JOURNAL OF LEARNING ANALYTICS

Volume 11(2), 197–214. https://doi.org/10.18608/jla.2024.8287

Learners' Linguistic Alignment and Physiological Synchrony: Identifying Trigger Events that Invite Socially Shared Regulation of Learning

Joni Lämsä[1], Justin Edwards[2], Eetu Haataja[3], Marta Sobocinski[4], Paola R. Peña[5], Andy Nguyen[6] and Sanna Järvelä[7]

Abstract

The theory of socially shared regulation of learning (SSRL) suggests that successful collaborative groups can identify and respond to trigger events stemming from cognitive or emotional obstacles in learning. Thus, to develop real-time support for SSRL, novel metrics are needed to identify different types of trigger events that invite SSRL. Our aim was to apply two metrics derived from different data streams to study how trigger events for SSRL shaped group linguistic alignment (based on audio data) and physiological synchrony (based on electrodermal activity data). The data came from six groups of students ($N = 18$) as they worked face-to-face on a collaborative learning task with one cognitive and two emotional trigger events. We found that the cognitive trigger event increased linguistic alignment in task-description words and led to physiological out-of-synchrony. The emotional trigger events decreased out-of-synchrony and increased high-arousal synchrony at the physiological level but did not affect linguistic alignment. Therefore, different metrics for studying markers and responses to different types of trigger events are needed, suggesting the necessity for multimodal learning analytics to support collaborative learning.

Notes for Practice

- Different types of trigger events (e.g., cognitive or emotional) arise in collaborative learning, requiring regulatory responses within a group.
- Collaboration analytics requires metrics to identify these types of trigger events and group responses to those to provide support with actionable insights.
- To identify cognitive and emotional trigger events that invite socially shared regulation of learning, we applied two metrics: group linguistic alignment (based on audio data) and physiological synchrony (based on electrodermal activity data).
- Cognitive trigger events increased the linguistic alignment of the task-description words among group members.
- Emotional trigger events became visible in the physiological synchrony of the group members.

Keywords: Collaboration analytics, collaborative learning, linguistic alignment, multimodal learning analytics, physiological synchrony, self-regulated learning (SRL), socially shared regulation of learning (SSRL)

Submitted: 06/11/2023 — **Accepted:** 05/06/2024 — **Published:** 04/07/2024

Corresponding author [1]Email: joni.lamsa@oulu.fi Address: Learning and Educational Technology (LET) Research Lab, P.O. Box 2000, FI-90014, University of Oulu, Oulu, Finland. ORCID iD: https://orcid.org/0000-0001-7995-4090
[2]Email: justin.edwards@oulu.fi Address: Learning and Educational Technology (LET) Research Lab, P.O. Box 2000, FI-90014, University of Oulu, Oulu, Finland. ORCID iD: https://orcid.org/0000-0003-1487-9207
[3]Email: eetu.haataja@oulu.fi Address: Research Unit of Population Health (PopH), P.O. Box 5000, FI-90014, University of Oulu, Oulu, Finland. ORCID iD: https://orcid.org/0000-0001-8280-1546
[4]Email: marta.sobocinski@oulu.fi Address: Learning and Educational Technology (LET) Research Lab, P.O. Box 2000, FI-90014, University of Oulu, Oulu, Finland. ORCID iD: https://orcid.org/0000-0001-9591-9113
[5]Email: paola.pena@ucdconnect.ie Address: UCD School of Information and Communication Studies. University College Dublin, Belfield, Dublin 4, Ireland. ORCID iD: https://orcid.org/0000-0002-6371-8074
[6]Email: andy.nguyen@oulu.fi Address: Learning and Educational Technology (LET) Research Lab, P.O. Box 2000, FI-90014, University of Oulu, Oulu, Finland. ORCID iD: https://orcid.org/0000-0002-0759-9656
[7]Email: sanna.jarvela@oulu.fi Address: Learning and Educational Technology (LET) Research Lab, P.O. Box 2000, FI-90014, University of Oulu, Oulu, Finland. ORCID iD: https://orcid.org/0000-0001-6223-3668

1. Introduction

The success of collaborative learning depends on individual- and group-level task engagement, which covers the interplay of learners' cognitive, metacognitive, emotional, motivational, and social processes and is influenced by contextual affordances and constraints. This phenomenon has been conceptualized as socially shared regulation of learning (SSRL; Hadwin et al., 2017). Due to the complexity of the phenomenon, the regulation of learning in collaborative settings presumes shared monitoring and control of these processes at individual and social levels. In collaborative learning, many events stemming from the ongoing situation and learning context may serve as catalysts, requiring regulatory responses within a group. These events have been designated as "triggers" (Järvelä, Nguyen, & Hadwin, 2023). Accordingly, SSRL theory (Hadwin et al., 2017) suggests that a distinguishing characteristic of successful groups is their ability to identify these trigger events, whether they pertain to cognitive, emotional, or motivational obstacles, and to adaptively and strategically respond to them through joint negotiations. It has been found that linguistically, joint negotiations result in linguistic alignment (Dideriksen et al., 2023; Pickering & Garrod, 2004), and physiologically, interacting partners exhibit physiological synchrony under certain conditions (Dindar et al., 2020; Schneider et al., 2020). In this paper, we define "negotiated sameness" as the co-constructed state within a group that results from joint negotiations that are essential for SSRL. Thus, linguistic alignment and physiological synchrony among the group members could be used to operationalize this state.

Over the past decade, researchers have increasingly studied collaborative learning with various physiological, behavioural, and contextual data streams to design and implement adequate, personalized support for learners and groups (see recent reviews by Febriantoro et al., 2023; Schneider et al., 2021). From the SSRL perspective, empirical evidence concerning the signals that recognize trigger events, as well as traces on student and group responses to these events, is vital for providing the necessary support. When viewed through the lens of jointly negotiated goals and (the enactment of) joint task plans, these trigger events may temporarily disrupt the negotiated sameness within a group, which can be evidenced at the physiological (Schneider et al., 2020) or contextual level (Vuorenmaa et al., 2023). Although methods based on multimodal learning analytics can offer real-time insights on collaborative learning processes based on physiological or behavioural data (such as electrodermal activity [EDA] or log data, Martinez-Maldonado et al., 2021; Wang et al., 2024), contextual data, such as video and audio recordings of group interactions, are essential to derive a meaningful interpretation of student behaviours or physiological responses (Schneider et al., 2021). Namely, video and audio recordings enable researchers not only to study student behaviours during collaborative learning tasks but also to understand how those behaviours are influenced by and embedded in the learning context (Molenaar et al., 2023).

