



# Facial Expression Recognition for Examining Emotional Regulation in Synchronous Online Collaborative Learning

Duong Ngo<sup>1</sup> · Andy Nguyen<sup>2</sup>  · Belle Dang<sup>2</sup> · Ha Ngo<sup>3</sup>

Accepted: 3 November 2023 / Published online: 2 January 2024  
© The Author(s) 2023

## Abstract

Artificial intelligence (AI) has been recognised as a promising technology for methodological progress and theoretical advancement in learning sciences. However, there remains few empirical investigations into how AI could be applied in learning sciences research. This study aims to utilize AI facial recognition to inform the learning regulation behaviors in synchronous online collaborative learning environments. By studying groups of university students ( $N=36$ ) who participated in their online classes under the COVID-19 social distancing mandates, we strive to understand the interrelation between individual affective states and their collaborative group members. Theoretically underpinned by the socially shared regulation of learning framework, our research features a cutting-edge insight into how learners socially shared regulation in group-based tasks. Findings accentuate fundamental added values of AI application in education, whilst indicating further interesting patterns about student self-regulation in the collaborative learning environment. Implications drawn from the study hold strong potential to provide theoretical and practical contributions to the exploration of AI supportive roles in designing and personalizing learning needs, as well as fathom the motion and multiplicity of collaborative learning modes in higher education.

**Keywords** Facial expression recognition · Emotional regulation · Socially shared regulation · Synchronous online learning · Computer-supported collaborative learning

---

✉ Andy Nguyen  
andy.nguyen@oulu.fi

<sup>1</sup> Hanoi University of Science and Technology, Hanoi, Vietnam

<sup>2</sup> Learning & Educational Technology (LET) Research Lab, University of Oulu, Oulu, Finland

<sup>3</sup> Faculty of Education, Victoria University of Wellington, Wellington, New Zealand

## Introduction

The COVID-19 pandemic has had a significant impact on higher education, prompting rapid shifts towards online instruction worldwide over the past few years. This transformation necessitates the adoption of innovative pedagogical approaches that can accommodate students' individualization and diverse motivations. Given the emerging context of remote teaching and learning, collaborative learning has become a popular choice to promote engagement among learners. Self-regulation is a critical component of individual learning success whereas socially shared regulation of learning (SSRL) is the key factor contributing to group performance in collaborative learning (Järvelä et al., 2019). The sudden changes to educational practices caused by the pandemic have further emphasized the need to gain a deeper understanding of and support socially shared regulation of regulation, which has become increasingly important in the new learning environment. However, it has been challenging to assess and support students' socially shared regulation in collaborative learning due to the unobservability of the cognitive and emotional processes and the dynamics of collaborative interactions (Järvelä & Bannert, 2021). Previously published studies are mostly confined to self-reported surveys and interviews. Despite recent attempts in exploring the use of multimodal data including physiological data (e.g., Dindar et al., 2022), video and audio data (e.g., Isohätälä et al., 2018) to study SSRL, one of the most significant discussions in SSRL research is the need for innovative methods to analyze new data modalities (Azevedo & Gašević, 2019). Recent advances in Artificial Intelligence (AI) in general, and computer vision facial recognition in specific, along with its applications in learning sciences have enabled novel methods for examining self-regulation and offered new insights into the self-regulated learning process (Nguyen et al., 2021, 2022b). To date, there is a dearth of empirical evidence that identifies how AI methods can be leveraged to investigate students' socially shared regulation in online collaborative learning. Addressing this research gap, our study aims to employ AI facial expression recognition to explore students' emotional regulation in synchronous computer-supported collaborative learning (CSCL).

Emotional regulation is one of the most challenging aspects of socially shared regulation, given the diverse involvement of learners at the group level and the individualized, sometimes idiosyncratic co- and socially shared emotion regulation strategies that are pertinent to each challenge (Järvenoja et al., 2019). However, prior studies that focus on investigating the factors associated with SSRL have more frequently employed traditional data collection instruments such as surveys, interviews, or a combination of both in self-reports (Järvelä et al., 2019; Kwon et al., 2014). Järvelä et al. (2019) suggested that regulation is “not linear and involves cyclical adaptation, which “requires multiple data channels” to feature not only individual, social learning activities but also their interactions with the learning context (p. 429). Given the multiplicity of results obtained, the authors also share this viewpoint and would argue that little advancement following this route can be achieved due to the unavoidable subjectivity of participants embedded in the qualitative data amassed. The limited range of methodologies

used in previous studies highlights the need for novel approaches that utilize emerging technologies in data collection, which can provide objective insights into the SSRL process.

Furthermore, in the light of recent turbulence caused by Covid-19 pandemic over the last few years, it is undeniable that the dynamics of teaching and learning have evolved dramatically, especially in the way education is delivered. Technology-featured and technology-aided models such as hybrid learning and blended learning have become the norm, leading to an increase in collaborative learning in virtual environments. Despite this transformation, research on the social interactions of learners during the collaborative learning process in a computer-supported environment is limited to traditional questionnaires (Panadero et al., 2016). While previous research has explored SSRL in asynchronous CSCL, there is a gap in knowledge regarding learning regulation in synchronous CSCL. Prior to the COVID-19 pandemic, most research on regulation in CSCL focused on asynchronous learning settings, such as those found on learning management systems (LMS) or Massive Open Online Courses (MOOCs). However, with the increased need for synchronous CSCL on platforms such as Microsoft Teams and Zoom, there is a growing demand for understanding SSRL in these contexts. Despite the importance of SSRL, especially during the COVID-19 pandemic, a literature review reveals a lack of evidence on SSRL synchronous CSCL, and limited attention has been given to emotional regulation in this context. Consequently, this study aims to address this gap and provide new insights into SSRL in synchronous CSCL.

In this study, we aim to explore the ways that thirty-six students in a higher education setting collaboratively study English and complete their assigned speaking tasks through a virtual learning platform (Microsoft Teams) and use digital data (recordings of the sessions) to track their socio-emotional interactions that led to self-regulation in learning. Our contribution is reflected in two aspects. Methodologically, it updates the ongoing literature with an AI-based real-time checking emotion and data collection tool towards an emerging field that has fast becoming a key instrument in understanding SSRL and, more broadly, high-level of cognitive learning. Upon the attainment of this goal, which is tracking learners' emotion during their learning process, implications at a theoretical ground regarding socially shared learning can be further investigated since emotional process is instrumental to successful learning. The dynamic and cyclical nature of this process has long served as a great hindrance to traditional methods such as interviews or surveys to capture accurately and fully. The utilization of AI tools in exploring qualitative processes of learning offers significant added value in interpreting unobservable processes such as learning will also pave ways to applicability in other interdisciplinary studies of related fields. Our study, which focuses on socially regulated learning in an online, computer assisted environment targeting collaborative work within learning groups, is also anticipated to provide practical ramifications to an evolving context of education that encapsulates a growing demand in online learning during and post-Covid era. Collectively, positioning as one of the first studies to initiate the use of AI to conduct research directly on Teams and via the platform of video-based qualification enables us to examine theories that are pertinent to regulation at group level in a context of synchronous collaborative learning.

