mimicked more by participant B compared to vice versa. A<B shows that facial emotions of B is mimicked more by participant A compared to vice versa.

from the CRQA results on surrogate data (n=24 dyads). The test results showed that actual CRQA results (M=16.5%; SD=0,615) were higher than the surrogate CRQA results (M=15.8%; SD=0,056) although the difference was not statistically significant ($t_{(45)}$=0,404, p=0,688). The insignificant result can be explained with the categorical nature of video-based facial emotions recognition results. The facial recognition algorithm applied in the current study classified facial emotional valence as positive, negative and neutral. The limited amount of possible emotional valence categories might have led to inflated CRQA results for the surrogate data. Moreover, as summarized, majority of the detected emotions were categorized as negative emotional valence. The dominance of negative emotional valence in the dataset further increases the likelihood of inflated CRQA results for the surrogate data. This claim is supported by previous research showing that if oscillators visit identical phases in time, it is more likely to accept null hypothesis when testing actual data against surrogate data [78].

C. Leaders and followers in the groups, based on video-coding

The leader and follower students in a group were identified from video coding results by calculating the duration they have led others during the task. For example, in Table III, we see that in Group 1, Member 1 (M1) led Member 2 (M2) 198 seconds during the whole task whereas M2 led M1 462 seconds.

Thus, we can assume that M2 was a leader and M1 was a follower during the collaboration. On the other hand, in Group 1, it can be also seen that M1 was leading M3 in a longer duration (184 s,) compared to vice versa (83 s,). Thus, although M1 was a follower of M2, he/she was a leader for M3 during the collaboration.

D. Validating leaders and followers identified by CRQA with video-coding results

The assumption was that if a group member leads others during the collaborative work, this would be reflected in emotional mimicry of the leading group member by the others

TABLE III
LEADING AND FOLLOWING DURATIONS AMONG THE GROUP MEMBERS

	M 1 > M 2	M2 > M1	M1 > M3	M3 >M1	M2>M3	M3>M2
Group 1	198	462	184	83	652	35
Group 2	187	540	205	4	864	5
Group 3	227	93	204	235	87	144
Group 4	121	119	160	282	60	314
Group 5	419	301	469	450	350	328
Group 6	32	574	36	896	155	378
Group 7	44	95	41	133	63	183
Group 8	281	166	260	398	126	396

M= Member; A >B this means that Member A leads Member B during the collaborative task.

as well. In order to check this assumption, we matched the emotional mimicry results (Table II) with the video coding of student interactions (Table III) and explored whether the leader and follower students are same in both analyses. Table IV displays the leader and followers in terms of CRQA results and video coding of interactions among the group members.

Table IV shows that leader and follower students identified by CRQA of emotional mimicry and video-coding of observed interactions were same in both analyses in 70% (n=17/24) of the cases. However, emotional mimicry and video-coding results indicated different leader and follower students in seven cases. These findings partially support our hypothesis that facial emotions of the leading students in a collaborative task are mimicked by the following students in the same group, although not always.

VI. DISCUSSION

This study shows that facial emotions mimicry among the group members can be an indicator of leader and follower students in collaborative learning settings. More specifically, our findings revealed that if a student leads the collaborative task, his/her facial emotions are mimicked more by the followers in the same group. However, the leading students mimic the facial emotions of the following students to a lesser extent. These findings are in line with the past research that found significant associations between emotional mimicry and status relationships among the interacting individuals. It has been found that in social settings individuals with less dominance have a tendency to attend the emotions of the individuals with more dominance [39], [40]. In collaborative learning, leading the group members with verbal utterances or physical actions can be regarded as a sign of power and

dominance [42], [43]. Based on this, we can assume that the more a student takes a leading position when proceeding with the collaborative task, the more his/her facial emotions are mimicked by the others. Although our assumption was supported in majority of the dyadic interactions among the group members, there were several exceptions as well.

