JOURNAL OF LEARNING ANALYTICS

Volume 12(1), 238–252. https://doi.org/10.18608/jla.2025.8661

Examining the Interplay of Gaze and Verbal Interactions in Socially Shared Regulation of Learning: A Transmodal Analysis (TMA) Study

Andy Nguyen[1], Yeyu Wang[2], Ridwan Whitehead[3], Muhammad Ashiq[4], Sanna Järvelä[5] and David Williamson Shaffer[6]

Abstract

This paper presents a transmodal analysis (TMA) study that investigates the interplay between gaze and verbal interactions through regulation-triggering events within the context of socially shared regulation of learning (SSRL) in face-to-face collaborative settings with shared computer-mediated materials. In face-to-face collaborative learning environments, gaze interaction serves as a pivotal mechanism for both directing attention and communicating emotional states. This aspect of non-verbal communication is critical for aligning cognitive processes and establishing an emotional connection among learners, thereby improving the effectiveness of the learning environment. Despite substantial research on verbal interactions in SSRL, the role of non-verbal cues, particularly gaze, has been less studied. This study examines 3,523 gaze and verbal interactions from twenty-four high school students engaged in eight collaborative learning groups. This study demonstrates the application of TMA method, which reveals epistemic networks across various data modalities, each operating on different temporal scales. Furthermore, our study highlights the critical role of gaze interactions in conveying cognitive and emotional cues, enhancing verbal communication for SSRL.

Notes for Practice

- The transmodal analysis (TMA) method provides a valuable framework for analyzing multimodal data in learning processes, offering a comprehensive lens to capture and interpret the interaction of various modes, particularly gaze and verbal interactions.
- Different regulation-triggering events, such as cognitive or emotional challenges, occur in collaborative learning, prompting the group to implement suitable regulatory responses.
- After cognitive regulation–triggering events, students often aim for mutual understanding by making eye contact with their peers or briefly pausing their thought process by looking at surrounding objects before engaging in metacognitive interactions.
- After emotional regulation–triggering events, students often engage in more in-depth discussions about planning and assessing their tasks

Keywords: Multimodal learning analytics, socially shared regulation of learning, transmodal analysis

Submitted: 11/09/2024 — **Accepted:** 19/12/2024 — **Published:** 23/03/2025

Corresponding author [1]Email: andy.nguyen@oulu.fi Address: Learning and Education Technology Research Lab, University of Oulu, Pentti Kaiteran katu 1, 90570 Oulu, Finland. ORCID iD: https://orcid.org/0000-0002-0759-9656

[2]Email: ywang2466@wisc.edu Address: Department of Educational Psychology, University of Wisconsin-Madison, Educational Sciences, 880, 1025 W. Johnson St., Madison, WI 53706, USA. ORCID iD: https://orcid.org/0000-0003-1978-5453

[3]Email: ridwan.whitehead@oulu.fi Address: Learning and Education Technology Research Lab, University of Oulu, Pentti Kaiteran katu 1, 90570 Oulu, Finland. ORCID iD: https://orcid.org/0009-0002-2888-7304

[4]Email: ashiq@wisc.edu Address: Department of Educational Psychology, University of Wisconsin-Madison, Educational Sciences, 880, 1025 W. Johnson St., Madison, WI 53706, USA.

[5]Email: sanna.jarvela@oulu.fi Address: Learning and Education Technology Research Lab, University of Oulu, Pentti Kaiteran katu 1, 90570 Oulu, Finland. ORCID iD: https://orcid.org/0000-0001-6223-3668

[6]Email: dws@education.wisc.edu Address: Department of Educational Psychology, University of Wisconsin-Madison, Educational Sciences, 880, 1025 W. Johnson St., Madison, WI 53706, USA. ORCID iD: https://orcid.org/0000-0001-9613-5740

1. Introduction

Self-regulated learning (SRL) is widely acknowledged as an effective and impactful approach to learning (Hadwin et al., 2018), and its significance has grown even more critical in the expansion of generative artificial intelligence (Giannakos et al., 2024). SRL refers to the process by which individuals actively manage their cognitive, motivational, and emotional resources to achieve learning goals. It involves setting objectives, monitoring progress, employing strategies to overcome challenges, and reflecting on outcomes. SRL is pivotal for fostering autonomy and adaptability in learners. Building upon SRL theory, socially shared regulation of learning (SSRL) extends the principles of SRL to group contexts, emphasizing the co-regulation of learning processes among group members. The concept of SSRL is rooted in the idea that successful collaboration relies not only on individual self-regulation in learning but also on the group members' ability to co-regulate learning with each other hence together manage the group learning processes collectively (Hadwin et al., 2018). Recent years of empirical research have evidenced that SSRL is a crucial aspect of collaboration that involves the collective control of cognitive, motivational, and emotional processes to achieve shared learning goals. SSRL involves the processes through which group members mutually adjust their behaviours and strategies, communicate their understanding, set shared goals, and monitor their progress toward these goals (Volet et al., 2009). These processes may enhance learning outcomes (De Backer et al., 2022) as well as promote a supportive and engaged learning atmosphere by sharing emotional and motivational states among the group members (Bakhtiar et al., 2018).

SSRL is characterized by the social interactions among group members as they negotiate understanding, share ideas, and coordinate their efforts (Hadwin et al., 2018). Over the years, several studies have contributed to our understanding of SSRL through examining social interactions within collaborative learning (Isohätälä et al., 2020; Vuorenmaa et al., 2023). Nevertheless, many of these studies have primarily focused on verbal interactions, often neglecting the non-verbal aspects of social interactions. Non-verbal interactions, including facial expressions, gestures, eye contact, and body language, can convey essential information about learners' cognitive and emotional states (Purwati et al., 2019), fostering a deeper understanding among group members and facilitating the SSRL process. Moreover, non-verbal cues can also influence group dynamics, such as building rapport, establishing trust, and maintaining group cohesion, which are critical factors in successful SSRL.

Recently, the literature on SSRL has increasingly recognized the importance of examining non-verbal interactions, as these non-verbal signals play a crucial role in social interactions and are likely to have a significant impact on collaborative learning and SSRL (Järvelä et al., 2023; Nguyen, Li, et al., 2023; Zhou et al., 2023). For example, facial expressions can reveal a learner's emotional process in learning (Nguyen, Li, et al., 2023), their confusion or comprehension, prompting other group members to adjust their explanations or provide further support. Yet, despite the growing emphasis on the role of non-verbal interactions in SSRL, comprehensive research exploring the relationship between verbal and different non-verbal interactions within SSRL is still required (Järvelä et al., 2023).

In line with this research direction, the aim of this study is to examine the role of gaze and verbal interactions in SSRL. Gaze, in particular, is a powerful non-verbal signal, providing cues about attention, interest, and understanding (Richardson & Dale, 2005; Schneider & Pea, 2017). Among non-verbal interactions in collaborative learning, gaze interaction has been recognized as a significant means of communication that can affect the regulatory dynamics in a group (Richardson & Dale, 2005). Its significance emanates from its ability to offer insight into what learners are paying attention to at a particular moment. In collaborative learning, gaze patterns can impact the dynamics of the group interactions (Schneider & Pea, 2017) and, subsequently, the process of SSRL. As participants immerse themselves in SSRL, certain gaze behaviours act as regulatory indicators, assisting in offering support or modifying the group's strategy for a given task. Notwithstanding the existing literature on eye gaze dynamics in collaborative learning, there is a noticeable gap in the literature when it comes to a holistic understanding of the interplays between gaze and verbal interactions within the context of SSRL.