In response to the challenges associated with the labour-intensive, manual analysis of group interactions from video and audio recordings, especially in face-to-face settings (Emara et al., 2021; Järvelä, Nguyen, & Hadwin, 2023; Lämsä et al., 2021), we apply a novel method for operationalizing the negotiated sameness at the contextual level. This method for capturing linguistic alignment relies on a linguistic analysis of transcriptions of group interactions. We employ this method with the aim of studying how trigger events for SSRL shape groups' negotiated sameness at the contextual (linguistic alignment) and physiological (physiological synchrony) levels. In the following section, we elaborate on the trigger event framework for studying SSRL. Afterward, we introduce how studying linguistic alignment and physiological synchrony can help us capture trigger events and group responses to them.

2. Literature Review

2.1. Trigger Events for Socially Shared Regulation of Learning

One of the challenges of studying regulation in collaborative learning is to understand how and when learners regulate their cognitive and socioemotional processes in response to changing conditions. SSRL involves multiple individual regulatory processes that contribute to joint regulation through negotiation and adaptation, when needed, to overcome challenges and achieve a common goal (Hadwin et al., 2017). However, SSRL is not a linear process but rather a dynamic and responsive one influenced by different events that occur during collaboration. To identify and make visible the events that invite — or "trigger" — regulation, Järvelä, Nguyen, and Hadwin (2023) created a conceptual framework of trigger events that empirically identifies specific kinds of events (or triggers) eliciting regulatory responses in collaborative learning, as evidenced in recent studies (Sobocinski et al., 2020; Vuorenmaa et al., 2023).

A trigger event can be classified according to the target of regulation that it invites, which can be cognitive, emotional, or motivational (Järvelä, Nguyen, & Hadwin, 2023). Cognitive triggers, whose sources can be either internal or external, are events that require learners to monitor and control their cognition. For example, an internal cognitive trigger might be a task definition that no longer reflects the group's understanding of the task. That is, if a group constructs a task definition early on that builds on conceptual misconceptions, and when it progresses with the task, it corrects its misconception, leading to a need to update the task definition. An external cognitive trigger might be a teacher giving feedback to learners about their progress.

Emotional triggers are related to events that require the monitoring and control of emotions, and like cognitive ones, their source can be either internal or external. Emotional trigger events can arise, for example, from interpersonal conflicts or externally scripted activities that create opportunities to monitor and control socioemotional aspects in interaction (Näykki et al., 2017). Certain trigger events can have both emotional and cognitive aspects. For example, if an enacted strategy repeatedly fails to produce the desired result, this event would require learners to both overcome their frustration (emotion regulation) and adapt the strategies they use (cognitive regulation). Trigger events can affect both individual- and group-level processes by evoking either self- or socially shared regulatory processes.

The regulation of learning, by definition, requires adaptation or change (Järvelä & Hadwin, 2013). To understand individual and group regulatory responses to the trigger events, there is a need to study the regulatory processes that these events evoke. These regulatory responses can be either adaptive or maladaptive (Sobocinski et al., 2020). When a discrepancy between the current state and judgment about goal attainment is perceived, adaptive regulatory response implies strategic changes in cognition, behaviour, motivation, or emotion (Miller & Hadwin, 2015). Adaptive regulatory response occurs when individuals' self-regulatory processes collectively contribute to controlling cognitive or emotional conditions through joint negotiations to optimize collaboration (Hadwin et al., 2017; Miller & Hadwin, 2015). In contrast, maladaptive regulatory response occurs when individuals or groups do not recognize the need for a regulatory response (i.e., they are not metacognitively aware of the trigger event) or they fail to take the regulatory actions needed (i.e., they do not have enough metacognitive knowledge to take the necessary control; Sobocinski et al., 2020). During the last decade, multimodal measures of collaborative learning have become increasingly affordable for researchers and practitioners, leading to their better understanding of the markers of these trigger events (Haataja et al., 2022) as well as the adaptive or maladaptive responses to the cognitive (Emara et al., 2021) or emotional trigger events (Törmänen et al., 2023) that groups face.

2.2. Multimodal Measures for Studying Trigger Events for Socially Shared Regulation of Learning

"Sharing" is central to collaborative learning, since joint goals and plans to achieve them are not achieved by accident but result from purposeful, externalized interactions within a group (Hadwin et al., 2017). Thus, group members' adaptive regulatory responses to trigger events that occur in collaborative learning during task completion and enactment derive from joint negotiations (Vuorenmaa et al., 2023). These joint negotiations result in achieving negotiated sameness. However, trigger events for SSRL may temporarily disrupt this negotiated sameness; after becoming jointly aware of the trigger event, a group must strategically respond and adapt to this event to achieve the negotiated sameness again. In the following section, we discuss how linguistic alignment can be used to operationalize this negotiated sameness at the contextual level, facilitating the identification of trigger events for SSRL and group responses to these events.

2.2.1. Linguistic Alignment

Spoken conversation is a collaborative process by which speakers coordinate their use of language to reach mutual understanding (Clark, 1996). This coordination is achieved through multiple cognitive mechanisms, among which is linguistic alignment (Pickering & Garrod, 2004). Alignment works as an automatic mechanism by which those who converse are primed to reuse their addressees' linguistic forms, resulting in a shared representations of the situation (Pickering & Garrod, 2004). It has been observed that alignment occurs at different linguistic levels across spoken interaction. Studies have shown that people tend to reuse their partners' lexical choices (Brennan & Clark, 1996), syntactic structures (Branigan et al., 2000), and accents (Giles et al., 1991), among other communicative signals.

A central feature of alignment is that the reuse of linguistic and non-linguistic forms occurs during an interaction and not in isolation (Pickering & Garrod, 2004), leading partners to take up joint actions together. Previous research has shown there to be a relationship between alignment and positive performance in joint tasks, in comparison with partners who do not align with each other (Dideriksen et al., 2023). Dideriksen et al. (2023) observed that interlocutors who performed better in task-oriented conversations showed richer lexical and syntactic alignment, suggesting that alignment permits higher complementary actions across dialogue partners.

Researchers have begun to investigate the implications of linguistic alignment in collaborative learning. For example, Haataja et al. (2022) found the equality of participation in metacognitive monitoring in collaborative learning to be associated with better group performance. Recent research has also demonstrated the value of investigating various levels of linguistic alignment in computer-supported collaborative learning (CSCL) to differentiate the phases of the collaborative learning process (Buseyne et al., 2024). Hayashi (2023) found evidence that linguistic alignment is related e.g., to higher reciprocal interactions and developing mutual understanding in a collaborative learning task. Likewise, research on collaborative learning has demonstrated correlations between linguistic alignment and learning outcomes and strong correlations between linguistic alignment and both gestural synchrony (i.e., coinciding changes in gesture) and gestural alignment (i.e., use of the same gestures; Sinclair & Schneider, 2021). This association may indicate that verbal data within collaborative learning can be used to measure the same underlying processes as nonverbal data, enabling triangulation between multimodal data streams.