## Theoretical Background

### Methodological Progress and Challenges in Studying SSRL

Socially shared regulation of learning (SSRL) is defined as a deliberate and strategic approach to planning, task enactment, reflection, and adaptation that takes place within a group context (Järvelä et al., 2019). Over the past ten years, there has been an increase in interest in understanding SSRL. Due to the lack of understanding regarding the mechanisms involved in the SSRL model within collaborative learning, recent efforts have been made to elucidate the regulation processes of cognition, motivation, and emotion in SSRL (Nguyen et al., 2023; Winne, 2019). It is recently suggested that the exploration and support of CoRL and SSRL have been hampered by the dynamic character of collaborative learning and the concealed processes of affective and cognitive alterations at the core of regulation. (Järvelä et al., 2019).

To elaborate on the issue, The previous literature has highlighted the notion that students utilize metacognition consistently to adjust their learning tactics, and these adaptations in the learning process may vary from cycle to cycle (Zimmerman & Schunk, 2011). Then, in order to accurately capture SSRL, it is necessary to comprehend and record each learner's varied and interconnected elements (such as emotion, motivation, and cognition) as well as their interactions and regulation with others within the social learning context (Järvelä et al., 2019). Accordingly, there was significant interest in finding new data collection methods in studying SSRL, with more emphasis on trace data or real-time measurements, and one such approach was multimodal data analytics. This includes analyzing eye-tracking data, screen recordings, and time-stamped descriptions of observed interactions between students and content using physiological sensors, think-aloud protocols, and interactions between students and machines (e.g., Taub & Azevedo, 2016).

Despite the benefits of using multimodal data analysis to investigate SSRL, there are limitations to this approach. Firstly, collecting multimodal data is costly, and ensuring that it captures students' real conceptions and intentions during the learning process is challenging (Fan et al., 2021). Secondly, because the results obtained from multimodal data analysis are fragmented by nature, it is difficult to measure and infer the interrelated aspects of the SSRL process. Thus, it is recommended that more comprehensive research techniques be employed to better examine SSRL in collaborative learning (Järvelä et al., 2019). Therefore, this study seeks to explore a new approach utilizing AI techniques to examine SSRL in synchronous CSCL.

### Artificial Intelligence (AI) Approaches for Examining SSRL Through Facial Emotion Recognition

Recent innovations in AI research, specifically in deep learning, have allowed AI algorithms to transform analytics-relevant tasks with fairly accurate results and minimal human effort. Notably, AI has made significant advances in both cognitive process analysis and real-time emotion recognition. In cognitive analysis, recent

works by Gao et al. (2021) and Debie et al. (2019) demonstrate how neural network pre-trained models can be used to predict and analyze cognitive processes using inputs from invasive and non-invasive cognitive monitoring. This provides SSRL researchers with various opportunities to better analyze and understand cognitive processes more holistically. Meanwhile, deep learning CNN models have significantly improved the performance of emotion recognition from video inputs, which has been a well-studied problem for decades. Those interested in the history and state-of-the-art AI approaches to facial emotion recognition can refer to the recent survey by Canal et al. (2022).

With the help of AI, researchers can better understand how students are regulating their learning in real-time and in different contexts, providing insights that were not possible before. Additionally, AI can analyze large datasets and generate accurate predictions, providing a more comprehensive understanding of the complex SSRL processes. When being used as complementary to multimodal analysis (Zhang et al., 2020), AI algorithms offer the capability not only to increase the accuracy of the frame-by-frame analysis of emotions but, but also offer more in-depth analysis of physiological sensors and audio data (Nguyen et al., 2023). Using state-of-the-art development in deep learning AI facial recognition algorithms, this study aims to better capture and understand student emotions involved in SSRL.

## Methods

### Participants and Procedures

Participants of the study were 36 college students from a higher education institution in Vietnam (19 females, 17 males), with a mixture of freshmen and sophomores. The study was carried out in an undergraduate English-Speaking Skills course that is part of an advanced program. This program is designed for non-native English speakers who study in small classes led by experienced lecturers, with intensive content in foreign languages. Additionally, the program provides students with opportunities to participate in research and interdisciplinary team projects from early on in their studies, and to practice solving practical problems faced by enterprises as part of their learning process. The advanced program's content is based on modern curricula from universities in developed countries such as the United States, Germany, Japan, and the United Kingdom. The program's teaching and learning materials are primarily delivered in English, with some special programs designed to be taught in Vietnamese following an English learning enhancement scheme. It is worth noting that the class was initially intended to be conducted in-person, but due to the complexity of the Covid-19 situation, with constant lockdowns in Vietnam during the study period, teaching and learning had to be shifted to online environments.

The data used in the current study includes recording from the small group discussions in a collaborative learning task in which students were arranged into 14 groups, each ranging between two to three members working together on an assigned topic by the instructor. In this study, we designed a collaborative learning task for ESL learners where a group of students are asked to discuss a topic at own

choice in English, while individuals within the group are encouraged to correct each other's phrasing and pronunciation. An example topic discussed by the students was "Impacts of Covid 19 on our life" or "How to study well at university?". The collaborative learning tasks for ESL was focusing on the discussion in English as the second language rather than the actual discussion content. Prior to the collaborative learning task, each group was given a topic by the instructor for preparation. In the collaborative learning session, students were asked to turn on the camera and be visible throughout the task, but no specific instructions regarding head or body location were announced to the participants. Since the group project was conducted remotely, the quality of the recording varies, since it is dependent on the student's personal webcam, ranging between  $640 \times 320$  pixels to  $1920 \times 1080$  pixels. The maximum time allowed for the collaborated learning task was 15 min, and the mean participant time for each student was around 13 min, resulting in a dataset with 14 distinct videos.

## Data Analysis

### Facial Expression Recognition

Facial Expression Recognition (FER) performed computationally is an ongoing challenging research topic in the machine learning research community. Since the work of Krizhevsky et al. (2012), there has been a resurgence of interest in convolutional neural networks for image analysis. In recent years, convolutional-based deep FER models have been shown to consistently outperform SIFT algorithms when it comes to scalability and generalizability in FER tasks (Goodfellow et al., 2013). Moreover, recent state-of-the-art advancements in deep learning models for video-based facial expression recognition analysis allowed us to classify the participants' facial emotions on a wider spectrum on a continuous range, which allows more flexible and robust frameworks for multi-class emotion classification by modification of the SoftMax classification layer (Giannopoulos et al., 2018). On top of that, there have also been innovations in deep learning-based FER in the wild research, which provided the ability to analyze human emotions directly from video recordings, without the need for explicit lab-control environments. These recent developments have provided us with the technical tools to better understand and analyze student's emotion under sequenced phases of recursive cognition of the learning regulation process (Nguyen et al., 2022a, b).