To address this research gap, the study aims to employ a novel trans-model analysis (TMA) approach to examine both verbal and gaze interactions in SSRL in responses to regulation-triggering events. The TMA approach allows for the consideration of the multimodality and multidimensionality of collaborative learning and facilitates the exploration of the complex interplay between these verbal and non-verbal interactions in SSRL processes. However, research also recognizes that it is challenging to identify regulatory moments due to the dynamics of collaborative learning settings and non-observable cognitive and emotional processes in SSRL (Järvelä & Bannert, 2021). Therefore, to better study the interplay between verbal and gaze interactions in SSRL, we examine these interactions throughout manifested cognitive and emotional regulation–triggering events. Regulation-triggering events, such as task-related challenges or interpersonal conflicts, can prompt the activation of SSRL processes that influence the overall collaborative learning process. In particular, they refer to events, situations, or incidents that may inhibit the learning process and thus require learners' regulatory responses (Järvelä et al., 2023; Järvelä & Hadwin, 2024).

The complex interplay between cognitive and emotional processes in collaborative learning necessitates a comprehensive examination of regulation-triggering events. As a result, this study aims to utilize TMA to advance the existing literature on

SSRL in collaborative learning by offering a detailed understanding of gaze and verbal interactions. Specifically, we aim to address the following research questions:

RQ1. How do gaze interactions complement or interact with verbal interactions to facilitate socially shared regulation of learning (SSRL) in a face-to-face collaborative learning environment?

RQ2. How do different types of regulation-triggering events impact the interplays between verbal and gaze interactions in a face-to-face collaborative learning environment?

a. How do the interplays between verbal and gaze interactions change after the cognitive regulation–triggering event?

b. How do the interplays between verbal and gaze interactions change after the emotional regulation–triggering event?

By doing so, this study not only adds a crucial layer to our understanding of SSRL but also aids in developing effective strategies for facilitating collaborative learning. This study aims to contribute to the literature on collaborative learning by offering a richer, more nuanced understanding of the regulatory processes involved in SSRL. Our findings can inform the development of targeted interventions and instructional strategies that can facilitate effective communication and regulation processes within collaborative learning environments. Gaining a deeper understanding of these verbal and gaze interactions and their possible implications for learning outcomes can inform the design of more effective instructional strategies and support systems for collaborative learning experiences.

2. Theoretical Framework

2.1. Socially Shared Regulation in Collaborative Learning

The theoretical underpinnings of self-regulated learning (SRL) provide the foundations for understanding and modelling socially shared regulation of learning (SSRL). SRL refers to an agentic process wherein learners strategically take control of their learning engagement and situations through active cycles of planning, performance, and reflection (Hadwin et al., 2018; Zimmerman, 2011). The regulation of learning in collaborative learning settings is seen as a dynamic, multifaceted process in which members engage in shared tasks, define common goals, plan collaboratively, and monitor each other's contributions to ensure that shared goals are achieved (Järvelä & Bannert, 2021). In collaborative learning, these strategic processes extend beyond the individual members of a group through the medium of social interaction (e.g., Huang & Lajoie, 2023). Although regulation of learning is a mental process that originates in learners' intentions, beliefs, and past sociohistorical experiences (Hadwin et al., 2018), regulation in collaborative learning is also an inherently interactive process.

The exploration of SSRL within collaborative learning environments has shown it to be a highly complex and dynamic process (Azevedo & Gašević, 2019; Järvelä et al., 2023). Challenges in studying and supporting SSRL lie in the inherent nature of the non-observable cognitive and emotional processes involved (Huang & Lajoie, 2023). Recently, learning analytics methods, particularly those utilizing machine learning and artificial intelligence techniques, offer novel opportunities to identify patterns and make predictions about learners' cognitive and emotional states (Azevedo & Gašević, 2019; Molenaar et al., 2023). This can enable a deeper investigation of the processes underpinning SSRL and provide valuable insights into how these processes contribute to collaborative learning.

Recently, Järvelä and Hadwin (2024) introduced an empirically testable model of triggers as metacognitive markers for regulation in learning. This lays the foundation for a conceptual framework aimed at enhancing our understanding of regulation within both individual and collaborative learning contexts. As such, it provides an SRL theory–guided conceptual framework for exploiting the wide array of multimodal data sources to study SSRL (Järvelä & Hadwin, 2024). The proposed trigger concept framework is also instrumental in operationalizing advanced learning technologies, such as artificial intelligence (AI) methodologies (Dang et al., 2024; Nguyen et al., 2022) to afford empirical insights into these regulatory patterns, which paves the way for a comprehensive understanding of SSRL processes. This trigger concept framework provides a theory-guided path for advanced research into regulation during complex learning in both individual and collaborative learning contexts, offering empirical markers to identify regulation occurrences.

Although the trigger concept framework proposed by Järvelä et al. (2023) and Järvelä and Hadwin (2024) presents a promising tool for the identification of traces, sequences, patterns, and models of SSRL, additional empirical evidence is necessitated to support it. A challenge to the collection of this evidence is the infrequency and paucity with which SSRL is observed in authentic learning environments, as highlighted in previous studies. To address this obstacle and generate substantiating evidence for the trigger concept framework, this study employed an experimental design, focusing on regulation-triggering events. This approach allowed for the systematic capture of trigger signals and their corresponding responses. Through this strategic experimental design, we aim to gather relevant data on the complex, interactive processes of SSRL, ultimately contributing to the enhancement and validation of the trigger concept framework.

2.2. Verbal and Gaze Interactions for SSRL in a Face-to-Face Collaborative Learning Environment

SSRL is, by nature, a collaborative endeavour, which means that learners must be able to communicate effectively to achieve shared learning goals. In understanding the complexities of SSRL, it is crucial to investigate social interactions through which collaborative learners coordinate their cognitive and affective resources (Vuorenmaa et al., 2023). Verbal interactions, generally foregrounded in SSRL research, contribute significantly to collaborative learning. While prior studies on verbal interactions for SSRL have significantly enriched our understanding of SSRL dynamics, an exclusive focus on the verbal domain provides only a partial view of SSRL processes, sidelining the critical non-verbal components. Recently, the importance of non-verbal interactions, such as gazes, facial expressions, gestures, and body language, in conveying cognitive and emotional information and facilitating SSRL is gaining recognition (Nguyen, Li, et al., 2023). They can indicate a learner's comprehension or confusion, reveal their level of engagement, or signal a shift in mood or motivation. This information is crucial for effective collaboration, as it allows group members to gauge each other's states and adjust their actions accordingly. Thus, a comprehensive understanding of SSRL necessitates a holistic examination of both verbal and non-verbal interactions within learning contexts.

Gaze interactions, as an important part of non-verbal interactions, can signal the intent to speak, show agreement or disagreement, or indicate confusion and the need for help or feedback, which can then be addressed through verbal interaction (Zhou et al., 2023). Schneider et al. (2018) pioneered an investigative approach that amalgamated video and eye-tracking data to explore gaze synchrony during collaborative problem-solving. The premise being that concurrent gaze patterns can signify shared attention, a fundamental element for effective SSRL. Recently, Zhou et al. (2023) distinguished two unique gaze pattern types that corresponded with varying degrees of mutual comprehension and satisfaction in collaboration. The first, termed peer-interaction focused, emphasizes the social facets of collaboration, while the second, known as resource-interaction focused, places importance on resource allocation and task completion. Although this significance of gaze interactions has been established in general communication (Purwati et al., 2019; Westland, 2015) and collaborative learning (Schneider & Pea, 2017; Zhou et al., 2023), their specific role in SSRL has been relatively underexamined.

3. Methods

3.1. Participants and Procedures

The study involved high school students (N=24) who were assigned to 8 small groups. Three students in a group were required to engage in a collaborative learning task, design a healthy breakfast smoothie while considering nutritional requirements. Each student had access to their own laptop while a group shared an online document to collectively design the smoothie recipe. The whole collaborative task lasted for 30–40 minutes. Following the completion of the first half-time of the learning task, a cognitive regulation–triggering event was introduced to the groups, followed by two emotional regulation–triggering events every three minutes. The data collection obtained ethics approval from the Ethics Committee of Human Sciences at the University of Oulu.