Emotional mimicry does not occur solely in relation to the events that are important for a group to complete the task [80]. Thus, several factors about the social structure of the group might explain the cases in the current study in which emotional mimicry and video coding results indicated inconsistent leader-follower students. For example, emotional mimicry occurs less in antagonistic and competitive situations [35]. Similarly, emotional mimicry decreases when people take a distance from each other or develop negative attitudes towards the others [81] On the other hand, it was found that people mimick others' emotions when they are in close relationship or wish to affiliate with them [22], [24]. In this study, students worked together with their group members for multiple sessions during a school term. Thus, we can assume that they established specific communication norms and feelings about each other prior to the collaborative exam. Considering this, it is possible that among some group members, emotional mimicry was affected by the negative attitudes or enmity between the collaborators. For example, in Group 6, video coding results showed that Member 2 was leading Member 1 for a much longer duration of time compared to vice versa. Contrary to this, CRQA results showed that Member 2 was mimicking Member 1's facial emotions more. However, when we look at the RR values, we see that the emotional mimicry between Member 1 and Member 2 in Group 6 is quite low compared to the other dyads in the whole sample. Thus, it is worth questioning whether there were negative feelings between these two members which led to low emotional mimicry that was also inconsistent with the video-coding results. Unfortunately, we did not measure the feelings of students towards other members before or after the collaboration. Thus, it is not possible for us to make assertive remarks about how the attitudes among the group members in the current study affected the emotional mimicry. It is worth addressing this issue in future studies.

The current study reflects a research agenda to apply objective and unobtrusive measures of emotions in collaborative learning settings and utilize emotional mimicry to explore how emotions are disseminated among the collaborators. Emotional contagion in collaborative learning involves alignment of group emotions at a specific valence and arousal level depending on the social context and task progress [54]. Emotional contagion at the group level occurs through a group member initiating a specific emotional expression and others following it. In other words, emotion contagion at the group level emerges as group members mimick each other's emotions until the group reaches to a stabile shared emotional state. Therefore, focusing on the moments in which emotional mimicry among the collaborating students starts, continues or ends might provide valuable qualitative insights about the state of collaboration at a specific time, which might add to understanding complexity in collaborative teams [82]. Nevertheless, using emotional mimicry to detect dynamics in collaborative processes towards success or failure in collaboration is yet to gain attention. Drawing on this, the current study can be considered as one of the first attempts to utilize emotional mimicry in understanding the collaborative learning processes. Future studies should take a step further and investigate whether and how emotional mimicry can provide a fine-grained and temporal understanding about emotional contagion processes and their relation to the task progress.

TABLE IV
DO CRQA RESULTS MATCH WITH VIDEO-CODING RESULTS?

Group	Members	Video-coding results		CRQA results		Match
		Leader	Follower	Leader	Follower	Yes / No
1	1&2	2	1	2	1	Yes
1	1&3	1	3	1	3	Yes
1	2&3	2	3	3	2	No
2	1&2	2	1	2	1	Yes
2	1&3	1	3	1	3	Yes
2	2&3	2	3	2	3	Yes
3	1&2	1	2	1	2	Yes
3	1&3	3	1	1	3	No
3	2&3	3	2	2	3	No
4	1&2	1	2	1	2	Yes
4	1&3	3	1	3	1	Yes
4	2&3	3	2	3	2	Yes
5	1&2	1	2	1	2	Yes
5	1&3	1	3	1	3	Yes
5	2&3	2	3	2	3	Yes
6	1&2	2	1	1	2	No
6	1&3	3	1	3	1	Yes
6	2&3	3	2	3	2	Yes
7	1&2	2	1	1	2	No
7	1&3	3	1	3	1	Yes
7	2&3	3	2	3	2	Yes
8	1&2	1	2	1	2	Yes
8	1&3	3	1	1	3	No
8	2&3	3	2	2	3	No

VII. LIMITATIONS AND FUTURE STUDIES

The first limitation of the study is high amount of missing values in terms of facial emotions. Therefore, the current findings are limited in terms of generalizability. The missing data might be due to recording facial emotions of the participants in each group with a single video-camera. A better approach in future studies would be to use several video-cameras for each group to record their naturalistic collaboration from different point of views. Multiple video recordings of the group members would be synchronized on the same time scale. Then, it might be possible to compensate a missing data for a participant at a specific time from multiple video recordings of the same group.