Given various available FER algorithms available in the literature for evaluating how learners socially shared regulate in group-based tasks in online learning under the regulated learning framework (Canal et al., 2022; Mellouk & Handouzi, 2020), it is important to choose a suitable FER model for practical setting. In an online-learning collaborative group-based task, given the often-small scope of the dataset ( $< 100$  examples per class), it is unlikely to be able to train a domain specific model from scratch or even fine-tune existing FER models. This means that the most practical approach to FER is to rely on publicly available pre-trained models for detecting emotional expression through the self-regulated collaborative learning process. To this endeavor,

several approaches applying AI exist, divided into two main categories: one-shot detectors, which attempt to directly locate the student's face and detect the student emotion at once, and two-shot detectors, which attempted to tackle this problem through two-steps: face-localization and emotion detection (Canal et al., 2022). We decided to adopt the second approach, given results from a recent survey in FER (Dalvi et al., 2021) suggested that the later approach achieves better performance on out-of-sample data.

As a result, this study applies the AI architecture using *img2pose* (Albiero et al., 2020) pre-trained on the Wider face dataset (Yang et al., 2016) as the facial localization module and Residual Masking Network (Pham et al., 2021) as our emotion detection module. *img2pose* is a state-of-the-art facial localization method with real-time inference, which had been shown to achieve reliable performance on the large-scale AFLW2000-3D dataset (Zhu et al., 2016) with a mean-squared error of 3.913. Furthermore, most benchmarking datasets (WIDER Face, FDDB (Jain & Learned-Miller, 2010), Pascal Face (Zhang et al., 2017) on the public benchmarking website *paper-withcode.com* also reported reliable results from *img2pose*, with the minimum accuracy rate reported of 82%. We believe *img2pose* will handle most student movements in our dataset. On the other hand, for the emotion detection module, Residual Masking Network is another well-known competitive method in FER, which achieved 73.28% accuracy on the classic FER2013 dataset. On the public ImageNet challenge, which also employs a more comprehensive image classification challenge than just facial emotion recognition (Deng et al., 2009), Residual Masking Network also achieved a 74.16% top –1 accuracy and 91.91% top –5 accuracy. In our study, we have opted to employ a common approach to emotional recognition, incorporating a classification that comprises seven distinct emotional classes as the output of our analytical pipeline. Specifically, these categories are: Anger, Disgust, Fear, Happiness, Sadness, Surprise, and Neutral.

We believe this approach could serve as a reliable starting point for qualitative analysis, given previous study (Canal et al., 2022) indicates human recognition accuracy of emotions on FER2013 was 65.5%. Our pipeline of processing video data to obtain facial expression recognition is described below (Fig. 1).

### Video Qualitative Analysis

In this study, a multistep analysis was used to investigate when, how, and which shared forms of regulation emerged and functioned during the group work. It also examined if there are challenges associated with its occurrence and what type of strategies the group employed to regulate their emotions. In preparing the data, 30-second segments were created for qualitative content analysis (Silverman, 2020). As a unit of analysis,

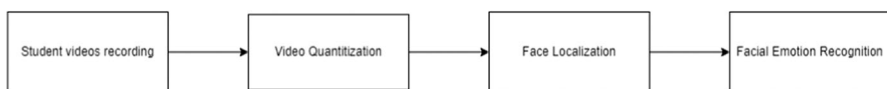


Fig. 1 The experiment pipeline

this provides a chronological overview of the group situation and a consistent and structured way to analyze the situation as it unfolds (Järvenoja et al., 2019).

In the first stage of analysis, each 30-second segment of the videos was coded drawing on self-regulated learning literature (Pintrich, 2000; Zimmerman & Schunk, 2011) and socially shared regulated learning (Järvelä & Hadwin, 2013) to see whether that is a regulatory episode (co-regulation and socially shared regulation) or non-regulatory episode. In this category, we distinguished the segment as either cognitive, socio-emotional or other interaction. The data were coded based on an adaptation of Malmberg et al. (2017) and Järvenoja et al. (2019) coding scheme as they were particularly relevant to our research questions and consistent with the theoretical framework of SSRL. In our analysis, each 30-second segment was coded with a single label. In instances where multiple characteristics were observable within a single segment, the code representing the strongest conceptual semantic meaning was selected. The coding definition is provided in Table 1 along with examples from our data that reflect these codes. While acknowledging its limitations, the choice to adopt a single code per segment was made to align our qualitative analysis more coherently with the subsequent automatic emotion analysis, thereby allowing for more straightforward comparisons and interpretations.

The qualitative video analysis was conducted by a single coder, who is also a researcher in the domain of socially shared regulation of learning. To validate the reliability of the video coding process, a second coder independently annotated a subset of the dataset, specifically focusing on two groups, which constitute 14.2% of the total dataset. Subsequently, a Cohen's Kappa reliability test was conducted. The results yielded a Cohen's Kappa score of 0.71, signifying moderate to high reliability.

### Process Mining to Examine the Patterns of Interactions for Regulation

To investigate the emotional regulation patterns in synchronous computer-supported collaborative learning (CSCL), we utilized a process-mining analysis. This involved the application of the Fuzzy Miner method, as developed by Günther and van der Aalst (2007), to the regulatory activities that had been qualitatively coded. The software used for the analysis was Fluxicon's Disco, which is a commonly used process mining tool in studies examining the process of learning events, as evidenced by prior research (Dindar et al., 2022; Nguyen et al., 2023).