Cognitive and emotional regulation–triggering events were developed to explore the impact of external factors on co-operative learning and the reactions of learners to these regulation-triggering events in a collaborative setting. The development of the regulation-triggering events followed the definitions and examples provided by Järvelä and Hadwin (2024). Specifically, the cognitive regulation–triggering event was designed to escalate task complexity, i.e., incorporating a customer's allergies to latex protein and dairy products, to require additional applying, elaborating, and connecting ideas. Conversely, the emotional regulation–triggering event was crafted to induce negative emotions by intensifying external pressures from others through unpleasant voice tones. In this study, emotional regulation–triggering events included a customer expressing impatience in an angry voice, rushing the completion of a task.

3.2. Qualitative Video Analysis

A thorough qualitative analysis of video data was carried out to determine various verbal and gaze interaction types for SSRL based on their characteristics. Within the theoretical framework of SSRL, to discern different interaction types related to regulation, a coding scheme from previous research was utilized (Isohätälä et al., 2020; Nguyen, Järvelä, et al., 2023). Four categories were incorporated: 1) metacognitive interaction, 2) cognitive interaction, 3) socio-emotional interaction, and 4) task execution interaction.

Gaze interactions were coded using the ELAN annotation software (Crasborn et al., 2006). Segments of gaze interaction were defined as the intervals between each change of gaze direction toward a point of interest. These segments were then labelled for their corresponding directions, namely, laptop, peer, and other. The code, *laptop*, refers to a learner looking at their laptop; *peer*, refers to a learner looking at one of their peers, and the code, *other*, refers to when a learner looked somewhere other than their laptop or their peers. This approach enabled the gain of data points related to group-level gaze interactions to a fine-grain detail without the need for using eye tracking devices. Table 1 shows the coding scheme with examples for video quantitative analysis.

Table 1. Coding Scheme for Video Quantitative Analysis

Code	Description	Example
Regulatory Characteristics of Verbal Interactions		
Metacognitive	Meta-level mental processes toward the control and monitoring of cognitive and emotional activities (orienting, planning, monitoring, evaluating, and regulating). The connection and reflection are aimed at task-related strategies, group processes, or dynamics.	*S1: By the way, I don't use this ingredient page at all, I just put it in there and see what happens.* *S2: If we just keep the ingredients the same, but increase their number in the same ratio, so then those percentages go absolutely nowhere.*
Cognitive	Interaction focuses on higher-order learning-related thinking skills such as understanding, analyzing, reasoning, and evaluating at the object-level related to task content.	*S2: Well, here are the others, here are all the chia seeds, hazelnut spread, whey protein powder.* *S1: But here would be pineapple or blueberry, then they would be the kind where there would be very little of everything.*
Socio-emotional	Action and interaction relevant to the expression of one's emotion in social contexts with clear negative/positive affect nature (e.g., showing gratitude, joking, disputing).	*S1: Oh good time, this guy first orders a smoothie, and then tells us to make it again and then complains that it's taking us a long time to make it. [Express annoyance with group shows shared feeling]*
Task Execution	Actions and interactions that primarily focus on carrying out task requirements, and completing the task, including, e.g., typing on the computer, reading the instructions.	*S1: Yeah, I'll change them to one hundred and twenty-five. [Inform current process]* *S2: One hundred and twenty-five. OK that should be twenty-five then.*
Gaze Interactions		
Peer	Participants gazing in the direction of one of their peers.	
Laptop	Participants gazing in the direction of their laptop.	
Other	Participants looking in any other direction than PEER or LAPTOP.	

In total, 3,523 interactions were coded, consisting of 1,551 verbal interactions ($f_{task\_execution} = 501$, $f_{metacognitive} = 505$, $f_{cognitive} = 456$, $f_{socio\text{-}emotional} = 89$) and 1,972 gaze interactions ($f_{laptop} = 780$, $f_{peer} = 884$, $f_{other} = 308$). To ensure the reliability of the qualitative coding, reliability testing involved an additional independent coder for 1,265 interactions (36% of the dataset), with high inter-reliability achieved (Cohen's Kappa = 0.88).

3.3. Ordered Network Analysis (ONA) and Transmodal Analysis (TMA)

Ordered network analysis (ONA) is a method to model ordered relationships of learning events unfolding over a complex learning process (Tan et al., 2023). As a unified method, ONA provides a visualization of connection patterns with underlying mathematical representations and statistical warrants. Due to its unique affordances, ONA has been widely applied in learning analytics to explore the patterns of different learning processes such as collaborative learning (Kang et al., 2024; Yan et al., 2023), self-regulated learning (Zhang et al., 2024), problem solving (Ruis et al., 2023), etc. For analytical procedures, ONA represents ordered relationships among binary codes for learning events within their recent temporal contexts using a fixed-length window. The window includes each learning event as a referring line and the previous lines within the recent temporal contexts. Based on the binary codes, ONA calculates ordered connections based on the co-occurances of codes between the referring line and its previous lines. That is, within a range of time, if there is a connection made between one event to its previous events, the connection between two corresponding codes is recorded as 1; otherwise, there is no connection made, recorded as 0. Then, ONA accumulates ordered connections for all windows for each unit. These ordered connections are represented in a high-dimensional space. To visualize the connection patterns, ONA can reduce the high-dimensional space to a two-dimensional space using dimensional reduction techniques, such as singular value decomposition and means rotations. As a result of ONA, each unit has a dual representation: 1) a pair of ONA scores in a two-dimensional space, and 2) an ordered network with nodes, bi-directional line weights, and network centoids in the co-registered space. Post-hoc statistical tests can be applied based on the ONA scores and line weights of each unit to compare statistical differences among learner groups, learning phases, or other metrics of comparison. Additionally, network representation provides insights based on different components in the visualizations. We adopt the original demonstration by Tan et al. (2023, p. 108) to explain the interpretation of an ONA plot from the following three components:

1. **Node positions**: The locations of nodes provide an interpretation on each axis in the two-dimensional space. That is, units located on the negative side of a dimension, tend to make more connections with codes also located on the negative side of the same dimension. For example, the small red dot in the bottom left quadrant in Figure 1 is the ONA score of a unit that makes more connections between code A and other codes.
2. **Node sizes**: The radius of inner coloured circles for nodes indicates the strength of self-references of the corresponding codes. That is, a larger inner radius of a code indicates that it occurs repeatedly within the temporal window. The radium of outer circles for nodes represents the total ingree of a code. That is, a larger outer radius of a node represents that the code is a common response to other codes within the corresponding temporal window. For example, code A has the strongest self-reference and the most common response to other codes.
3. **Bi-directional line weights**: Between any two codes, two triangles indicate the bi-directional connection strengths between them. The grey chevron indicates which direction has stronger connections. For example, in Figure 1, there are more connections from code B to code A.

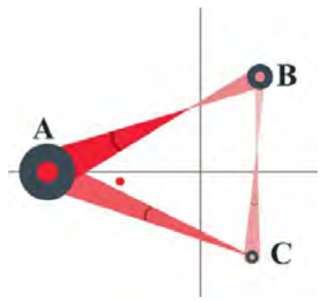

Figure 1. Ordered network analysis interpretation.

However, as described by Wang et al. (2023), ONA fails to account for diverging temporal relationships in complex learning contexts, especially when involving heterogenous data collected from different modalities. Thus, Wang and colleagues applied Transmodal Analysis (TMA) to augment ONA to resolve this methodological challenge. TMA is a conceptual and methodological framework to model human activities and processes across multiple modalities by augmenting existing state-dependent models (Shaffer et al., 2025). Instead of fusing data collected from different modalities, TMA represents the interdependence between multimodal events by specifying mathematical functions, such as temporal influence functions (TIFs), horizon of observational functions, and learner difference functions. Among the three mathematical functions, a TIF specifies how long a learning event of a modality can continue to impact future learning events. Due to incompatible data size across different modalities, TMA allows multiple TIFs to specify unique temporal influences. For example, *step functions* with different reference points can represent different sizes of fixed windows. In this paper, we applied a combination of TMA and ONA, termed as T/ONA, to analyze patterns of verbal and gaze interactions through social regulation. That is, T/ONA adopts the affordances of ONA for graphical visualizations and statistical warrants; at the same time, T/ONA calculates and accumulates connections, making for a finer granularity to account for the temporal differences of different modalities.