Another limitation of the study is that our facial emotions analysis algorithm currently categorizes the valence dimension of the facial emotions as positive, negative and neutral to

provide an overall view about the transitory nature of emotions during collaboration. However, a variety of emotions are experienced both at negative (e.g. boredom, anxiety, sadness, shame and guilt) and positive valences (e.g. enjoyment, pride and satisfaction) [51]. In addition, emotions differ from each other in terms of arousal. Future studies on emotional mimicry might employ more advanced methods to classify emotions with respect to their valence and arousal at a continuous scale rather than categorical.

The current study was conducted as part of a high school course and collaborative exams were graded by the subject teachers on a 6-point scale. The results revealed that all group performance scores varied between 5,5 and 6 points. Considering the limited variation among the group performance scores, it was not possible to investigate the relationship between emotional mimicry and collaborative performance. Future studies might utilize a more detailed metric with a larger scale to grade the collaborative task and, investigate whether and how emotional mimicry influences group performance.

Finally, sharing compatible goals have been also found to be a contributor of emotional mimicry [24]. Developing a shared understanding and deciding on shared goals are an essential characteristic of collaborative learning. Therefore, future studies can address how developing shared goals or not affects emotional mimicry during collaboration.

VIII. CONCLUSION

The study employs a video-based facial expression recognition method to capture the momentary changes of students' emotions during a collaborative exam task, and demonstrates a novel methodology, CRQA to detect the emotional contagion among the collaborating partners. Our findings show that emotional mimicry can reveal about the leader and follower relationships during collaborative learning to certain extent. However, the relationship between emotional mimicry and leader-follower interactions during collaboration can be affected by several confounding factors such as attitude of group members towards collaboration or affiliation among the group members. Therefore, the current study calls for future research to shed more light on the role of emotional mimicry in collaborative learning. We conclude that video-based facial expression method is a new promising approach for understanding dynamics in collaborative learning, which can contribute for future implications mitigating socioemotional challenges in collaborative learning.

REFERENCES

[1] N. Miyake and P. A. Kirschner, "The social and interactive dimensions of collaborative learning," in *The Cambridge Handbook Of The Learning Sciences*, R. K. Sawyer, Ed. NY: Cambridge University Press, 2014, pp. 418–438.

[2] P. Näykki, S. Järvelä, P. A. Kirschner, and H. Järvenoja, "Socio-emotional conflict in collaborative learning—A process-oriented case study in a higher education context," *Process. case study a High. Educ. Context*, vol. 68, pp. 1–14, 2014.

[3] O. B. Ayoko, V. C. Callan, and C. E. J.- Härtel, "Climate on Conflict and Team Members' Reactions to Conflict," *Small Gr. Res.*, vol. 39, no. 2, pp. 121–149, 2008.

[4] P. Ekman, "Facial Expressions of Emotion: New Findings, New Questions," *Psychol. Sci.*, vol. 3, no. 1, pp. 34–38, 1992.

[5] N. Bosch *et al.*, "Detecting student emotions in computer-enabled classrooms," in *IJCAI International Joint Conference on Artificial Intelligence*, 2016, pp. 4125–4129.

[6] M. Tomasello, *A natural history of human thinking*. Cambridge: Harvard University Press, 2014.

[7] R. S. Lazarus, *Emotion and adaptation*. New York, NY: Oxford University Press, 1991.

[8] A. Bakhtiar, E. A. Webster, and A. F. Hadwin, "Regulation and socio-emotional interactions in a positive and a negative group climate," *Metacognition Learn.*, vol. 13, no. 1, pp. 57–90, 2018.