### Learning Regulation Activities and Emotions Co-Occurrences Analysis

The results of facial expression recognition are resampled and matched into the 30-second segments of video qualitative coding for quantitative analysis that seeks to explore the relationship between detected facial emotions and regulatory activities. Descriptive statistics are reported for the distribution of emotions among different regulatory activities. Furthermore, the emotions of different members in each group are aligned for co-occurrences analysis. First, the emotion of each learner in a group is defined with the threshold of 0.5 ( $p > 0.5$ ). Kruskal-Wallis H was applied

**Table 1** Coding scheme for interaction characteristic and regulation types

Categories	Description	Examples
Cognitive	<p>Interaction focuses on the learning-related higher mental process toward metacognitive level (monitoring and controlling) when dialogues focus on:</p> <ul style="list-style-type: none"> <li>- Establishing task demands</li> <li>- Activating prior knowledge</li> <li>- Evaluating and selecting resources and strategies</li> <li>- Checking and evaluating task progression, solution, and overall performance</li> <li>- Prompt questions when notice a noticeable interval in task progression</li> </ul>	<p>It is also the topic we will discuss today, Covid19 pandemic and online learning. How many minutes [as required in the task] has it been?</p>
Socio-emotional	<p>Interaction focuses on strong expressions of socio-emotion with clear negative/positive effects to nature (e.g. showing gratitude, approving, joking, disputing, criticizing, being ironic). This could be task-related or non-task interaction. Expression included verbal (content, tones) and/or non-verbal indicators (e.g. smiling, laughing, frowning, sighing, moaning, facepalm).</p>	<p>S1: There is still a large gap between the rich and poor and I want to solve that problem. S2: God, God [facepalmed], no! You talked like my grandma. You must enjoy life! S3: Great, good job! [Thumb-up, smiling and laughing]</p>
Task execution	<p>Interaction that primary focus on carrying out task requirement, completing the task:</p> <ul style="list-style-type: none"> <li>- Stating and answering the questions prompts</li> <li>- No clear and/or strong affect is indicated</li> </ul>	<p>S1: If we have a chance to learn online a semester one more time, what would you change? S2: [...] I will change my timetable template, the way to study. I don't want to lose the scholarship...[...]</p>
Co-regulation	<p>Interaction between two students where one attempts to influence the other's metacognition, motivation, and/or emotion to assist and guide his/ her learning. This often starts with one student noticing an individual lacking/ obstacle in learning and seeking help OR this is noticed by another. Others initiate regulatory activities. The third member can contribute but only to reflect agreement and not add new ideas.</p>	<p>S1: What is next [in the task]? I forgot already. What is the topic? S2: I think today we will discuss some favorite hobbies or activities.</p>

**Table 1** (continued)

Categories	Description	Examples
Socially shared regulation	<p>Interaction where multiple students contribute to regulated collective activity regarding the group's metacognition, motivation, and/or motivation. After noticing obstacles to the group's progression,</p> <ul style="list-style-type: none"> <li>- Regulatory activities initiated by one student</li> <li>- At least two or more students participate in the discussion and contribute new ideas.</li> </ul>	<p>S1: [Expressing concern about the difficulty of the limited time]  S2: No. It's never too late.  S3: My friend told me that he talked to a foreigner every day.  S2: And after that, was his English better?  S3: Yes, now he can communicate fluently like a native.  S1: Right, we can still make it.</p>
Others	<p>Interaction that is not related to the task topic or objective. E.g., Talking about:</p> <ul style="list-style-type: none"> <li>- Introducing and closing</li> <li>- Out-of-school or school-related activities</li> </ul>	<p>Good evening, everyone, welcome to the video of team 13  I think we should set up a time tomorrow to watch a suggested movie together</p>

to test the difference of emotions co-occurrences between learning regulation behaviors.

## Results

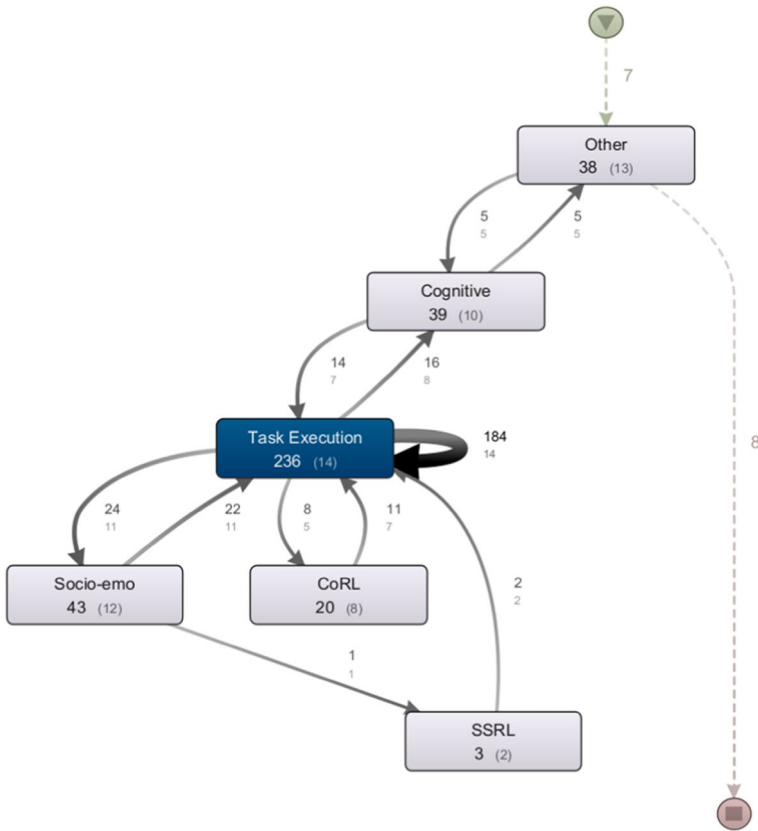
### Learning Regulation Behaviors in Synchronous Online Collaborative Learning

The results of video qualitative analysis for regulatory interactions and learning regulation are described in Table 2. In line with previous studies (e.g., Järvelä et al., 2016; Malmberg et al., 2017), most social interactions in collaborative learning are related to task execution ( $f=62.3\%$ ). While previous studies largely examined social interactions for regulation in face-to-face (Malmberg et al., 2017; Nguyen et al., 2022a, b) and asynchronous online collaborative learning (Iiskala et al., 2015; Lai & Hwang, 2016), our findings confirmed the similar learning regulation behaviors in the context of synchronous online collaborative learning. Furthermore, our findings reported rare occurrences of CoRL and SSRL in synchronous online collaborative learning ( $f=5.3\%$  and  $f=0.8\%$  respectively). These results reflect those of Malmberg et al. (2017) who also found that SSRL infrequently occurred in face-to-face collaborative learning. Furthermore, these results corroborate the ideas of Järvelä et al. (2020), who suggested that social interactions do not often lead to learning regulation. Comparison of the findings with those of other studies confirms the need for promoting learning regulation support in collaborative learning to enhance learning.