In addition to using linguistic alignment to operationalize negotiated sameness at the contextual level of collaborative learning, researchers have also studied the negotiated sameness of groups at the physiological level with the help of physiological synchrony. Because the trigger events for SSRL require a group to take a regulatory action and strategically adapt to the ongoing situation, previous findings have suggested that physiological arousal and synchrony can also be used as metrics to evidence trigger events for SSRL (Haataja et al., 2022; Järvelä, Nguyen, Vuorenmaa, et al., 2023) and a group's adaptive responses to those events (Mønster et al., 2016; Schneider et al., 2020; Sobocinski et al., 2020).

2.2.2. Physiological Synchrony

Physiological arousal and synchrony have been related to both (meta)cognitive and emotional learning constructs (see the recent systematic literature review by Febriantoro et al., 2023). It is, however, crucial to point out how arousal and synchrony differ. Where physiological arousal is often seen as an increased sympathetic nervous system activity (Quadt et al., 2022) measured with, for example, EDA or heart rate, physiological synchrony refers to the interdependence or association of such signals between individuals (Palumbo et al., 2017).

In general, it is considered that physiological arousal serves as an adaptive behaviour by affording physiological resources to meet the appraised task demands regardless of whether those demands are physical, cognitive, or socioemotional (Quadt et al., 2022). An example of manifested physiological arousal might be an increase in EDA, reflected as moist palms due to an anticipated exam (Roos et al., 2021), higher perceived task difficulty (Malmberg et al., 2022), higher mental effort (Dindar et al., 2020), and intensity of emotion (Roos et al., 2021). In addition to focusing on states with high physiological arousal, states with low physiological arousal also deserve attention. For example, changes in a group's physiological state can indicate smooth task progress (Sobocinski et al., 2020) and higher-quality interactions (Schneider et al., 2020) in collaborative learning. Similarly, Törmänen et al. (2023) observed changes in group physiological states, especially during positive socioemotional interactions and regulation episodes. Altogether, these findings may provide evidence for adaptive responses to trigger events for SSRL.

While physiological arousal has been studied for years, physiological synchrony has gained increasing research interest more recently in collaborative learning. It is often evidenced as an alignment of physiological activity between participants in social settings. Generally, moments of physiological synchrony with high arousal have been rare in collaborative learning (Li et al., 2023; Törmänen et al., 2023). When capturing physiological synchrony among collaborating students, it has been related to positive (Sharma et al., 2019) and negative emotional valence (Malmberg et al., 2019), and the associations have been absent in some studies (Dindar et al., 2020). Li et al. (2023) found that physiological synchrony with high arousal was associated with both the occurrence of emotional trigger events and student regulatory responses to those events.

In addition to emotional arousal and valence, higher physiological arousal and synchrony among students may also be coupled with cognitive trigger events in collaborative learning in which a group becomes aware of the need for a regulatory act and strategic adaptation (Haataja et al., 2022). When strategic adaptation manifests as new ways of working, the physiological synchrony derived from EDA seems to again decrease (Mønster et al., 2016). Similarly, Järvelä, Nguyen, Vuorenmaa et al. (2023) noted that high physiological arousal and synchrony among students followed group actions that targeted completing the task (task execution) and preceded interactions that targeted group or individual ongoing cognitive processes (metacognitive interaction).

Several studies (Haataja et al., 2022; Malmberg et al., 2023; Schneider et al., 2020; Sobocinski et al., 2020) have shown that physiological synchrony can vary over time. When it comes to trigger events for SSRL, on the one hand, it can be hypothesized that triggering a group to adapt by increasing task demands can potentially be reflected as increasing physiological synchrony and arousal (Järvelä, Nguyen, Vuorenmaa et al., 2023). On the other hand, a change from high arousal synchrony toward an out-of-synchrony state could indicate that the group is able to flexibly adapt to these arising task demands (Malmberg et al., 2023; Mønster et al., 2016). In contrast to changing states of synchrony or arousal, continuous high arousal synchrony might mean that the group is aware of their challenging situation but is unable to adapt (Haataja et al., 2022).

Even though some studies have found significant associations between physiological metrics (e.g., physiological synchrony and high arousal) and collaborative learning constructs (e.g., Liu et al., 2021), these associations have been absent (Schneider et al., 2021) or even opposite (Yan et al., 2023) in other studies. Thus, the results regarding associations between physiological metrics and collaboration quality and performance have been mixed (Dich et al., 2018; Schneider et al., 2021). The results achieved in different studies are not necessarily contradictory per se, but the coupling of temporal changes in physiological synchrony with contextual observations might be a more fruitful approach to understanding the phenomenon in depth than just coupling the aggregate synchrony measures with outcomes (Sung et al., 2023; Yan et al., 2023). When physiological metrics are developed and used for making inferences about the need for regulatory acts or the enactment of the needed acts in collaborative learning in the pursuit of designing and implementing support, contextual affordances and constraints of collaborative learning should be considered. This argument is supported by the literature review of Schneider et al. (2021), who found that metrics based on social interaction data (e.g., linguistic features) providing information about the context in which collaborative learning takes place were significantly associated with the studied collaborative learning constructs with only few exceptions.

Thus, our aim is to study how trigger events for SSRL shape groups' negotiated sameness at the contextual (linguistic alignment) and physiological levels (physiological synchrony). In other words, we demonstrate how linguistic alignment and physiological synchrony together can provide evidence for the presence of cognitive and emotional trigger events for SSRL and group responses to those events. We address the following research questions (RQs):

RQ1: How is a group's linguistic alignment shaped when presented with cognitive and emotional trigger events for SSRL?

RQ2: How is a group's physiological synchrony shaped when presented with cognitive and emotional trigger events for SSRL?

RQ3: How do a group's linguistic alignment and physiological synchrony manifest in their interactions when presented with cognitive and emotional trigger events for SSRL?

3. Methods

3.1. Participants, Context, and Research Design

We designed a collaborative learning task with cognitive and emotional trigger events for SSRL implemented in a research infrastructure for first-year high school students. The participants were from the same Finnish high school and had similar socioeconomic backgrounds. We did not use criteria regarding student background information or knowledge when forming the groups; instead, the students were divided into small groups of three learners at random. Three groups were simultaneously present in the research infrastructure, although each group had a separate room. Each small group worked collaboratively on a shared Google document with their own laptops. Interaction among the group members was in face-to-face format. The task (30–40 min) was to design a nutritious breakfast smoothie for a customer based on the nutritional needs described in the document. In addition, the groups were provided with the nutritional information of the different food items that they could use in their smoothie recipes.