As an emerging method in multimodal learning analytics, this is one of the early empirical studies that apply TMA to investigate the learning processes by engaging with multimodal data. Shaffer et al. (2025) first demonstrated the benefits of TMA, higher explanatory power, and higher goodness of fit in the context of a virtual internship for college engineering students. Additionally, Borchers et al. (2024) used T/ONA to demonstrate the role of teachers in an AI-supported classroom based on data collected from human–computer interaction and human–human interaction. According to the preliminary results, T/ONA is a promising method for understanding complex learning phenomena that engage with multimodal data. Thus, we applied T/ONA in this study to examine social interactions and regulations through verbal and gaze behaviours. T/ONA was conducted in this study using the R programming environment. The R package utilized for this analysis is available upon request by contacting the authors.

3.3.1. Model Configuration

To explore the patterns of connection-making for individuals when given different triggers for social regulation, we defined each individual student at each phase as the smallest unit of analysis. To investigate the verbal and non-verbal interactions during the collaborative task, we used T/ONA to represent ordered connections among the following four codes derived from discourse and three codes detected based on gaze fixation: *cognitive, task execution, metacognitive, socio-emotional, peers, laptop,* and *other* (see details in Table 1). Additionally, to explain the most variance between phases before receiving any regulation-triggering events and the phase after receiving all triggering events, we constructed T/ONA using means rotation

ISSN 1929-7750 (online).

of the first phase and the last phase. With input data collected from two modalities — dialogue and eye gaze — we specified different TIFs respectively. Based on a deep qualitative analysis of video data, we identified that the recent temporal contexts for a dialogue behaviour and a gaze behaviour are approximately 30 seconds and 10 seconds, respectively. That is, any dialogue event makes connections with learning events within 30 seconds. As gazing switches focus more rapidly compared to dialogue, we specified the TIF for gaze as 10 seconds to make connections with events and 10 seconds after.

3.3.2. Post-Hoc Statistical Tests

T/ONA adopts the dual representations given a unit as ONA; ONA scores detailed network structures with nodes and edges. To investigate the differences in connection making when introducing different social-regulation triggers, we conducted a statistical test on ONA scores for individual students across four phases. Due to the violation of normality and homogenous variance using a Shapiro-Wilk test (p = 0.160) and a Bartlett's test (p = 0.07), we conducted a non-parametric test — the Friedman test — to compare differences in ONA scores between phases. To control for type I errors due to multiple comparisons across groups, we further conducted a pairwise Wilcoxon signed rank test across all phases with Benjamini-Hochberg correction to compare differences between any adjacent phases. While the Friedman test reveals the significant changes across different phases at an individual level, we conducted further analysis to investigate which specific connections contributed to the differences at the individual level. Similarly, due to the violation of assumptions, we conducted a paired Wilcoxon signed rank test across all phases. Then, we interpreted significant differences in connection-making for any two adjacent phases along with the ONA plots to provide statistical warrants for the visualized differences.

4. Results and Findings

4.1. RQ1. How do Gaze Interactions Complement or Interact with Verbal Interactions to Facilitate SSRL in a Face-to-Face Collaborative Learning Environment?

Figure 2 provides a qualitative example showcasing verbal and gaze interactions following a cognitive regulation–triggering event. In the figure, the pink, green, and yellow columns represent the gaze points of the individual learners within a group, with their dialogues located to the right of these columns. Each box within the figure corresponds to a specific point of gaze or dialogue, with the dialogue type provided in parentheses.

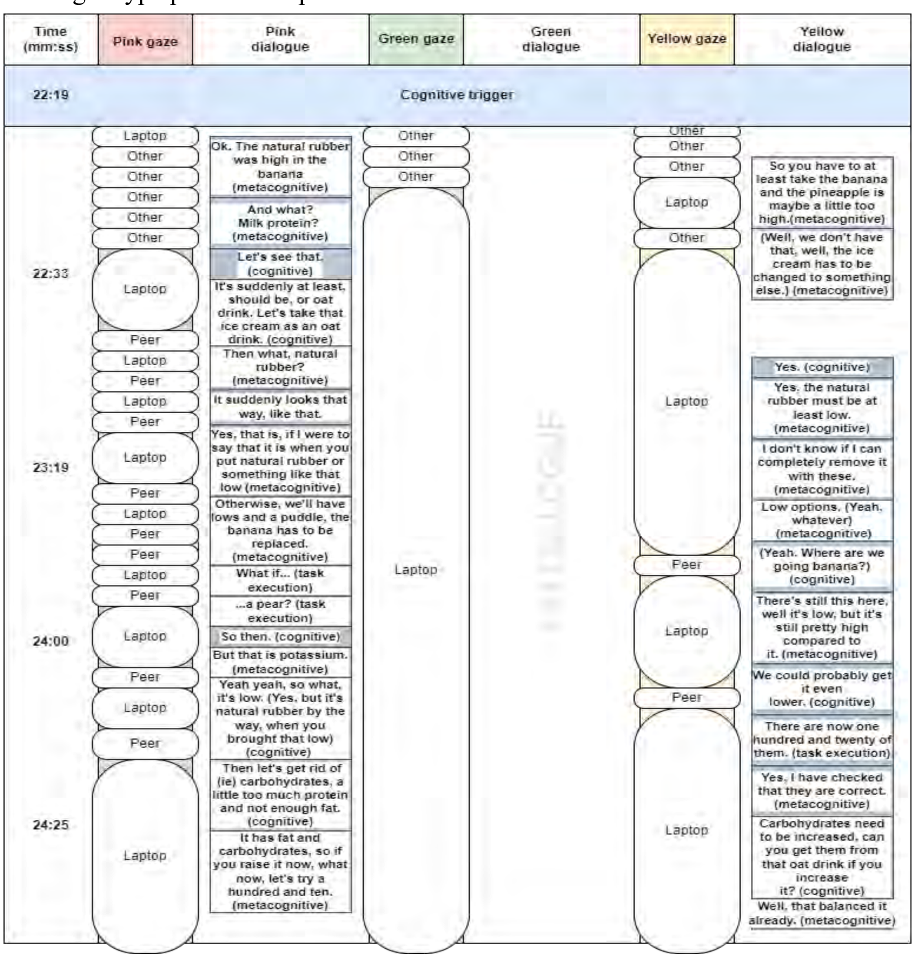

Figure 2. Gaze and dialogue in collaborative learning after a cognitive regulation–triggering event.

This qualitative example also illustrates the premise that the targeted intervention had a marked effect on learner interactions, specifically in the aspect of metacognition. Upon an initial glance, it becomes evident that there is a significant number of interactions between the pink and yellow learners, whereas the green learner does not engage in interactions, and is looking at their laptop. Nevertheless, pink and yellow learners are trying to navigate the updated requirements and associated issues with their current task. Verbal interactions concerning the task are linked to both gazing at the laptop and engaging with peers through eye contact. Apart from resource-focused interactions with the laptop, peer-gazing behaviour reflects the aim of cultivating a shared understanding among group members. Yet, in the observed scenario, SSRL was conducted by a subset of the group, not by all members, pointing to variations in engagement and strategic approaches within SSRL that require further study.

In Figure 3, we plotted a grand mean T/ONA network to show 1) general patterns of connection-making for all four phases (the grey network), and 2) trends of changes in connection-making across four phases based on the mean and standard deviation of ONA scores for each phase (coloured confidence intervals for phases 1, 2, 3, and 4). According to node position, *cognitive* and *task execution* interactions are located on the most negative side while *metacognitive* interactions are located on the most positive side along the MR dimension. According to the co-registration of ONA, the location of plotted ONA scores shows general patterns of connection-making. For example, units with low ONA scores on the MR dimension, located on the negative side, tend to make more connections with *cognitive* interaction and *task execution*; units plotted on the most positive side make more connections with *metacognitive* interactions. Additionally, gazing objects (*peer, laptop, other* objects) and *socio-emotional* locates in the middle of the plot, indicating that these codes are likely to make connections with codes on both the positive and negative sides.