To investigate the patterns of learning regulation behaviors in synchronous online collaborative learning, we conducted a process mining analysis. The results of this analysis are presented in Fig. 2. The literature review has demonstrated the significance of learning regulation for achieving success at both the individual and group levels, as evidenced by a number of reports (Dindar et al., 2022). While previous research has examined the patterns of learning regulation in face-to-face collaboration (Järvenoja et al., 2019) and asynchronous online collaborative learning (Iiskala et al., 2015), very little is known about whether similar patterns of learning regulation exist in the context of synchronous computer-supported collaborative learning (CSCL). Our study findings revealed that group learning patterns in synchronous CSCL frequently commence and conclude with other interactions unrelated to the

**Table 2** Learning regulation activities coding

Frequency	Percent	
Cognitive	39	10.3
Socio-emo	43	11.3
Task Execution	236	62.3
Other	38	10.0
CoRL	20	5.3
SSRL	3	0.8
Total	379	100.0



**Fig. 2** Learning regulation pattern in synchronous online collaborative learning

learning process. However, cognitive interactions set the stage for the learning regulation pattern in synchronous CSCL. Our study’s outcomes confirm the existence of similar results, whereby cognitive interactions initiate regulatory adaptation cycles (Nguyen et al., 2023). In accordance with extant literature on the SSRL (Dindar et al., 2020; Nguyen et al., 2023), our empirical investigations similarly reveal a limited incidence of SSRL within collaborative learning environments involving small groups (N=3). Previous research has indicated that while collaborative settings theoretically offer opportunities for shared regulation, the actual manifestation of these regulatory activities externalized in verbal interactions remains infrequent (Isohäätä et al., 2017). Our current study extends the scope of inquiry to encompass emotional regulation as operationalized through facial expressions. Interestingly, we found that co-regulation often iterates with task execution whereas socially shared regulation mainly occurs after socio-emotional interactions. However, with a small sample size, caution must be applied in generalizing the findings. Notwithstanding the relatively limited sample, this study offers interesting insights into learning regulation patterns in synchronous CSCL.

### Learning Regulation Activities and Emotions Co-Occurrences Analysis

The results of facial emotion recognition were aligned among group members and matched with learning regulation activities for conducting co-occurrences analysis. For each 30-second segment that was coded, we calculated the total counts of each type of emotion on a second-by-second basis. This granular approach enables a more nuanced understanding of the emotional dynamics within each coded interaction. Emotional synchrony was recognized whenever two or more group members shared the same emotion in each second frame. Figure 3 illustrates the distribution of shared emotion in different learning regulation activities.

Interestingly, the most frequently shared emotion expressed by the learners is *surprised*. In the context of clinical reasoning, Lajoie et al. (2021) has reported that *angry* emotion appeared most frequently in self-regulated learning while *surprised* tended to occur with the second highest frequency. In an effort to confirm accuracy and provide more insight into this finding difference, we have reviewed the video data for different shared emotion segments. We found out that a possible explanation for this might be due to the nature of the learning tasks. In our study, the students were asked to share opinions and initiate discussions in English in the context of learning English as a second language. Learners often expressed surprised emotions while listening to their group members' sharing information and thoughts. A further

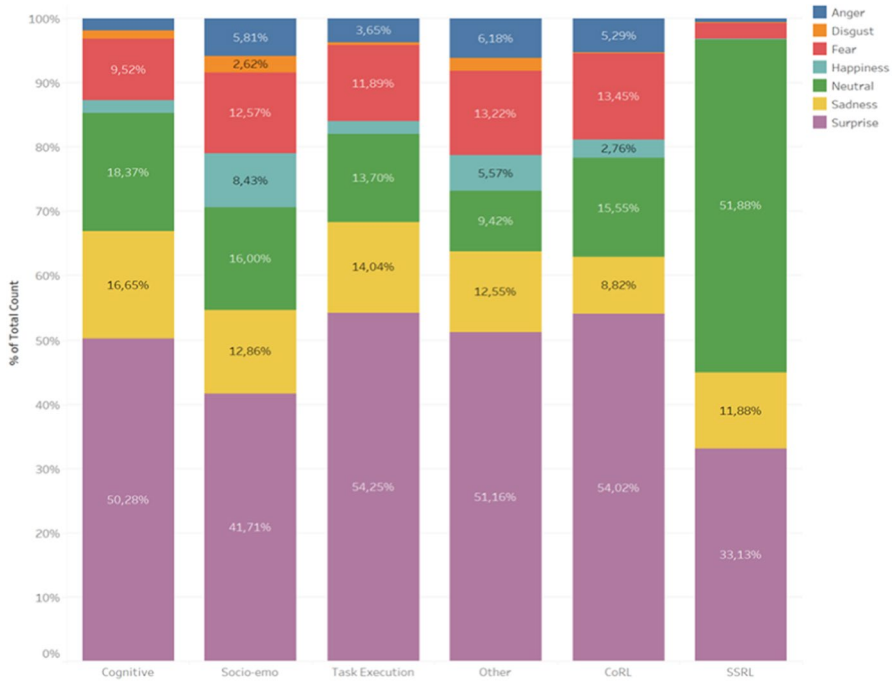


Fig. 3 Emotion distribution for different learning regulation activities

study with more focus on the effects of learning task design on emotional aspects of learning is therefore suggested.

Table 3 reported the descriptive statistics for emotional synchronies in different learning regulation activities. Mean refers to the average occurrence of emotional synchronies within the 30-second segment of interactions. The study sought to show evidence of difference in the number of emotional synchronies among different regulatory activities. We have acknowledged the limitation of this 30 s segments in the revision and future studies will apply more granular analysis for video coding. The emotional synchronies appear most often in SSRL segments ( $M=8.67$ ,  $SD=10.69$ ) while occurring least in socio-emotional interactions ( $M=4.49$ ,  $SD=4.32$ ). This result may be explained by the fact that SSRL is established based on the sharedness of emotion and cognition among group members (Järvelä et al., 2019). On the other hand, socio-emotional interactions are needed for negotiating between learners in a group while confronting emotional challenges in which the group members may not share similar emotions (Järvenoja et al., 2019). The Kruskal-Wallis H Test showed a slightly significant difference in the number of emotional synchronies among different regulatory activities with  $p=0.029$ . This combination of findings provides some support for the conceptual premise that AI techniques could be utilized to provide new insights into the process of learning regulation. Nevertheless, the close mean values for emotional synchrony frequency among interactions for SSRL ( $M_{SSRL} = 8.67$ ), Cognitive ( $M_{Cognitive} = 8.31$ ), and Task Execution ( $M_{Task Execution} = 8.50$ ) suggest that this metric is not sufficient to differentiate among these interactions. The issue is compounded by the large standard deviations, making it challenging to attribute a future mean value, such as 8.59, to any specific condition. These findings indicate a need for more discriminative measures in future research.