To study how trigger events for SSRL shape linguistic alignment and physiological synchrony, in this study, we focused on six groups (N = 18 students, comprising 14 males and 4 females) with cognitive and emotional trigger events. Each group received a cognitive trigger after 15 minutes of their collaboration that consisted of the customer sending a voice message telling them they had an allergy to latex and dairy products. This caused the groups to alter their smoothie recipes to exclude any ingredients that included the mentioned products. The groups also received additional emotional triggers every three minutes after the cognitive trigger had been introduced. The emotional triggers included the customer sending a voice message to the group asking them to hurry up. Even though the research design included three emotional triggers, most of the groups completed the task between the second and third emotional triggers. Thus, we decided to focus only on the cognitive trigger and two emotional triggers.

3.2. Data

To answer RQ1 about linguistic alignment, both group and individual microphones were used to capture audio data from the collaborative learning task, and these audio recordings were used for transcribing the interactions within the groups. We also used these transcriptions to illustrate how linguistic alignment and physiological synchrony manifest in group interactions when presented with cognitive and emotional trigger events for SSRL (RQ3). To answer RQ2 about physiological synchrony, physiological data (EDA) were collected with physiological Shimmer GSR3+ sensors. The sensors were synchronized before each session and attached to the participants' non-dominant hands so that gel electrodes were placed on the thenar and hypothenar eminences on their palms (Dawson et al., 2016). The sampling rate of the signals was set to 128 Hz. Due to problems with the EDA data collection, two groups had to be excluded from the physiological data analysis because EDA data were not available from all group members. Thus, we used data from four groups (n = 12) to address RQ2.

3.3. Data Analysis

To answer RQ1 about linguistic alignment, we measured levels of lexical alignment between group members. Prior work on lexical alignment in pairs of speakers has demonstrated that alignment in proportional usage of the 25 highest frequency words in a conversation is a strong proxy for overall alignment in a conversation while avoiding problems caused by sparsity of usage for rarer words, with this measure correlating strongly with group task success (Friedberg et al., 2012). Additionally, prior work on collaborative problem solving has likewise illustrated the significant relationship between task success and alignment on words from the task description (Friedberg et al., 2012; Rahimi et al., 2017). As such, this analysis considers alignment on both categories, that is, the 25 highest frequency words in each conversation and the words from the description of the task provided to each group.

Following prior alignment research (Friedberg et al., 2012) for a word (w) in each category, the alignment between a pair of speakers was measured using the following formula, where ALL is the total number of words from the respective speaker and count(w) is the number of times word w is spoken:

$$\text{align}(w) = -\left| \frac{\text{count}_{s1}(w)}{\text{ALL}_{s1}} - \frac{\text{count}_{s2}(w)}{\text{ALL}_{s2}} \right|$$

Note that the equation is negated, so that pairs of speakers with higher levels of alignment approach 0, representing perfect alignment for a given word, while perfectly misaligned speakers have scores approaching –1 for a given word. As in alignment research involving groups of more than two speakers (Friedberg et al., 2012), align(w) is computed for each category (c) of words by summing as follows:

$$\text{align}(c) = \sum_{w \in c} \text{align}(w)$$

This measure of alignment for a category of words was calculated for each pair of students in each group. Pair scores within a category were then averaged for that group to attain a single group score per category. Since Friedberg et al. (2012) found no significant differences between using a simple or a weighted average for group alignment scores, a simple average was used in this analysis.

While recent work on lexical alignment has extracted automatic measures of alignment at the level of shared expressions (Dubuisson Duplessis et al., 2017; Sinclair & Schneider, 2021), those measures have been developed for dyadic conversation and are not straightforwardly applied to groups of more than two speakers. Likewise, since those tools were developed for English corpora, the automated measurement of shared expressions has not been demonstrated as being fruitful for highly inflected languages such as Finnish.

Transcripts were preprocessed by removing punctuation; converting all words to lower-case; removing stop-words, nonspeech noises such as laughter, any part of the transcript indicated as not fully understood by transcribers, and any part of the transcript for which the identity of the speaker could not be determined; and lemmatizing all tokens. Tokenization was performed using the NLTK Python library version 3.7.[1] Where previously, the align(w) measurement has been applied to stemmed English-language tokens (Rahimi et al., 2017), lemmatization was performed, rather than stemming, for all tokens using the Libvoikko lemmatizer version 4.3.[2] Because Finnish is a highly inflected language, retaining the morphological equivalence between forms of words has been shown to improve the precision of computational linguistic measures applied to Finnish text (Hollink et al., 2004), motivating the decision to lemmatize tokens.

Changes in alignment were measured before and after the triggers, focusing on the cognitive trigger and the first two emotional triggers presented to each group. Three-minute windows of group discussions before and after each type of trigger were investigated in depth, following the granular analysis of these windows around triggers in prior work (Dang et al., 2023). The three-minute windows corresponded to the duration between trigger events, ensuring that all dialogue occurring after a trigger event was captured without any contamination from subsequent triggers. Additionally, maintaining consistent window sizes across triggers and groups was essential for comparing the effects of the different trigger events. Because the interactions differed in length, the mean amount of time per turn was first calculated (4.04 seconds), so these windows were operationalized as the 45 utterances immediately before and after a trigger so as to standardize the number of utterances included in the analysis of each group.

To answer RQ2 about physiological synchrony, we first detected the skin conductance responses (SCR) from the EDA data using the Ledalab toolbox (Benedek & Kaernbach, 2010). The signals were visually inspected for a lack of electrode contact, and we then removed small movement artifacts from the signal by applying the Butterworth low pass filter with frequency 1 and order 5. We used a threshold of 0.05 μS for detecting SCR peaks using continuous decomposition analysis (Benedek & Kaernbach, 2010; Society for Psychophysiological Research Ad Hoc Committee on Electrodermal Measures, 2012). After we had identified the peaks, we analyzed the shared physiological arousal events using the method developed by Dindar et al. (2022). We operationalized the events of physiological synchrony with high arousal by detecting the peaks of SCR and considered SCR peaks to co-occur within the group if they were manifested within a time window of 1.5 seconds by all the group members (Dindar et al., 2022). In addition to the events of physiological synchrony with high arousal, we detected the events of physiological synchrony with low arousal whenever no SCR peaks were manifested by any group members within the 1.5 second time window. Finally, if a group was not in synchrony with high or low arousal, we assumed it to be physiologically out of synchrony.

[1] nltk.org

[2] https://github.com/voikko/corevoikko

To study how the group physiological synchrony was shaped by cognitive and emotional triggers for SSRL, we studied the relative amount of time they spent in the three different physiological states (synchrony with high arousal, out of synchrony, synchrony with low arousal) three minutes before and after each trigger. As with measuring group linguistic alignment, we selected this three-minute time window, since it was the time interval between introducing consecutive trigger events.