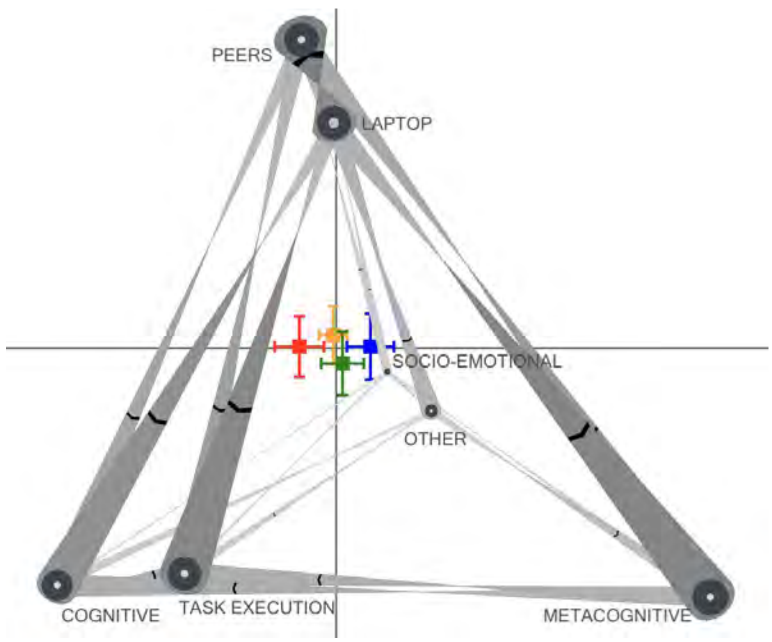

Figure 3. Grand mean ONA network.

Based on the trend along the MR dimension, we discovered the general patterns of learning progression as receiving more cognitive and emotional regulation–triggering events: students focused on *cognitive* interaction and *task execution* at the beginning (phase 1 indicated by red) and then made connections with social-emotional and gaze behaviours in the middle (phases 2 and 3 indicated by yellow and green), and eventually engaged with *metacognitive* interactions at the end of the task (phase 4 indicated by blue). Even though without a significant difference across phases according to a Friedman test ($\chi^2(23)$ = 22.620, p = 0.483), we can identify this trajectory visually from the plot: the mean of ONA scores increases from the most negative side to the post positive side on the MR dimension. We further conducted a pairwise Wilcoxon signed rank test across all phases with Benjamini-Hochberg corrections. While there are no significant differences in connection-making before and after receiving a cognitive regulation–triggering event (W = 84, p = 0.073, r = 0.440) and before and after receiving the first emotional regulation–triggering event (W = 116, p = 0.345, r = 0.230), there is a significant difference before and after receiving the second emotional regulation–triggering event (W = 75, p = 0.047*, r = 0.500). Thus, individual Wilcoxon signed rank tests suggest that the last emotional regulation–triggering event has a strong effect in changing connection-making from *cognitive* focused to *metacognitive* focused interactions.

Throughout the session, students made strong connections between verbal and gaze interactions. For example, there are three common responses after non-verbal interactions of gazing at their *laptop: cognitive, metacognitive*, and *task execution* interactions. After students gazed at the task interface on their *laptop*, they then immediately engaged in *metacognitve* and *cognitive* interactions. Additionally, reasoning through *cognitive* or *metacognitive* aspects and *task execution* were common responses to gazing toward *peers*. This indicates that students tend to make eye contact with their peers before starting a conversation about *cognitive/metacognitive* reasoning or *task execution*. This points toward using gaze to start conversations/indicate readiness to interact. Furthermore, there are strong bi-directional connections between gazing toward *peers* and *laptop* screens, indicating the interweaving of social interactions with peers and cognitive processing on their screens.

4.2. RQ2. How do Different Types of Regulation-Triggering Events Impact the Interplays Between Verbal and Gaze Interactions?

4.2.1. a. How do the Interplays Between Verbal and Gaze Interactions Change After the Cognitive Regulation–Triggering Event?

We further uncovered patterns of connection-making in a finer granularity. To measure the differences in connection strength, we conducted individual Wilcoxon signed rank tests on ONA line weights for any adjacent phases. Furthermore, to compare different connection-making patterns, a subtracted plot was produced by subtracting the mean line weights of one group from those of another group. In a subtracted plot, the colour of the edge shows corresponding strong connections as a comparison between two groups. In Figure 4, we constructed a subtracted plot to compare connection-making before (indicated by red) and after (indicated by yellow) receiving a cognitive regulation–triggering event.

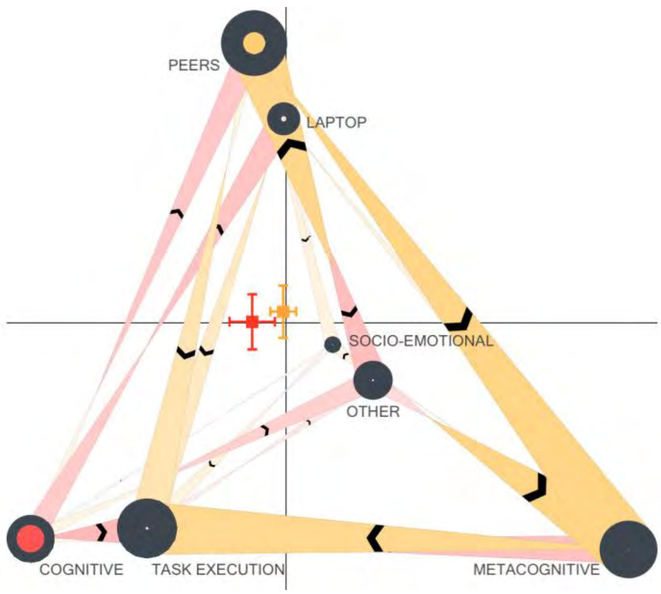

Figure 4. Comparison before and after the cognitive regulation–triggering event.

For connections made before receiving a cognitive regulation–triggering event, students made significantly more connections from *cognitive* interactions to gazing at their *laptop* ($W = 231$, $p = 0.019$*, $r = 0.540$), *peers* ($W = 219$, $p = 0.049$*, $r = 0.460$), and *other* ($W = 163$, $p = 0.007$**, $r = 0.716$). That is, in the initial stage, students tended to first provide *cognitive* interactions then followed by gaze interactions with their peers or looking at the task interface on their screen. However, after receiving the cognitive regulation–triggering event, students made significantly more connections with *metacognitive* interactions. For example, students made significantly more connections from *metacognitive* interactions to *task execution* after receiving the cognitive regulation–triggering event ($W = 61$, $p = 0.035$*, $r = 0.518$). This connection indicates that students switched strategies in their tasks by focusing on monitoring and planning before executing a specific task, which is aligned with the learning objective of designing cognitive regulation–triggering events.

Additionally, *metacognitive* interactions are common responses to non-verbal cues, such as gazing at their *peers* ($W = 70$, $p = 0.021$*, $r = 0.533$) and *other* objects in the learning environment ($W = 15$, $p < 0.001$***, $r = 0.900$). These two connections indicate that students tended to check intercomprehension with other peers via eye contact or pausing verbal interactions by gazing at other objects before providing *metacognitive* interactions. Furthermore, after receiving a cognitive regulation–triggering event, there are significantly more self-transitions of gazing at *peers* ($W = 40$, $p < 0.001$***, 0.733), which suggested that students tended to make more gazing interactions with their peers as prompted by the cognitive regulation–triggering event.

4.2.2. b. How do the Interplays Between Verbal and Gaze Interactions Change After the Emotional Regulation–Triggering Event?