[9] N. M. Webb, M. Ing, N. Kersting, and K. M. Nemer, "Help seeking in cooperative learning groups," in *Strategic help seeking: Implications for learning and teaching*, S. A. Karabenick and R. S. Newman, Eds. NY: Routledge, 2006, pp. 45–115.

[10] L. Linnenbrink-Garcia, T. K. Rogat, and K. L. Koskey, "Affect and engagement during small group instruction," *Contemp. Educ. Psychol.*, vol. 36, no. 1, pp. 13–24, 2011.

[11] S. M. Fiore, M. A. Rosen, K. A. Smith-Jentsch, E. Salas, M. Letsky, and N. Warner, "Toward an understanding of macrocognition in teams: Predicting processes in complex collaborative contexts," *Hum. Factors*, vol. 52, no. 2, pp. 203–224, 2010.

[12] J. N. Ervin, J. A. Bonito, and J. Keyton, "Convergence of intrapersonal and interpersonal processes across group meetings," *Commun. Monogr.*, vol. 84, no. 2, pp. 200–220, 2017.

[13] M. Cukurova, R. Luckin, E. Millán, and M. Mavrikis, "The NISPI framework: Analysing collaborative problem-solving from students' physical interactions," *Comput. Educ.*, vol. 116, pp. 93–109, 2018.

[14] S. L. Do and D. L. Schallert, "Emotions and Classroom Talk: Toward a Model of the Role of Affect in Students' Experiences of Classroom Discussions," *J. Educ. Psychol.*, vol. 96, no. 4, p. 619, 2004.

[15] D. Keltner and J. J. Gross, "Functional accounts of emotions," *Cogn. Emot.*, vol. 13, no. 5, pp. 467–480, 1999.

[16] N. H. Frijda, "The laws of emotion," *Am. Psychol.*, vol. 43, no. 5, p. 349, 1988.

[17] L. F. Barrett, "Discrete emotions or dimensions? The Role of valence focus and arousal focus," *Cogn.*

Emot., vol. 12, no. 4, pp. 579–599, 1998.

[18] H. Stadthagen-Gonzalez, C. Imbault, M. A. P. Sánchez, and M. Brysbaert, "Norms of valence and arousal for 14,031 Spanish words," *Behav. Res. Methods*, vol. 49, no. 1, pp. 111–123, 2017.

[19] C. von Scheve and S. Ismer,"Towards a theory of collective emotions," *Emot. Rev.*, vol. 5, no. 4, pp. 406–413, 2013.

[20] K. Peters and Y. Kashima, "A multimodal theory of affect diffusion," *Psychol. Bull.*, vol. 141, no. 5, pp. 966–992, 2015.

[21] B. Parkinson, "Interpersonal emotion transfer: Contagion and social appraisal," *Soc. Personal. Psychol. Compass*, vol. 5, pp. 428–439, 2011.

[22] J. L. Lakin and T. L. Chartrand, "Using nonconscious behavioral mimicry to create affiliation and rapport," *Psychol. Sci.*, vol. 14, no. 4, pp. 334–339, 2003.

[23] T. K. Rogat and K. R. Adams-Wiggins, "Interrelation between regulatory and socioemotional processes within collaborative groups characterized by facilitative and directive other-regulation," *Comput. Human Behav.*, vol. 52, pp. 589–600, 2015.

[24] U. Hess and A. Fischer, "Emotional Mimicry as Social Regulation," *Personal. Soc. Psychol. Rev.*, vol. 17, no. 2, pp. 142–157, 2013.

[25] S. T. Hawk, A. H. Fischer, and G. A. Van Kleef, "Face the noise: Embodied responses to nonverbal vocalizations of discrete emotions," *J. Pers. Soc. Psychol.*, vol. 102, no. 4, pp. 796–814, 2012.

[26] M. Paetzel, G. Varni, I. Hupont, M. Chetouani, C. Peters, and G. Castellano, "Investigating the Influence of Embodiment on Facial Mimicry in HRI Using Computer Vision-Based Measures."