Since changes in a group's physiological state can indicate (mal)adaptive responses to trigger events for SSRL (Schneider et al., 2020; Sobocinski et al., 2020; Törmänen et al., 2023), we also calculated transition probabilities among the different physiological states for each of the 3-minute time windows. To investigate how the trigger events shaped these transition probabilities, we calculated the differences before and after each trigger event. We visualized these differences by plotting heatmaps for one cognitive and two emotional trigger events. To investigate which differences in the transition probabilities were statistically significant, we applied the bootstrap method (Mooney & Duval, 1993) and created 1,000 shuffled time series of the group physiological states (so-called bootstrap samples). We then selected a six-minute segment from these shuffled time series and calculated the difference in the transition probabilities in the latter three-minute part of this segment compared to the former three-minute part. Finally, we calculated the proportion of times when the observed difference in each transition probability around the trigger event was more probable (for positive observed transition probability) or less probable (for negative observed transition probability) based on the simulated bootstrap sample. We assumed that the difference was statistically significant when $p < 0.05$, that is, when less than 5% of the observations in the bootstrap sample were larger (for positive observed transition probability) or smaller (for negative observed transition probability) than the observed sample.

To answer RQ3 and illustrate how a group's linguistic alignment and physiological synchrony manifest in their interactions when presented with cognitive and emotional trigger events for SSRL, we selected a transcription excerpt around the cognitive trigger and the second emotional trigger event. The duration of the excerpts was approximately two minutes. By highlighting task-description words and the most common words from the transcriptions, we elaborated on these excerpts from the perspective of (the lack of) changes in the linguistic alignment around these trigger events. We also elaborated on the changes in the target group physiological arousal and synchrony around these events.

4. Results

4.1. Linguistic Alignment After the Cognitive and Emotional Trigger Events

To ensure that linguistic alignment was different from the amount expected due to chance, the corpus of transcripts of the six groups was first compared against a shuffled corpus. Following the procedures used in prior alignment research (Dubuisson Duplessis et al., 2017; Sinclair & Schneider, 2021), for each transcript, a shuffled version was created by which each real utterance in sequence was interleaved with an utterance randomly selected from the real corpus. Overall, mean alignment was significantly higher for real interactions than for shuffled interactions both on the 25 most common words in the corpus (M $\text{align(c)}_{\text{Real}} = -0.608$, M $\text{align(c)}_{\text{shuffled}} = -1.180$, Wilcoxon rank-sum test U = 53, $p = 0.0025$, $r = 0.65$) and on the task-description words (M $\text{align(c)}_{\text{Real}} = -0.197$, M $\text{align(c)}_{\text{shuffled}} = -0.392$, Wilcoxon rank-sum test U = 55, $p = 0.010$, $r = 0.74$).

Likewise, to ensure that alignment immediately before and after triggers was different from the amount expected due to chance, the 45 turns ending in the real utterance before each trigger and the 45 turns beginning with the utterance after each trigger were selected for comparison. For each trigger and category of words, alignment was significantly higher for real scripts than for shuffled scripts (Wilcoxon rank-sum test $p < 0.01$ in all cases). This indicates that the alignment observed in the interactions was not due to chance and that changes in alignment around triggers were not arbitrarily associated with the position of a trigger in the script.

For task-description words, there was a significant difference in alignment before and after cognitive triggers, with alignment increasing significantly after cognitive triggers (M $\text{align(c)}_{\text{pre-cog}} = -0.014$, M $\text{align(c)}_{\text{post-cog}} = -0.006$, Wilcoxon rank-sum test U = 23, $p = 0.010$, $r = 0.741$). No significant difference in alignment on the most common words was observed before and after cognitive triggers (M $\text{align(c)}_{\text{pre-cog}} = -0.077$, M $\text{align(c)}_{\text{post-cog}} = -0.073$, Wilcoxon rank-sum test U = 36, $p = 0.630$, $r = 0.139$). For emotional triggers, there were no significant differences in alignment before and after either of the emotional triggers on either of the word categories. This result is summarized in Table 1 and visualized in Figure 1. Each group's linguistic alignment is separately presented in the Appendix.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Table 1. Group Linguistic Alignment*

	Pre-cognitive	Post-cognitive	*p*-value	Pre-emotional1	Post-emotional1	*p*-value	Pre-emotional2	Post-emotional2	*p*-value
align(c)$_{\text{25MostCommon}}$	–0.077	–0.073	.63	–0.073	–0.067	.20	–0.069	–0.058	.77
align(c)$_{\text{TaskDescription}}$	–0.014	–0.006	.01	–0.007	–0.011	.07	–0.012	–0.010	.57

**Note: Alignment is measured in the 25 most common words and task-description words before and after the cognitive trigger event and two emotional trigger events with the associated P-values.*

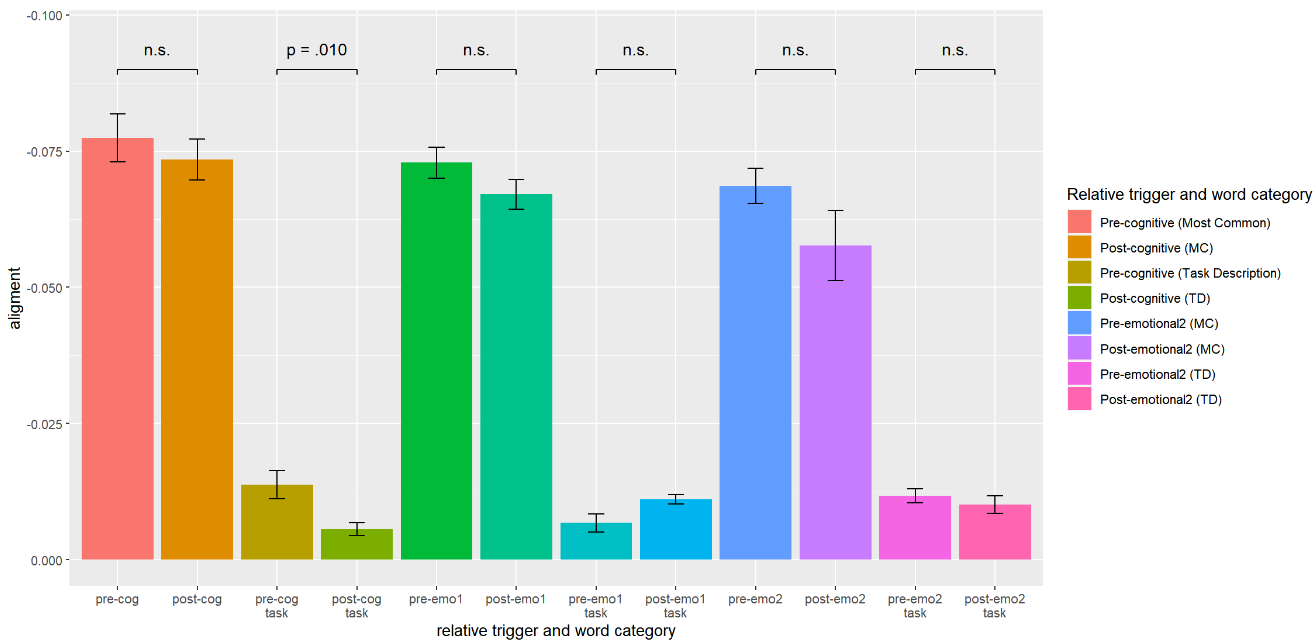


Figure 1. Group linguistic alignment in the most common words and task-description words before and after the cognitive trigger event and two emotional trigger events.