In Figure 5, we constructed a subtracted plot to compare connection-making after receiving a cognitive regulation–triggering event (indicated by yellow) and after receiving the first emotional regulation–triggering event (indicated by green). According to the plot and statistical testing on line weights, there are significantly more connections from *peers* to *cognitive* interactions after receiving a cognitive trigger ($W = 208$, $p = 0.035^*$, $r = 0.507$). That is, students tend to first make gaze interactions with their *peers* before their *cognitive* interactions. However, after receiving an emotional regulation–triggering event, students made significantly more connections that are *socio-emotional* and *metacognitive* focused. For example, there are significantly more connections from *task execution* to both *socio-emotional* ($W = 1$, $p = 0.035^*$, $r = 0.929$) and *metacognitive* ($W = 33$, $p = 0.004^{**}$, $r = 0.714$). That is, after being prompted by the first emotional regulation–triggering event, students focused on *socio-emotional* talks and *metacognitive* interactions after *executing tasks*. Additionally, there are significantly more self-transitions within *metacognitive* interactions ($W = 77$, $p = 0.037^*$, $r = 0.487$). That is, students were engaged in a deep discussion of planning and evaluating their task given an emotional regulation–triggering event.

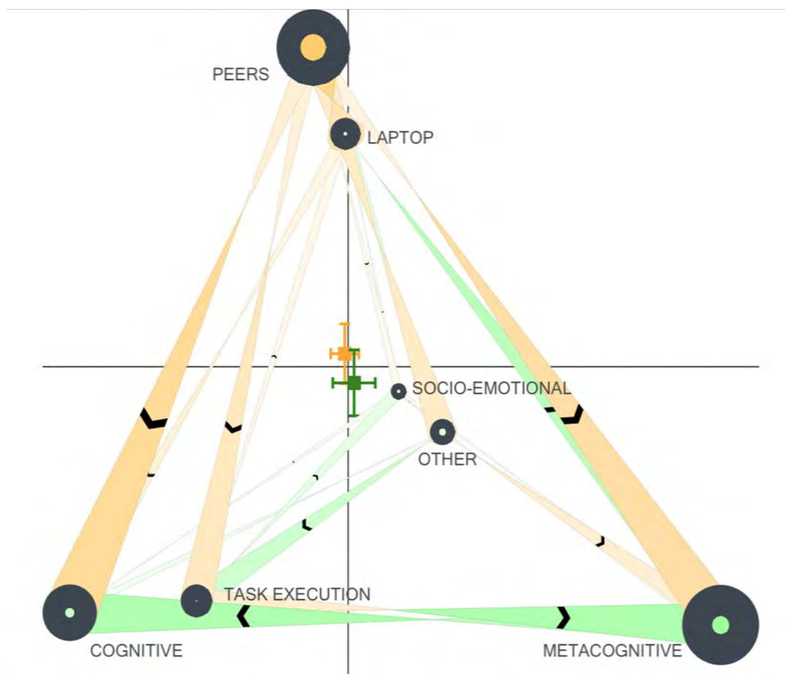

Figure 5. Comparison before and after the first emotional regulation–triggering event.

In Figure 6, we constructed a subtracted network before and after the repeated emotional regulation–triggering event. Before receiving the second triggering event, there is a significantly stronger connection from *task execution* to *cognitive* interactions ($W = 197$, $p = 0.023^*$, $r = 0.577$). That is, students tended to focus on *task execution* based on the prompt of the first emotional trigger and then explaining the *cognitive* interactions of their execution. However, after receiving the second emotional trigger, there are significantly more connections from verbal *socio-emotional* cues to eye-contact with *peers* ($W = 29$, $p = 0.015^*$, $r = 0.661$), which indicates the focus of social interactions between verbal and gazing interactions. Additionally, there is significantly more connection from *socio-emotional* cues to gazing at the task interface on their *laptop* ($W = 43$, 0.038^*, $r = 0.547$), suggesting that students focused on checking the tasks on their *laptop* after engaging in *socio-emotional* interactions. The act of establishing eye contact after giving verbal socio-emotional cues can be perceived as a non-verbal communication strategy to reinforce or verify the emotions or sentiments expressed verbally (Westland, 2015). Students, when confronted with emotional regulation–triggering events, may inherently lean toward seeking social support or validation from their peers.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License (CC BY 4.0)

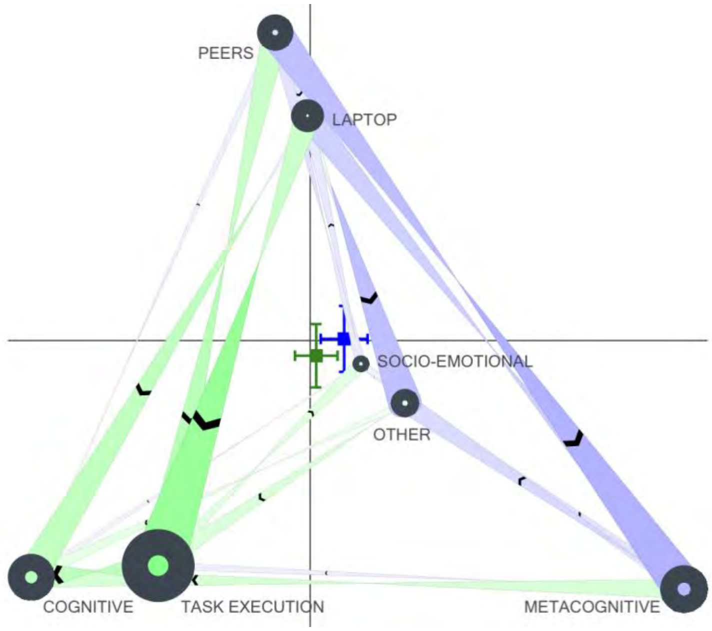

Figure 6. Comparison before and after the second emotional regulation–triggering event.

5. Discussion and Implications

In this study, the use of the innovative method of TMA allowed for the mapping and analysis of communication patterns, thus revealing the importance of both verbal and gaze interactions in SSRL. It can be particularly useful when studying phenomena that have both quantitative and qualitative dimensions, as it allows for building data to conduct more situated and interpretive analyses as well as applying statistical methods to analyze patterns and trends.

The methodological contributions of this study are showcased through the effective use of TMA, which addresses the inherent challenges of capturing the temporal and dynamic aspects of SSRL. In collaborative learning, learners engage in SSRL by continuously interacting with each other and their environment, recalibrating their strategies based on feedback, jointly setting and adjusting goals, and monitoring their collective progress (Hadwin, 2021; Molenaar et al., 2012). The complexity here arises from the interconnectedness of these processes: each interaction or adjustment can significantly influence the others, culminating in a web of regulatory activities that are in a state of constant flux. Given that SSRL processes evolve over time in context, it is crucial to track these changes throughout the duration of collaborative tasks. Traditional methods based on frequency might fall short in accurately representing the complexities of SSRL. For instance, TMA allows for the incorporation of a variety of data sources, such as speech, gestures, gaze, and digital interactions. In socially shared regulation, learners regulate their collective processes through a combination of verbal and non-verbal cues, task-based actions, and interaction with digital tools. While prior methods such as epistemic network analysis (ENA) primarily focuses on discourse and cognitive connections between ideas, transmodal analysis expands this by including multiple modes of interaction, making it more suitable for studying complex, embodied regulation processes. In this study, gaze interactions can play a critical role in SSRL, signalling attention shifts, agreement, or disagreement, which may not be captured through verbal interactions alone. TMA thus provides a richer dataset for understanding how learners jointly monitor, evaluate, and adjust their collective learning strategies.