[27] M. Paetzel, G. Varni, I. Hupont, M. Chetouani, C. Peters, and G. Castellano, "The Attribution of Emotional State - How Embodiment Features and Social Traits Affect the Perception of an Artificial Agent."

[28] M. Chetouani, G. Varni, and I. Hupont, "Computational Study of Primitive Emotional Contagion in Dyadic Interactions," vol. 3045, no. c, 2017.

[29] E. Hatfield, L. Bensman, P. D. Thornton, and R. L. Rapson, "New Perspectives on Emotional Contagion: A Review of Classic and Recent Research on Facial Mimicry and Contagion," *Interpersona An Int. J. Pers. Relationships*, vol. 8, no. 2, pp. 159–179, 2014.

[30] E. Hatfield, M. Forbes, and R. L. Rapson, "Collective Emotions: Perspectives from Psychology, Philosophy, and Sociology," *Collect. Emot. Perspect. from Psychol. Philos. Sociol.*, p. 447, 2014.

[31] U. Hess and A. Fischer, "Emotional mimicry as social regulation," *Personal. Soc. Psychol. Rev.*, vol. 17, no. 2, pp. 142–157, 2013.

[32] M. Sonnby-Borgström, P. Jönsson, and O. Svensson, "Gender differences in facial imitation and verbally reported emotional contagion from spontaneous to emotionally regulated processing levels," *Scand. J. Psychol.*, vol. 49, no. 2, pp. 111–122, 2008.

[33] V. Der Schalk *et al.*, "Convergent and divergent responses to emotional displays of ingroup and outgroup," *Emotion*, vol. 11, no. 2, p. 286, 2011.

[34] M. Weisbuch and N. Ambady, "Affective divergence: automatic responses to others' emotions depend on group membership," *J. Pers. Soc. Psychol.*, vol. 95, no. 5, p. 1063, 2008.

[35] P. Weyers, A. Mühlberger, A. Kund, U. Hess, and P. Pauli, "Modulation of facial reactions to avatar emotional faces by nonconscious competition priming," *Psychophysiology*, vol. 46, no. 2, pp. 328–335, 2009.

[36] E. R. Smith, C. R. Seger, and D. M. Mackie, "Can emotions be truly group level? Evidence regarding four conceptual criteria," *J. Pers. Soc. Psychol.*, vol. 93, no. 3, p. 431, 2007.

[37] R. Porat, E. Halperin, and M. Tamir, "social regulator," *J. Pers. Soc. Psychol.*, vol. 110, no. 2, p. 167, 2016.

[38] L. Z. Tiedens and A. R. Fragale, "Power Moves : Complementarity in Dominant and Submissive Nonverbal Behavior," *J. Pers. Soc. Psychol.*, vol. 84, no. 3, pp. 558–568, 2003.

[39] E. Carr, P. Winkielman, and C. Oveis, "UC San Diego," *J. Exp. Psychol. Gen.*, vol. 143, no. 3, p. 997, 2014.

[40] S. Connelly, B. Gaddis, and W. Helton-Fauth, "A Closer Look at the Role of Emotions in Transformational and Charismatic Leadership," in *Monographs in Management and Leadership*, F. J. Y. Bruce J. Avolio, Ed. Emerald Group Publishing Limited, 2013, pp. 299–327.

[41] R. Yamaguchi and M. L. Maehr, "Children?s emergent leadership: the relationships with group characteristics and outcomes," *Small Gr. Res.*, vol. 35, no. 4, pp. 388–406, 2004.

[42] P. E. Slater, "Role differentiation in small groups," *Am. Sociol. Rev.*, vol. 20, no. 3, pp. 300–310, 1955.

[43] R. J. Bruno, "Social differentiation, participation inequality and optimal collaborative learning online," Purdue University, 2010.

[44] N. Webb, K. Nemer, A. Chiznik, and B. Sugrue, "Equity Issues in Collaborative Group Assessment : Group Composition and Performance," 1997.