## Discussion

The main goal of the current study was to provide empirical evidence of how AI methods can be implemented to investigate interrelation between individual affective states and their shared regulatory activities in CSCL. While previous studies have provided evidence for learning regulation in face-to-face collaborative learning and asynchronous CSCL, much less is known about the phenomenon in synchronous

**Table 3** Emotional synchronies in learning regulation activities

Activities	Absolute frequency of emotional synchronies	Mean	Std. deviation
Cognitive	324	8.31	7.95
Socio-emo	193	4.49	4.32
Task execution	2006	8.50	7.78
Other	224	5.89	7.29
CoRL	117	5.85	6.74
SSRL	26	8.67	10.69
Total	2890	7.63	7.50

CSSL. Accordingly, this study seeks to address the research gap by utilizing AI-enhanced techniques to examine learning regulation in the context of synchronous CSSL. Furthermore, this is needed as prior studies have noted the importance of developing innovative methods to capture all the phases, facets, and changes over time of regulatory activities to advance the field of study (Hadwin et al., 2018; Järvelä et al., 2019). The study introduces an approach for analyzing this relationship by aligning the core emotions, captured by FER with its corresponding interaction when examining co-regulation and socially shared regulation in a synchronous CSSL environment. In this regard, computer vision and machine learning are proving useful in detecting micro expressions, moods, and temperaments of learners by automatically detecting nonverbal behavior and affect (Behera et al., 2020).

Despite a growing interest that recognizes the intertwined between learners' emotions and cognition that guide the learning process (Nguyen et al., 2022a, b, 2023; Woolf et al., 2009), studies of emotion regulation are still few in comparison to other constructs (motivation, cognition, metacognition). Saariaho et al. (2016)'s study examines the emotional landscape that student-teachers experience during self- and co-regulated learning. While their results confirm prior findings on positive emotions are emphasized as essential elements across all regulatory phases, the retrospective approach of the study to analyze the sketched emotion visualization of past experience arguably can affect data accuracy. In a study carried out by Ucan and Webb (2015), emotion regulation was examined along with motivational and metacognitive control. Our study illuminates the need for more refined empirical work to distinguish the high values of emotional synchrony between SSRL and other interactions such as Cognitive and Task Execution. The current data points to a trend but does not offer a statistically significant basis for differentiation, a point that should be considered when interpreting these findings.

Järvelä et al. (2020) suggests that "with the aid of advanced technologies, multidisciplinary collaboration between the learning sciences, affective computing and machine learning can help to study these complex phenomena" (p. 2392). The complex phenomena refer to SSRL in collaborative learning in which learners actively engage, sustain, and regulate their cognition, emotions, motivation, and behaviors towards the accomplishment of their learning goals. In the context of synchronous CSSL, this multidisciplinary study offered empirical evidence for this research proposition. The study does not only contribute to our understanding of learning regulation but also provides a methodological approach proven to be useful in expanding our understanding of how AI could inform research in learning sciences. In this study, we purposely focus on shared regulations and the affect aspect, adapting Järvenoja et al. (2019) method of studying emotion regulation from video data, however, with the added alignment of unobtrusive measures of emotion from FER. The method of coding for regulatory episodes, metacognitive processes, emotion regulation strategies during collaborative learning revealed the dynamic relationship between emotions and regulation.

The study contributes to the literature on educational technology, which has endorsed the use of AI as a promising tool for transforming education. Although the literature has shown that AI technology has brought great opportunities for improving learning and teaching, the design and development of AI-enhanced support for

learning and teaching have faced several challenges (Järvelä et al., 2020; Ouyang & Jiao, 2021; Roll & Wylie, 2016). For instance, as AI in education is an interdisciplinary area of research, there is still a divide between those who understand AI the methods and techniques and those who know how AI could be applied to offer benefits to learning and teaching. Furthermore, there remains a paucity of evidence on the systematic application approach of AI in educational research. Utilizing the role of AI models as scientific tools (Baker, 2000), this study attempts to bridge the gap between AI machine learning and learning research. Despite the importance of learning regulation and collaborative learning, in reviewing the literature, there have not been many attempts on using AI technologies to advance our understanding and support learning regulation in synchronous CSCL. Accordingly, our work hopes to find a methodological and theoretical grounding for further development and implementation of AI in education to support learning regulation.

## Limitation and Future Directions

We acknowledge a few limitations to our work. Firstly, the study is confined to a limited sample size with participants from a higher education setting and observation was conducted online based on recordings of the language sessions. Nevertheless, their cognitive, socio-emotional, co-regulation and socially shared regulation was examined through the group dynamics consisting of fixed analysis units, which equated to a reasonably sufficient amount of data to validate findings. Secondly, due to the restricted quantity of problem-solving tasks and the predominance of certain modes of learning task designs, some key emotions may have overridden the overall results. Therefore, there is a need for future research to diversify examination in terms of emotional regulation in learning across different contexts and learner characteristics, in order to compare, contrast and cross-validate findings obtained from this study. Thirdly, although our research adopted the video analysis approach with the 30s segment, a technique that has been approved by previous studies (Järvenoja et al., 2019), there may be scope to examine different approaches that can provide insights at different levels of granularity. Lastly, due to the distinct nature of our data collection method which espoused the use of AI to conduct research directly on Teams and via the platform of video-based qualification, risks may arise from spontaneous technical issues while facial expression may not fully reflect the regulatory processes. In response, efforts have been made in the data analysis phase which employed external coders to evaluate and infer a person's subjectively experienced emotions to verify results. We have also partially indicated the emotional processes/temporal emotions of students through the session, and triangulated with coding of their interactions help to establish the link between emotion and their regulatory interactions. Attempts have also been made to allow inferences about differences detected in learners and processes across the learning sessions. Regardless of these limitations, this work provides valuable insights into facial expression recognition for emotional regulation in online learning within a collaborative, socially shared learning context. Moreover, as the ultimate goal of this study is to inform learning regulation through the application of AI facial expression recognition, our results

shed light on the potential of this methodological approach, rather than proposing a new model. Therefore, future scholars are encouraged to consider extending the sample size of students and testing the model across stages in diversified classroom settings (e.g., no internet connection, interruption during teaching and learning activities, constrained teacher - student interactions, etc.). Potential approaches for future work may also encompass the possibility of adding other covariates (e.g., only females or males, different ages, etc.) to determine the effectiveness of this methodology across contexts and genders as well as to avoid sample bias.

**Funding** Open Access funding provided by University of Oulu (including Oulu University Hospital). This research has funded by the Research Council of Finland (a.k.a. Academy of Finland) grant 350249.

**Data Availability** The data that support the findings of this study could be available with a data transfer agreement on request from the corresponding author, Andy Nguyen (andy.nguyen@oulu.fi). The data are not publicly available due to restrictions apply to the availability of these data as their containing information that could compromise the privacy of research participants.