Drawing on trigger events to frame sampling and analysis about SSRL and regulatory activities recognizes that change over time in context matters: regulatory acts are not just random patterns or dominant sequences of actions but are also intentional strategic responses to situations that demand action or remediation; regulation is not what happens most frequently or all the time but rather what is activated when a situation demands intervention; and regulation involves the dynamic interplay of cognition, motivation, and emotion within and across regulatory activities (Järvelä & Hadwin, 2024; Winne, 2022). Therefore, previous studies have advocated for new approaches that incorporate data from various modalities to more effectively investigate SSRL (Azevedo & Gašević, 2019; Järvelä & Bannert, 2021). Building on this notion, our study showcases how the innovative TMA techniques provide a thorough analysis of SSRL, incorporating data from both verbal and gaze interactions.

In line with previous research (e.g., Ertl et al., 2006; Isohätälä et al., 2020; Vuorenmaa et al., 2023), we found that the interactions were instrumental in navigating the cognitive and emotional challenges posed by the regulation-triggering events. For instance, when faced with task-related challenges, group members used verbal communication to clarify shared understandings, seek collective solutions, and adjust their strategies. This aligns with previous research emphasizing the importance of social interactions in SSRL (Isohätälä et al., 2020). Nevertheless, our study extends this understanding by highlighting the ways in which the interplays between verbal and gaze interactions contribute to the regulation of learning processes in response to cognitive and emotional regulation–triggering events, where the gaze interactions used in conjunction with verbal interactions help to construct meaning.

Our examination of gaze interactions also revealed their significant role in conveying cognitive and emotional information, which often complements or enhances verbal exchange. This aligns with findings from prior research, which indicate that eye contact is a primary non-verbal means of establishing shared understanding, and mutual acknowledgment in various social contexts (Purwati et al., 2019). In the context of SSRL, our study shows that after being presented with a cognitive regulation–triggering event, students often sought mutual understanding with their classmates by making eye contact or momentarily halted their thought process by looking at other items around them before offering metacognitive interactions.

Our results significantly emphasized the role of emotional regulation–triggering events in shaping learner interactions for SSRL. The observed shifts in student behavioural patterns, especially in their approach to forming bonds via socio-emotional interactions following an event that triggers emotional regulation, further highlight the central role of emotions in guiding shared learning. The influence of emotions on learning has been a longstanding area of interest in educational research (Boekaerts & Pekrun, 2016; Nguyen, Li, et al., 2023). Moreover, students engage in more in-depth discussions about planning and assessing their tasks when faced with an emotional regulation–triggering event.

The insights drawn from this study have several implications. For teachers, gaining a comprehensive understanding of SSRL dynamics through research utilizing our approach can empower them to design collaborative tasks that account for potential regulation-triggering events. By anticipating these events, they can scaffold learning experiences to either harness or mitigate their impact. In terms of technology, there is an opportunity for educational technology developers to design tools that track and give immediate feedback on group dynamics, particularly in authentic collaborative learning settings. Such tools or applications can assess both verbal and non-verbal interactions, presenting recommendations and action points to optimize SSRL processes in everyday classrooms.

6. Conclusion, Limitations, and Future Directions

Our study highlights the critical role of both verbal and non-verbal interactions, particularly gaze behaviours, for SSRL in collaborative learning settings with computer-mediated materials. In particular, the study features the complex dynamics and adaptive nature of SSRL in response to regulation-triggering events. This study also demonstrates an innovative analytics method with TMA to address both the nature of temporality and multimodality of interactions for SSRL. Our study reveals how learners adapt, recalibrate, and negotiate their shared understanding in real-time, especially when confronted with regulation-triggering events. By recognizing and understanding these nuances, educators and researchers can better examine and support SSRL in collaborative learning.

However, these findings must be contextualized within the study's limitations. The relatively small sample size might impact the generalizability of our results. Despite these limitations, this research makes a significant contribution to the SSRL field, particularly regarding the integration of verbal and gaze interactions in studying these phenomena.

Our study also raises several questions for future research. For instance, how does the design of different regulation-triggering events influence the interplay between verbal and non-verbal interactions? How can we effectively support learners in navigating these triggering events and enhancing their SSRL processes? These questions warrant further investigation to deepen our understanding of SSRL and inform the design of effective instructional strategies and support systems for collaborative learning. Future research could explore how different regulation-triggering events influence learner regulation processes and how these triggers interact with verbal and non-verbal interaction patterns. Such an approach can provide a more holistic picture of SSRL processes. Subsequent studies could expand the application of T/ONA methods to examine SRL at the individual level. Moreover, future research could extend this approach to develop automated tools for teacher use in classrooms, integrating TMA algorithms and AI methods to automate the coding and interpretation of interactions. Additionally, further investigation into ethics and privacy concerns associated with such technologies would be essential to ensure responsible and equitable implementation. We hope that our findings will inspire further research in this area, contributing to the ongoing efforts to enhance collaborative learning experiences.

Declaration of Conflicting Interest

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

Funding

The publication of this article received financial support from the Research Council of Finland, grant numbers 324381 and 350249, and the University of Oulu profiling project Profi7 Hybrid Intelligence – 352788.

Acknowledgments

This work was carried out with the support of LeaF Research Infrastructure, University of Oulu, Finland.

References

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

Bakhtiar, A., Webster, E. A., & Hadwin, A. F. (2018). Regulation and socio-emotional interactions in a positive and a negative group climate. *Metacognition and Learning, 13*(1), 57–90. https://doi.org/10.1007/s11409-017-9178-x

Boekaerts, M., & Pekrun, R. (2016). Emotions and emotion regulation in academic settings. In L. Corno & E. M. Anderman (Eds.), *Handbook of educational psychology* (3rd ed., pp. 76–90). Routledge. https://doi.org/10.4324/9781315688244

Borchers, C., Wang, Y., Karumbaiah, S., Ashiq, M., Shaffer, D. W., & Aleven, V. (2024). Revealing networks: Understanding effective teacher practices in AI-supported classrooms using transmodal ordered network analysis. *Proceedings of the 14th Learning Analytics and Knowledge Conference* (LAK '24), 18–22 March 2024, Kyoto, Japan (pp. 371–381). ACM Press. https://doi.org/10.1145/3636555.3636892

Crasborn, O., Sloetjes, H., Auer, E., & Wittenburg, P. (2006). Combining video and numeric data in the analysis of sign languages with the ELAN annotation software. In C. Vettori (Ed.), *Proceedings of the LREC2006 2nd Workshop on the Representation and Processing of Sign Languages: Lexicographic Matters and Didactic Scenarios* (LREC2006), 28 May 2006, Genoa, Italy (pp. 82–87). European Language Resources Association (ELRA).

Dang, B., Nguyen, A., & Järvelä, S. (2024). The unspoken aspect of socially shared regulation in collaborative learning: AI-driven learning analytics unveiling "silent pauses." *Proceedings of the 14th Learning Analytics and Knowledge Conference* (LAK '24), 18–22 March 2024, Kyoto, Japan (pp. 231–240). ACM Press. https://doi.org/10.1145/3636555.3636874

De Backer, L., Van Keer, H., De Smedt, F., Merchie, E., & Valcke, M. (2022). Identifying regulation profiles during computer-supported collaborative learning and examining their relation with students' performance, motivation, and self-efficacy for learning. *Computers & Education, 179*, 104421. https://doi.org/10.1016/j.compedu.2021.104421

Ertl, B., Fischer, F., & Mandl, H. (2006). Conceptual and socio-cognitive support for collaborative learning in videoconferencing environments. *Computers & Education, 47*(3), 298–315. https://doi.org/10.1016/j.compedu.2004.11.001

Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernandez-Leo, D., Järvelä, S., Mavrikis, M., & Rienties, B. (2024). The promise and challenges of generative AI in education. *Behaviour & Information Technology*. https://doi.org/10.1080/0144929X.2024.2394886

Hadwin, A. F. (2021). Commentary and future directions: What can multi-modal data reveal about temporal and adaptive processes in self-regulated learning? *Learning and Instruction, 72*, 101287. https://doi.org/10.1016/j.learninstruc.2019.101287

Hadwin, A., Järvelä, S., & Miller, M. (2018). Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 83–106). Routledge.