[45] C. Shirky, *Here Comes Everybody*. New York: The Penguin Press. Shirky, 2008.

[46] V. M. Eguiluz, M. G. Zimmermann, C. J. Cela-Conde, and M. S. Miguel, "Cooperation and the Emergence of Ecosystem Institutions," *Am. J. Sociol.*, vol. 110, no. 4, pp. 997–1008, 2005.

[47] A. B. Bakker, "Flow among music teachers and their students: The crossover of peak experiences," *J. Vocat. Behav.*, vol. 66, no. 1, pp. 26–44, 2005.

[48] A. C. Frenzel, T. Goetz, O. Ludtke, R. Pekrun, and R. E. Sutton, "Emotional transmission in the classroom: exploring the relationship between teacher and student enjoyment," *J. Educ. Psychol.*, vol. 101, no. 3, p. 705, 2009.

[49] E. S. Becker, T. Goetz, V. Morger, and J. Ranellucci, "The importance of teachers' emotions and instructional behavior for their students' emotions -

An experience sampling analysis," *Teach. Teach. Educ.*, vol. 43, pp. 15–26, 2014.

[50] W. N. Morris, *Mood: The frame of mind*. Berlin Heidelberg New York: Springer, 1989.

[51] R. Pekrun, T. Goetz, W. Titz, and R. P. Perry, "Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research," *Educ. Psychol.*, vol. 37, no. 2, pp. 91–105, 2002.

[52] P. A. Schutz, J. Y. Hong, D. I. Cross, and J. N. Osbon, "Reflections on investigating emotion in educational activity settings," *Educ. Psychol. Rev.*, vol. 18, no. 4, pp. 343–360, 2006.

[53] H. Järvenoja, S. Järvelä, and J. Malmberg, "Supporting groups' emotion and motivation regulation during collaborative learning," *Learn. Instr.*, 2017.

[54] H. Järvenoja *et al.*, "Capturing motivation and emotion regulation during a learning process," *Front. Learn. Res.*, vol. 6, p. 3, 2018.

[55] M. G. Frank, M. Herbasz, K. Sinuk, A. Keller, and C. Nolan, *I see how you feel: Training laypeople and professionals to recognize fleeting emotions*. Sheraton, New York City: International Communication Association, 2009.

[56] M. H. Chiu, H. L. Liaw, Y. R. Yu, and C. C. Chou, "Facial micro?expression states as an indicator for conceptual change in students' understanding of air pressure and boiling points," *Br. J. Educ. Technol.*, vol. 50, no. 1, pp. 469–480, 2019.

[57] H. L. Liaw, M. H. Chiu, and C. C. Chou, "Using facial recognition technology in the exploration of student responses to conceptual conflict phenomenon," *Chem. Educ. Res. Pract.*, vol. 15, no. 4, pp. 824–834, 2014.

[58] K. Bahreini, R. Nadolski, and W. Westera, "Communication skills training exploiting multimodal emotion recognition," *Interact. Learn. Environ.*, vol. 25, no. 8, pp. 1065–1082, 2017.

[59] J. Whitehill, Z. Serpell, Y. C. Lin, A. Foster, and J. R. Movellan, "The faces of engagement: Automatic recognition of student engagementfrom facial expressions," *IEEE Trans. Affect. Comput.*, vol. 5, no. 1, pp. 86–98, 2014.

[60] A. Psaltis, K. C. Apostolakis, K. Dimitropoulos, and P. Daras, "Multimodal student engagement recognition in prosocial games," *IEEE Trans. Games*, vol. 10, no. 3, pp. 292–303, 2018.

[61] H. Farmer, E. W. Carr, M. Svartdal, P. Winkielman, and A. F. D. C. Hamilton, "Status and power do not modulate automatic imitation of intransitive hand movements," *PLoS One*, vol. 11, p. 4, 2016.