## Declarations

**Conflicting Interests** The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Albiero, V., Chen, X., Yin, X., Pang, G., & Hassner, T. (2020). img2pose: Face Alignment and Detection via 6DoF, Face Pose Estimation. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Azevedo, R., & Gašević, D. (2019). Analyzing multimodal multichannel data about self-regulated learning with advanced learning technologies: Issues and challenges. *Computers in Human Behavior*, *96*, 207–210. <https://doi.org/10.1016/j.chb.2019.03.025>
- Baker, M. (2000). *The roles of models in Artificial Intelligence and Education research: A prospective view*. <https://www.semanticscholar.org/paper/The-roles-of-models-in-Artificial-Intelligence-and-Baker/5adc4bd6aa6aec5ddd00b9aa63326dc2298e2b28>. Accessed 11 May 2022.
- Behera, A., Matthew, P., Keidel, A., Vangorp, P., Fang, H., & Canning, S. (2020). Associating facial expressions and Upper-Body gestures with learning tasks for enhancing intelligent tutoring systems. *International Journal of Artificial Intelligence in Education*, *30*(2), 236–270. <https://doi.org/10.1007/s40593-020-00195-2>
- Canal, F. Z., Müller, T. R., Matias, J. C., Scotton, G. G., de Sa Junior, A. R., Pozzebon, E., & Sobieranski, A. C. (2022). A survey on facial emotion recognition techniques: A state-of-the-art literature review. *Information Sciences*, *582*, 593–617. <https://doi.org/10.1016/j.ins.2021.10.005>

- Dalvi, C., Rathod, M., Patil, S., Gite, S., & Kotecha, K. (2021). A survey of AI-based facial emotion recognition: Features, ML & DL Techniques, age-wise datasets and future directions. *Ieee Access : Practical Innovations, Open Solutions*, 9, 165806–165840. <https://doi.org/10.1109/ACCESS.2021.3131733>
- Debie, E., Rojas, R. F., Fidock, J., Barlow, M., Kasmarik, K., Anavatti, S., Garratt, M., & Abbass, H. A. (2019). Multimodal fusion for objective assessment of cognitive workload: A review. *IEEE Transactions on Cybernetics*, 1–14. <https://doi.org/10.1109/tcyb.2019.2939399>
- Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A Large-Scale Hierarchical Image Database. *IEEE Conference on Computer Vision and Pattern Recognition*.
- Dindar, M., Järvelä, S., & Haataja, E. (2020). What does physiological synchrony reveal about meta-cognitive experiences and group performance? *British Journal of Educational Technology*, 51(5), 1577–1596.
- Dindar, M., Järvelä, S., Nguyen, A., Haataja, E., & Çini, A. (2022). Detecting shared physiological arousal events in collaborative problem solving. *Contemporary Educational Psychology*, 69, 102050. <https://doi.org/10.1016/j.cedpsych.2022.102050>
- Fan, Y., Matcha, W., Uzir, N. A., Wang, Q., & Gašević, D. (2021). Learning analytics to reveal links between learning design and self-regulated learning. *International Journal of Artificial Intelligence in Education*, 31(4), 980–1021. <https://doi.org/10.1007/s40593-021-00249-z>
- Gao, Z., Dang, W., Wang, X., Hong, X., Hou, L., Ma, K., & Perc, M. (2021). Complex networks and deep learning for EEG signal analysis. *Cognitive Neurodynamics*, 15(3), 369–388. <https://doi.org/10.1007/s11571-020-09626-1>
- Giannopoulos, P., Perikos, I., & Hatzilygeroudis, I. (2018). Deep learning approaches for facial emotion recognition: A case study on FER-2013. In *Advances in hybridization of intelligent methods: Models, systems and applications* (pp. 1–16). Cham: Springer.
- Goodfellow, I. J., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B., Cukierski, W., Tang, Y., Thaler, D., Lee, D. H., Zou, Y., Ramaiah, C., Feng, F., Li, R., Wang, X., Athanasakis, D., Shave-Taylor, J., Milakov, M., Park, J., & Bengio, Y. (2013). Challenges in Representation Learning: A Report on Three Machine Learning Contests. In M. Lee, A. Hirose, Z.-G. Hou, & R. M. Kil (Eds.), *Neural Information Processing* (pp. 117–124). Springer. [https://doi.org/10.1007/978-3-642-42051-1\\_16](https://doi.org/10.1007/978-3-642-42051-1_16)
- Günther, C. W., & van der Aalst, W. M. P. (2007). Fuzzy mining – adaptive process simplification based on multi-perspective metrics. In G. Alonso, P. Dadam, & M. Rosemann (Eds.), *Business Process Management* (pp. 328–343). Springer. [https://doi.org/10.1007/978-3-540-75183-0\\_24](https://doi.org/10.1007/978-3-540-75183-0_24)
- Hadwin, A. F., Bakhtiar, A., & Miller, M. (2018). Challenges in online collaboration: Effects of scripting shared task perceptions. *International Journal of Computer-Supported Collaborative Learning*, 13(3), 301–329. <https://doi.org/10.1007/s11412-018-9279-9>
- Iiskala, T., Volet, S., Lehtinen, E., & Vauras, M. (2015). Socially shared metacognitive regulation in asynchronous CSCL in science: Functions, evolution and participation. *Frontline Learning Research*, 3(1), 78–111. <https://doi.org/10.14786/flr.v3i1.159>
- Isohäätä, 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.
- Isohäätä, J., Näykki, P., Järvelä, S., & Baker, M. J. (2018). Striking a balance: Socio-emotional processes during argumentation in collaborative learning interaction. *Learning Culture and Social Interaction*, 16, 1–19. <https://doi.org/10.1016/j.lcsi.2017.09.003>
- Jain, V., & Learned-Miller, E. (2010). *FDDB: A Benchmark for Face Detection in Unconstrained Settings* (UMass Amherst Technical Report).
- Järvelä, S., & Bannert, M. (2021). Temporal and adaptive processes of regulated learning—what can multimodal data tell? *Learning and Instruction*, 72, 101268. <https://doi.org/10.1016/j.learninstruc.2019.101268>
- Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist*, 48(1), 25–39. <https://doi.org/10.1080/00461520.2012.748006>
- Järvelä, S., Malmberg, J., & Koivuniemi, M. (2016). Recognizing socially shared regulation by using the temporal sequences of online chat and logs in CSCL. *Learning and Instruction*, 42, 1–11. <https://doi.org/10.1016/j.learninstruc.2015.10.006>
- Järvelä, S., Järvenoja, H., & Malmberg, J. (2019). Capturing the dynamic and cyclical nature of regulation: Methodological Progress in understanding socially shared regulation in learning. *International Journal of Computer-Supported Collaborative Learning*, 14(4), 425–441. <https://doi.org/10.1007/s11412-019-09313-2>
- Järvelä, S., Gašević, D., Seppänen, T., Pechenizkiy, M., & Kirschner, P. A. (2020). Bridging learning sciences, machine learning and affective computing for understanding cognition and affect in