4.2. Physiological Synchrony After the Cognitive and Emotional Trigger Events

Figure 2 shows the relative amount of time four groups spent in the different physiological states (synchrony with low arousal, out of synchrony, synchrony with high arousal) three minutes before and after the trigger events for SSRL. First, between-group comparisons showed that regardless of the phase of the collaborative learning process, whether before the cognitive trigger or after the second emotional trigger, groups spent varying amounts of time in synchrony and out of synchrony. For example, while Group 2 spent 34% (26%) of the time in high arousal synchrony (low arousal synchrony) before the cognitive trigger, Group 4 only spent 17% (39%) of the time in high arousal synchrony (low arousal synchrony).

Second, within-group comparisons revealed that the physiological synchrony trajectories for some groups remained relatively stable before and after the cognitive and emotional triggers for SSRL. For instance, Group 1 consistently spent 18–23% of the time in high arousal synchrony and 30–35% of the time in low arousal synchrony, regardless of the phase of the learning process. In contrast, the amount of time Group 2 spent in high arousal synchrony decreased from 34% to 24% when comparing the phases before the cognitive trigger and after the second emotional trigger, with a concurrent increase in out-of-synchrony time over this period. In particular, the cognitive trigger did not seem to increase synchrony with high arousal, while the emotional triggers had the potential to increase synchrony with high arousal within some groups.

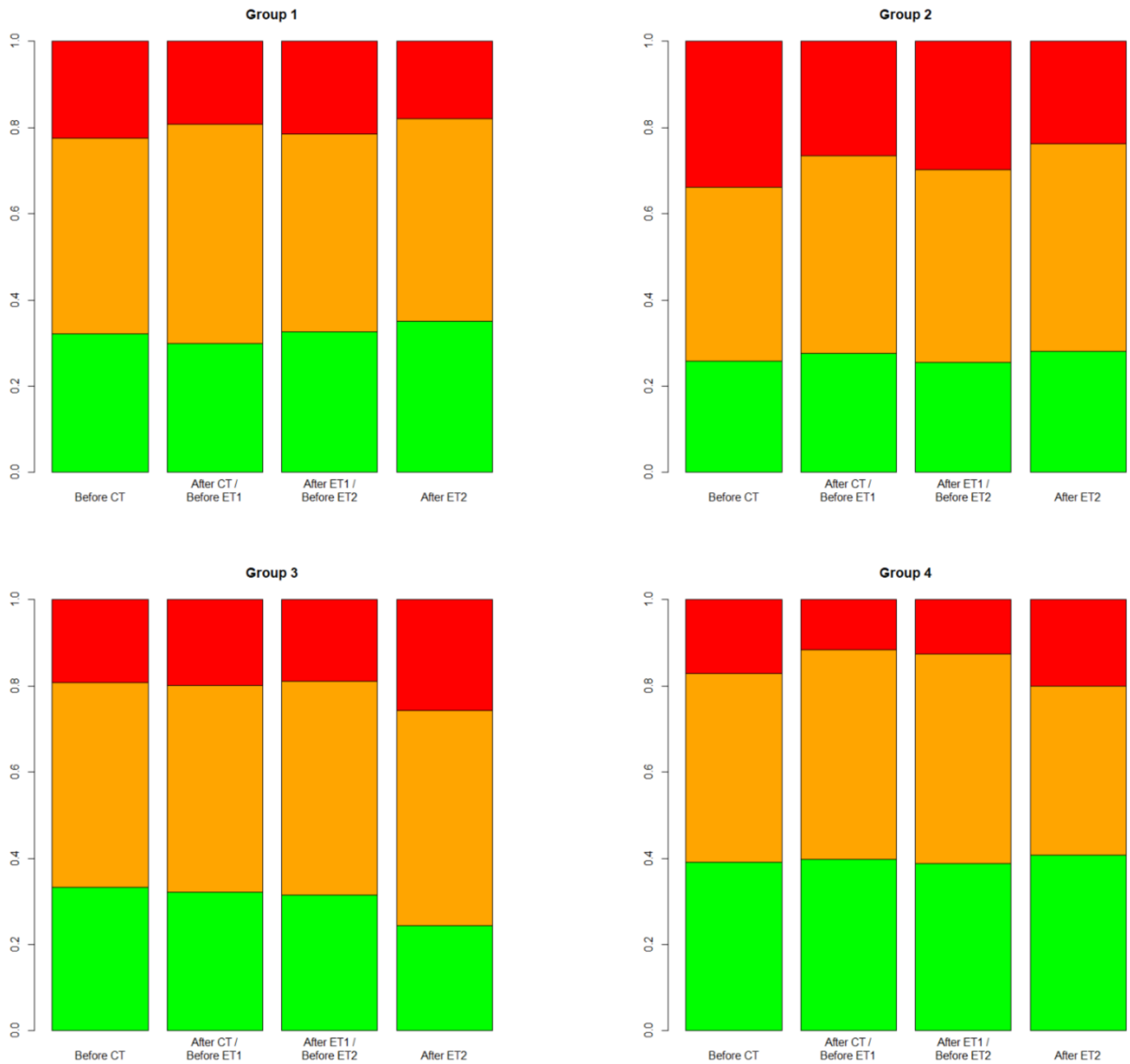


Figure 2. The relative amount of time the groups were in the three physiological states: synchrony with high arousal (red), out of synchrony (orange), and synchrony with low arousal (green) three minutes before and after the cognitive trigger (CT) and two emotional triggers (ET1 and ET2).

Figure 3 shows changes in transition probabilities among three physiological states for four groups before and after each trigger event for SSRL. First, the concurrent presence in the out-of-synchrony state (transition from out of synchrony to out of synchrony) was more probable after the cognitive trigger than before it (p = .04, Figure 3a). The cognitive trigger also reduced synchronous transitions from low/high arousal to low/high arousal (the four cells in the bottom left corner of Figure 3a) and increased the likelihood of transitions to an out-of-synchrony state.

Second, the transition from high arousal synchrony to an out-of-synchrony state was less probable after the first emotional trigger than before it (p = .05, Figure 3b). The first emotional trigger also promoted synchronous transitions from low/high arousal to low/high arousal (Figure 3b).