[62] C. Anderson, S. E. Spataro, and F. J. Flynn, "Personality and organizational culture as determinants of influence," *J. Appl. Psychol.*, vol. 93, no. 3, p. 702, 2008.

[63] H. Järvenoja and S. Järvelä, "Emotion control in collaborative learning situations: Do students regulate emotions evoked by social challenges," *Br. J. Educ. Psychol.*, vol. 79, no. 3, pp. 463–481, 2009.

[64] M. I. Coco, R. Dale, and N. Nguyen, "Cross-

[65] recurrence quantification analysis of categorical and continuous time series : an R package," vol. 5, no. June, pp. 1–14, 2014.

A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic, "Robust discriminative response map fitting with constrained local models," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2013, pp. 3444–3451.

[66] C. A. Corneanu, M. O. Simón, J. F. Cohn, and S. E. Guerrero, "Survey on RGB, 3D, Thermal, and Multimodal Approaches for Facial Expression Recognition: History, Trends, and Affect-Related Applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 8, pp. 1548–1568, 2016.

[67] F. Zhang, T. Zhang, Q. Mao, and C. Xu, "Joint pose and expression modeling for facial expression recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 3359–3368.

[68] J. Chen, Z. Chen, Z. Chi, and H. Fu, "Facial expression recognition in video with multiple feature fusion," *IEEE Trans. Affect. Comput.*, vol. 9, no. 1, pp. 38–50, 2018.

[69] X. Huang, S. Wang, X. Liu, G. Zhao, X. Feng, and M. Pietikainen, "Spontaneous Facial Micro-Expression Recognition using Discriminative Spatiotemporal Local Binary Pattern with an Improved Integral Projection," vol. 10, no. 1, pp. 32–47, 2019.

[70] X. Huang, A. Dhall, G. Zhao, R. Goecke, and M. Pietikäinen, "Riesz-based Volume Local Binary Pattern and A Novel Group Expression Model for Group Happiness Intensity Analysis," in *In BMVC*, 2015, pp. 31–34.

[71] A. Dhall, J. Joshi, K. Sikka, R. Goecke, and N. Sebe, "The more the merrier: Analysing the affect of a group of people in images," *2015 11th IEEE Int. Conf. Work. Autom. Face Gesture Recognition, FG 2015*, vol. 1, pp. 1–8, 2015.

[72] X. Zhu and D. Ramanan, "Face detection, pose estimation, and landmark localization in the wild BT - Proc. of. Conf. on Computer Vision and Pattern Recognition," *2012 IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012.

[73] M. I. Coco and R. Dale, "Performance in a Collaborative Search Task : The Role of Feedback and Alignment," *Top. Cogn. Sci.*, vol. 10, pp. 55–79, 2018.

[74] E. B. M. A. Riley and G. van orden, *Tutorials in contemporary nonlinear methods*. National Science Foundation, 2005.

[75] D. G. Stephen, J. A. Dixon, and R. W. Isenhower, "Dynamics of representational change: Entropy, action, and cognition," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 35, no. 6, p. 1811, 2009.

[76] K. J. Rohlfing, G. Leonardi, I. Nomikou, J. Ra?czaszek-Leonardi, and E. Hüllermeier, "Multimodal turn-taking: motivations, methodological challenges, and novel approaches," *IEEE Trans. Cogn. Dev. Syst.*, pp. 1–12, 2019.

[77] U. Dimberg and M. Thunberg, "Rapid facial reactions to emotional facial expressions," *Scand. J. Psychol.*,

vol. 39, no. 1, pp. 39–45, 1998.

[78] G. Lancaster, D. Iatsenko, A. Pidde, V. Ticcinelli, and A. Stefanovska, "Surrogate data for hypothesis testing of physical systems," *Phys. Rep.*, vol. 748, pp. 1–60, 2018.

[79] M. M. Louwerse, R. Dale, E. G. Bard, and P. Jeuniaux, "Behavior matching in multimodal communication is synchronized," *Cogn. Sci.*, vol. 36, no. 8, pp. 1404–1426, 2012.