- collaborative learning. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.12917>. bjet.12917-bjet.12917.
- Järvenoja, H., Näykki, P., & Törmänen, T. (2019). Emotional regulation in collaborative learning: When do higher education students activate group level regulation in the face of challenges? *Studies in Higher Education*, 44(10), 1747–1757. <https://doi.org/10.1080/03075079.2019.1665318>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 25. <https://papers.nips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html>. Accessed 11 May 2022.
- Kwon, K., Liu, Y. H., & Johnson, L. P. (2014). Group regulation and social-emotional interactions observed in computer supported collaborative learning: Comparison between good vs. poor collaborators. *Computers and Education*, 78, 185–200. <https://doi.org/10.1016/j.compedu.2014.06.004>
- Lai, C.-L., & Hwang, G.-J. (2016). A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. *Computers & Education*, 100, 126–140. <https://doi.org/10.1016/j.compedu.2016.05.006>
- Lajoie, S. P., Zheng, J., Li, S., Jarrell, A., & Gube, M. (2021). Examining the interplay of affect and self regulation in the context of clinical reasoning. *Learning and Instruction*, 72, 101219. <https://doi.org/10.1016/j.learninstruc.2019.101219>
- Malmberg, J., Järvelä, S., & Järvenoja, H. (2017). Capturing temporal and sequential patterns of self-, co-, and socially shared regulation in the context of collaborative learning. *Contemporary Educational Psychology*, 49, 160–174. <https://doi.org/10.1016/j.cedpsych.2017.01.009>
- Mellouk, W., & Handouzi, W. (2020). Facial emotion recognition using deep learning: Review and insights. *Procedia Computer Science*, 175, 689–694. <https://doi.org/10.1016/j.procs.2020.07.101>
- Nguyen, A., Järvelä, S., Wang, Y., Róse, C., Malmberg, J., & Järvenoja, H. (2021). Multimodal deep learning model for detecting types of interactions for regulation in collaborative learning. *Proceedings of Annual Meeting of the International Society of the Learning Sciences (ISLS)*. <https://2021.isls.org/proceedings/>. Accessed 7 May 2022.
- Nguyen, A., Hong, Y., Dang, B., & Nguyen, P. (2022a). Emotional regulation in synchronous online collaborative learning: A facial expression recognition study. *ICIS 2022 Proceedings*. [https://aisel.aisnet.org/icis2022/learning\\_iscurricula/learning\\_iscurricula/12](https://aisel.aisnet.org/icis2022/learning_iscurricula/learning_iscurricula/12). Accessed 1 May 2022.
- Nguyen, A., Järvelä, S., Wang, Y., & Róse, C. (2022b). Exploring socially shared regulation with an AI deep learning approach using multimodal data. *Proceedings of International Conferences of Learning Sciences (ICLS)*. <https://2022.isls.org/proceedings/>. Accessed 5 May 2022.
- Nguyen, A., Järvelä, S., Rosé, C., Järvenoja, H., & Malmberg, J. (2023). Examining socially shared regulation and shared physiological arousal events with multimodal learning analytics. *British Journal of Educational Technology*, 54(1), 293–312. <https://doi.org/10.1111/bjet.13280>
- Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, 2, 100020. <https://doi.org/10.1016/j.caeai.2021.100020>
- 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. <https://doi.org/10.1080/00313831.2015.1066436>
- Pham, L., Vu, T. H., & Tran, T. A. (2021). Facial expression recognition using residual masking network. *2020 25th International Conference on Pattern Recognition (ICPR)*, 4513–4519. <https://doi.org/10.1109/ICPR48806.2021.9411919>
- Pintrich, P. R. (2000). Chapter 14—The Role of Goal Orientation in Self-Regulated Learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 451–502). Academic Press. <https://doi.org/10.1016/B978-012109890-2/50043-3>
- Roll, I., & Wylie, R. (2016). Evolution and revolution in Artificial Intelligence in Education. *International Journal of Artificial Intelligence in Education*, 26(2), 582–599. <https://doi.org/10.1007/s40593-016-0110-3>
- Saariaho, E., Pyhäälto, K., Toom, A., Pietarinen, J., & Soini, T. (2016). Student teachers' self- and co-regulation of learning during teacher education. *Learnings: Research and Practice*, 2(1), 44–63. <https://doi.org/10.1080/23735082.2015.1081395>
- Silverman, D. (Ed.). (2020). *Qualitative research*. Sage.
- Taub, M., & Azevedo, R. (2016). Using Eye-Tracking to Determine the Impact of Prior Knowledge on Self-Regulated Learning with an Adaptive Hypermedia-Learning Environment. In A. Micarelli, J. Stamper, & K. Panourgia (Eds.), *Intelligent Tutoring Systems* (pp. 34–47). Springer International Publishing. [https://doi.org/10.1007/978-3-319-39583-8\\_4](https://doi.org/10.1007/978-3-319-39583-8_4)

- Ucan, S., & Webb, M. (2015). Social Regulation of Learning during Collaborative Inquiry Learning in Science: How does it emerge and what are its functions? *International Journal of Science Education*, 37(15), 2503–2532. <https://doi.org/10.1080/09500693.2015.1083634>
- Winne, P. H. (2019). Paradigmatic dimensions of Instrumentation and Analytic methods in research on self-regulated learning. *Computers in Human Behavior*, 96, 285–289. <https://doi.org/10.1016/j.chb.2019.03.026>
- Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., & Picard, R. (2009). Affect-aware tutors: Recognising and responding to student affect. *International Journal of Learning Technology*, 4(3/4), 129–164. <https://doi.org/10.1504/IJLT.2009.028804>
- Yang, S., Luo, P., Loy, C. C., & Tang, X. (2016). WIDER FACE: A face detection benchmark. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, 5525–5533. <https://doi.org/10.1109/CVPR.2016.596>
- Zhang, S., Zhu, X., Lei, Z., Shi, H., Wang, X., & Li, S. Z. (2017). S<sup>3</sup>FD: Single shot scale-invariant face detector. *IEEE International Conference on Computer Vision (ICCV)*, 2017, 192–201. <https://doi.org/10.1109/ICCV.2017.30>
- Zhang, Z., Cao, T., Shu, J., & Liu, H. (2020). Identifying key factors affecting college students' adoption of the e-learning system in mandatory blended learning environments. *Interactive Learning Environments*, 1–14. <https://doi.org/10.1080/10494820.2020.1723113>
- Zhu, X., Lei, Z., Liu, X., Shi, H., & Li, S. Z. (2016). Face alignment across large poses: A 3d solution. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 146–155). IEEE.
- Zimmerman, B. J., & Schunk, D. H. (2011). *Handbook of self-regulation of learning and performance*. <https://psycnet.apa.org/record/2011-12365-000>. Accessed 5 May 2022.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.