Third, the second emotional trigger stimulated transitions to a high arousal synchrony state from low arousal synchrony and out-of-synchrony states (Figure 3c; p < .001 for the transition from low arousal synchrony to a high arousal synchrony state and p = .02 for the transition from out-of-synchrony to a high arousal synchrony state). This trigger event also decreased the concurrent presence in the out-of-synchrony state (indicated by the negative transition probability from out of synchrony to out of synchrony in Figure 3c; p < .001) and the low arousal synchrony state (p = .04). Moreover, the transitions from out of synchrony to a low arousal synchrony state and from high arousal synchrony to out-of-synchrony states were statistically more probable after the second emotional trigger than before it (p = .006 and p = .02, respectively).

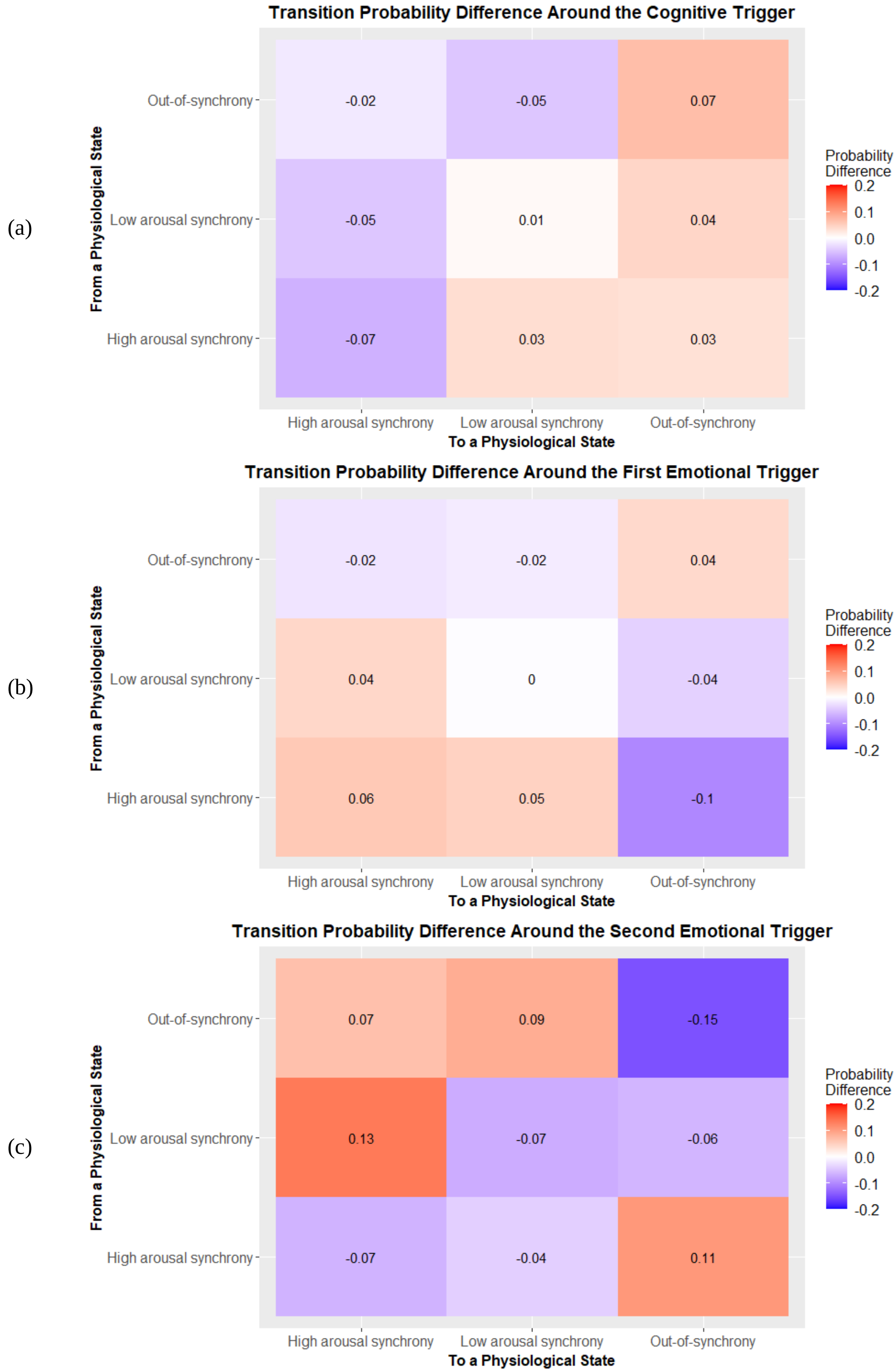


Figure 3. The heatmaps show differences in transition probabilities among the three physiological states, i.e., synchrony with high arousal, synchrony with low arousal, and out-of-synchrony: (a) after the cognitive trigger compared to before it; (b) after the first emotional trigger compared to before it; and (c) after the second emotional trigger compared to before it. Values in the cells represent the extent of the difference in probability; positive values (red) indicate a higher probability and negative values (blue) indicate a lower transition probability after the specific trigger.

4.3. Changes in Linguistic Alignment and Physiological Synchrony in Group Interactions

Figure 4 shows how Group 3's linguistic alignment manifests in their interactions when presented with cognitive trigger event for SSRL (receiving a voice message telling them the customer had an allergy to latex and dairy products). This excerpt demonstrates how linguistic alignment appears in the group interaction before and after the cognitive trigger. At the beginning of the excerpt, there is a lack of linguistic alignment: both students 1 and 2 used task-description words that other group members did not use (marked with red colour Figure 4). For example, when student 1 stated that "We now need fat and protein," student 2 replied that "What about quark — it has 295 calories," illustrating the lack of linguistic alignment. After the cognitive trigger event, the group shows a higher level of linguistic alignment, as they jointly redefined their task and worked to replace dairy products with other suitable ingredients. The linguistic alignment of task-description words (proteins and contain, marked with blue in Figure 4) coincides with shared task enactment. These changes in the linguistic alignment on task-description words did not become visible in the relative amount of time Group 3 spent in the different physiological states (Figure 2). After the cognitive trigger event, Group 3 spent longer periods of time in the out-of-synchrony state without transitioning to other physiological states compared to before the event: this finding is evidenced by the relative number of transitions (34% and 45%) from the out-of-synchrony state to the out-of-synchrony state before and after this trigger.