[80] H. M. Weiss and R. Cropanzano, "Affective events theory: a theoretical discussion of the structure, causes and consequences of affective experience at work," in *Research in organizational behavior*, B. M. Staw and L. L. Cummings, Eds. Greenwich, CT: JAI Press, 1996.

[81] U. Hess, S. Houde, and A. Fischer, "Do we mimic what we see or what we know," in *Collective Emotions,*, C. V Scheve and M. Salmela, Eds. UK: Oxford, 2014, pp. 94–107.

[82] S. Decuyper, F. Dochy, and P. den Bossche, "Grasping the Dynamic Complexity of Team Learning: An Integrative Model for Effective Team Learning in Organisations," *Educ. Res. Rev.*, vol. 5, no. 2, pp. 111–113, 2010.

Muhterem Dindar received his PhD degree in Computer Education and Instructional Technology from Anadolu University, in 2015. He is currently working as a postdoctoral researcher at the Learning and Educational Technology (LET) Research Unit at the University of Oulu. His research interests include complex problem solving, multimodal collaboration analytics, and gamified learning.

Sanna Järvelä is a professor in learning sciences and educational technology and a head of the Learning and Educational Technology Research Unit (LET) in the University of Oulu. Her main research interests deal with self-regulated learning, computer supported collaborative learning and on-line learning processes. Järvelä and her research group is internationally recognized in theoretical and methodological advancement of social aspects of self-regulated learning and processes oriented and multimodal methods. Järvelä is the member of the Finnish Academy of Science and Letters. Järvelä has published more than 150 scientific papers in international refereed journals and about 50 book chapters and three edited books, she has 8608 citations in Google Scholar and her h-index is 48. She is the current EARLI (European Association for Research on Learning and Instruction) president.

Sara Ahola received the B.S. degree in educational psychology from University of Oulu, Finland, in 2018. She is currently working toward the M.S. degree, with her master's thesis considering self-efficacy and group level regulation of learning. She is a research assistant in the Learning and Educational Technology research unit (LET) in the Faculty of Education at University of Oulu, Finland.

Xiaohua Huang received the B.S. degree in communication engineering from Huaqiao University, Quanzhou, China in 2006. He received his Ph.D degree in Computer Science and Engineering from the University of Oulu, Oulu, Finland in 2014. He was a research assistant in Southeast University since 2006. He had been a scientist researcher in the Center for Machine Vision and Signal Analysis at the University of Oulu in 2015-2018. He is now an independent researcher at University of Oulu. He has authored or co-authored more than 20 papers in journals and conferences, and has served as a reviewer for journals and conferences. His current research interests include facial expression recognition, micro-expression analysis, group-level emotion recognition, multi-modal emotion recognition and texture classification.

Guoying Zhao (IEEE Senior member) is currently a Professor with the Center for Machine Vision and Signal Analysis, University of Oulu, Finland, where she has been a senior researcher since 2005 and an Associate Professor since 2014. She received the Ph.D. degree in computer science from the Chinese Academy of Sciences, Beijing, China, in 2005. She has authored or co-authored more than 190 papers in journals andconferences. Her papers have currently over 9600 citations in Google Scholar (h-index 43). She was co-publicity chair for FG2018, has served as area chairs for several conferences and is associate editor for Pattern Recognition, IEEE Transactions on Circuits and Systems for Video Technology, and Image and Vision Computing Journals. She has lectured tutorials at ICPR 2006, ICCV 2009, SCIA 2013 and FG 2018, authored/edited three books and eight special issues in journals. Dr. Zhao was a Co-Chair of many International Workshops at ECCV, ICCV, CVPR, ACCV and BMVC. Her current research interests include image and video descriptors, facial-expression and micro-expression recognition, gait analysis, dynamic-texture recognition, human motion analysis, and person identification. Her research has been reported by Finnish TV programs, newspapers and MIT Technology Review.