Huang, X., & Lajoie, S. P. (2023). Social emotional interaction in collaborative learning: Why it matters and how can we measure it? *Social Sciences & Humanities Open, 7*(1), 100447. https://doi.org/10.1016/j.ssaho.2023.100447

Isohätälä, J., Näykki, P., & Järvelä, S. (2020). Cognitive and socio-emotional interaction in collaborative learning: Exploring fluctuations in students' participation. *Scandinavian Journal of Educational Research, 64*(6), 831–851. https://doi.org/10.1080/00313831.2019.1623310

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. (2024). Triggers for self-regulated learning: A conceptual framework for advancing multimodal research about SRL. *Learning and Individual Differences, 115*, 102526. https://doi.org/10.1016/j.lindif.2024.102526

Järvelä, S., Nguyen, A., & Hadwin, A. (2023). Human and artificial intelligence collaboration for socially shared regulation in learning. *British Journal of Educational Technology, 54*(5), 1057–1076. https://doi.org/10.1111/bjet.13325

Kang, J., Zhou, Y., Rajarathinam, R. J., Tan, Y., & Shaffer, D. W. (2024). Unveiling joint attention dynamics: Examining multimodal engagement in an immersive collaborative astronomy simulation. *Computers & Education, 213*, 105002. https://doi.org/10.1016/j.compedu.2024.105002

Molenaar, I., de Mooij, S., Azevedo, R., Bannert, M., Järvelä, S., & Gašević, D. (2023). Measuring self-regulated learning and the role of AI: Five years of research using multimodal multichannel data. *Computers in Human Behavior, 139*, 107540. https://doi.org/10.1016/j.chb.2022.107540

Molenaar, I., Roda, C., van Boxtel, C., & Sleegers, P. (2012). Dynamic scaffolding of socially regulated learning in a computer-based learning environment. *Computers & Education, 59*(2), 515–523. https://doi.org/10.1016/j.compedu.2011.12.006

Nguyen, A., Hong, Y., Dang, B., & Nguyen, P. T. B. (2022). Emotional regulation in synchronous online collaborative learning: A facial expression recognition study. *ICIS 2022 Proceedings*, 9–14 December 2022, Copenhagen, Denmark. Association for Information Systems. https://aisel.aisnet.org/icis2022/learning_iscurricula/learning_iscurricula/12

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

Nguyen, A., Li, Y., Järvelä, S., Liu, Y., & Zhao, G. (2023). Emotional synchrony through regulatory triggers in collaborative learning: A facial expression recognition study. *Proceedings of the 17th International Conference of the Learning Sciences* (ICLS 2023), 10–15 June 2023, Montreal, QC, Canada (pp. 1038–1041). International Society of the Learning Sciences. https://doi.org/10.22318/icls2023.547768

Purwati, P., Rahmiati, R., Asra, S., & Fadlia, F. (2019). The impact of non-verbal communication in the learning process. *Journal of Education, Linguistics, Literature and Language Teaching, 2*(2), 1–12.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science, 29*(6), 1045–1060. https://doi.org/10.1207/s15516709cog0000_29

Ruis, A. R., Tan, Y., Brohinsky, J., Yang, B., Wang, Y., Cai, Z., & Shaffer, D. W. (2023). Thin data, thick description: Modeling socio-environmental problem-solving trajectories in localized land-use simulations. In G. Arastoopour Irgens & S. Knight (Eds.), *Advances in quantitative ethnography: 5th international conference, ICQE 2023, Melbourne, VIC, Australia, October 8–12, 2023, proceedings* (pp. 349–364). Springer Cham. https://doi.org/10.1007/978-3-031-47014-1_24

Schneider, B., & Pea, R. (2017). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. In M. Orey & R. M. Branch (Eds.), *Educational media and technology yearbook* (Vol. 40, pp. 99–125). Springer Cham. https://doi.org/10.1007/978-3-319-45001-8_7

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (2018). Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. *International Journal of Computer-Supported Collaborative Learning, 13*(3), 241–261. https://doi.org/10.1007/s11412-018-9281-2

Shaffer, D. W., Wang, Y., Marquart, C., Ruis, A. R., & Cai, Z. (2025). Transmodal analysis. *Journal of Learning Analytics, 1–22*. https://doi.org/10.18608/jla.2025.8423.

Tan, Y., Ruis, A. R., Marquart, C., Cai, Z., Knowles, M. A., & Shaffer, D. W. (2023). Ordered network analysis. In C. Damşa & A. Barany (Eds.), *Advances in quantitative ethnography: 4th international conference, ICQE 2022, Copenhagen, Denmark, October 15–19, 2022, proceedings* (pp. 101–116). Springer Cham. https://doi.org/10.1007/978-3-031-31726-2_8

Volet, S., Summers, M., & Thurman, J. (2009). High-level co-regulation in collaborative learning: How does it emerge and how is it sustained? *Learning and Instruction, 19*(2), 128–143. https://doi.org/10.1016/j.learninstruc.2008.03.001

Vuorenmaa, E., Järvelä, S., Dindar, M., & Järvenoja, H. (2023). Sequential patterns in social interaction states for regulation in collaborative learning. *Small Group Research, 54*(4), 512–550. https://doi.org/10.1177/10464964221137524

Wang, Y., Shah, M., Jimenez, F. A., Wilson, C., Ashiq, M., Eagan, B., & Shaffer, D. W. (2023). Developing nursing students' practice readiness with Shadow Health® Digital Clinical Experiences™: A transmodal analysis. In G. Arastoopour Irgens & S. Knight (Eds.), *Advances in quantitative ethnography: 5th international conference, ICQE 2023, Melbourne, VIC, Australia, October 8–12, 2023, proceedings* (pp. 365–380). Springer Cham. https://doi.org/10.1007/978-3-031-47014-1_25

Westland, G. (2015). *Verbal and non-verbal communication in psychotherapy.* W. W. Norton.

Winne, P. H. (2022). Modeling self-regulated learning as learners doing learning science: How trace data and learning analytics help develop skills for self-regulated learning. *Metacognition and Learning, 17*(3), 773–791. https://doi.org/10.1007/s11409-022-09305-y

Yan, L., Tan, Y., Swiecki, Z., Gašević, D., Shaffer, D. W., Zhao, L., Li, X., & Martinez-Maldonado, R. (2023). Characterizing individual-level collaborative learning behaviors using ordered network analysis and wearable sensors. In G. Arastoopour Irgens & S. Knight (Eds.), *Advances in quantitative ethnography: 5th international conference, ICQE 2023, Melbourne, VIC, Australia, October 8–12, 2023, proceedings* (pp. 66–80). Springer Cham. https://doi.org/10.1007/978-3-031-47014-1_5

Zhang, J., Borchers, C., & Barany, A. (2024). Studying the interplay of self-regulated learning cycles and scaffolding through ordered network analysis across three tutoring systems. In Y. J. Kim & Z. Swiecki (Eds.), *Advances in quantitative ethnography: 6th international conference, ICQE 2024, Philadelphia, PA, USA, November 3–7, 2024, proceedings, part I* (pp. 231–246). Springer Cham. https://doi.org/10.1007/978-3-031-76335-9_17

Zhou, Q., Bhattacharya, A., Suraworachet, W., Nagahara, H., & Cukurova, M. (2023). Automated detection of students' gaze interactions in collaborative learning videos: A novel approach. In O. Viberg, I. Jivet, P. J. Muñoz-Merino, M. Perifanou, & T. Papathoma (Eds.), *Responsive and sustainable educational futures: 18th European conference on technology enhanced learning, EC-TEL 2023, Aveiro, Portugal, September 4–8, 2023, proceedings* (pp. 504–517). Springer Cham. https://doi.org/10.1007/978-3-031-42682-7_34

Zimmerman, B. J. (2011). Motivational sources and outcomes of self-regulated learning and performance. In D. H. Schunk & B. Zimmerman (Eds.), *Handbook of self-regulation of learning and performance* (pp. 49–64). Routledge. https://doi.org/10.4324/9780203839010