S1: So actually, now we need to add these others, we now need fat and protein,
especially fat for this task.
S3: In other words, more peanut butter.
S1: On the other hand, should we just change the milk to something else. Where…
S2: Uh, well, what about quark, for example, was it in there anyway?
S1: Quark definitely has more… quark...
S2: It has 295 calories.
S1: I'll just see how that looks.
S2: The most fat, well not for us, well, Greek yogurt could be good, it's still
quite healthy and then it has quite a lot of those nutrients.
S3: Well, let's put it in.
S2: Let's try Greek yogurt.
S1: That's it.
S1: Oh, now ours has fat! Now there is enough fat.
S3: Put in it a little less.
S2: Oops!
S3: Now it's quite a lot.
S1: Isn't that just the right amount of fat in relation to our carbohydrates?
S3: Yes.
S2: Now we should…

Cognitive trigger

S3: Where does it say latex?
S1: Yeah, I don't know.
S2: That last one of those nutrients.
S1: If milk proteins are not a good thing, then we need to change our Greek
yogurt.
S2: Okay, well.
S1: Or is this milk protein?
S2: It probably is.
S1: Yeah, well, it definitely has milk protein in it.
S2: What about, for example, oat milk?
S1: Well, it probably wouldn't have it. Let's try it. what does it contain?
S2: Oat or almond?
S1: Oat milk.
S1: Oat milk contains nothing.
S2: Well, it contains…
S3: Protein.
S1: Well, it contains a little bit of everything.
S2: What does AB yogurt contain? There must be milk in it.

Figure 4. Excerpt of Group 3's interactions before and after the cognitive trigger event. Task-description words marked in red were not used by other group members. Other students utilized the task-description words highlighted in blue.

Figure 5 shows how Group 4 is finishing the task and fulfilling the rest of the nutrition requirements by adding and removing different ingredients to the smoothie around the second emotional trigger event (receiving a voice message asking them to hurry up). The linguistic alignment on task-description words did not change after this trigger event even though the alignment of the most common words slightly increased (repetitive use of the most common words marked with blue in Figure 5). At the same time as the second emotional trigger event, student 1 stated "Oh good timing, be quiet" to the customer, indicating frustration with constant hurrying up. After this event, the group continued the task enactment as before and recognized that they have not yet met all the standards of the task "Now it is good, but we need more—calories." Even though linguistic alignment did not significantly change after the second emotional trigger event, the relative amount of time the group spent in physiological synchrony with high arousal increased after this event, while out-of-synchrony state was observed less (Figure 2). The probability of the transitions from synchrony with low physiological arousal to synchrony with high physiological arousal also increased (22% vs. 36%) and the probability of concurrently being in out-of-synchrony state decreased (41% vs. 30%) after this trigger event compared to before it.

```
S2: Okay, now [it] went over...
S3: ...kilocalories.
S2: You can at least remove the peanut butter.
S3: Less.
S3: Still less.
S2: Put 15, for example.
S3: Yeah, it is good.
S1: We need fewer carbs.
S3: And I put more of that protein.
S1: Where in this is most of the protein?
S2: If you add a bit of oat milk and take out a bit of those fruits?
S3: However, there should be more than 250 [grams] of fruits.
S2: Not like that.
S3: But yes, they still exist.
S2: We need a bit more protein. If I put in a bit more of this oat milk.
S1: There is approximately a bit more protein in this-
S3: And now there are a bit too few kilocalories.
S1: Oh yeah, appears so!
S2: If there is peanut butter, add extra.
S3: And now there are too many carbohydrates, where did they come from? Oh,
    never mind. [It] is good now.
S1: If we want to- Well, a lot- there are at least a lot of kilocalories in
    flaxseed oil.
S3: Yes, you can put it in.
S3: Except that we are almost full now.
Emotional Trigger
S1: Oh good timing, be quiet.
S2: Now [it] is good, but we need more...
    S1: Yeah.
S2: ...calories.
S3: So-
S1: Ah, what do we want?
S3: They should be added somehow in proportion so that circle does not go away.
S2: Now I want something with fewer calories.
S3: If I add 20 grams of pineapple?
S2: If there was that kind [of thing], where there are a lot of calories, but
    not much else. You could just add [that].
S3: Yes, if there is something. Wait, I'll look.
S1: Ah lots- what? So kilocalories, but nothing else.
S3: Right.
S2: For example put in that ice cream.
S3: But, is it…
S1: But doesn't it have fat?
S3: It is milk…
S1: Oh, right.
S2: Oh, right
S3: And was it- was it that they did not want any milk, or that the milk
    protein is not suitable?
S3: I don't remember how it was.
S1: It says yes, milk protein allergy.
```

Figure 5. Excerpt of Group 4's interactions before and after the second emotional trigger event. The most common words have been marked in blue and those were used by other group members.

5. Discussion

We aimed to study how trigger events for SSRL shape groups' negotiated sameness at the contextual (linguistic alignment) and physiological levels (physiological synchrony). First, we found that the cognitive trigger increased the linguistic alignment of the task-description words (i.e., students used more similar words for describing the task after the cognitive trigger than before it; RQs1&3, see Figure 4). The cognitive trigger also more frequently led to a concurrent presence in the out-of-physiological-synchrony state (RQ2). Second, neither the first nor the second emotional triggers led to any significant changes in linguistic alignment (RQ1s&3, see Figure 5). The first emotional trigger promoted synchronous transitions from low/high physiological arousal to low/high physiological arousal and decreased the probability of the transition from high arousal synchrony to an out-of-synchrony state (RQ2). The second emotional trigger decreased the probability of being concurrently in out-of-synchrony and low arousal synchrony states while stimulating transitions to a high arousal synchrony state from low arousal synchrony and out-of-synchrony states (RQ2).

The trigger events for SSRL were part of the treatment of the groups in this study. Although trigger events for SSRL can also originate from external sources in authentic learning environments, such as when a teacher provides revised task instructions (cf. the cognitive trigger when the customer sent a voice message telling the students they have an allergy to latex and dairy products) or when a group member suddenly needs to leave earlier due to a doctor's appointment (cf. the emotional trigger when the customer sent a voice message to the group asking them to hurry up), previous research has primarily concentrated on trigger events stemming from internal sources (Emara et al., 2021; Haataja et al., 2022). In the case of trigger events originating from external sources, the source of the trigger event itself facilitates (or co-regulates) collective awareness of the need for a regulatory act and invites regulatory response at the group level (Hadwin et al., 2017). Instead of an external collective invitation to regulate learning, trigger events originating from internal sources may necessitate the gradual development of collective awareness regarding the need for regulatory acts. This awareness can emerge through social interactions involving certain group members inviting (i.e., co-regulating) others to engage in regulatory activities (Hadwin et al., 2017).

These differences in the trigger events should be considered when interpreting the findings of this study. In previous studies, cognitive trigger events from internal sources have followed group monitoring that they are not progressing in the task according to the standards (metacognitive monitoring with negative valence; Haataja et al., 2022). In our study, external cognitive trigger (a voice message telling them the customer had an allergy to latex and dairy products) only updated the standards of the task. This difference may explain the difference in the physiological arousal and synchrony: That is, while the cognitive trigger following metacognitive monitoring with negative valence has increased high arousal physiological synchrony (Haataja et al., 2022), the external cognitive trigger in this study did not lead to an increase in physiological synchrony. Since the linguistic alignment of the task-description words increased after the external cognitive trigger, this might have related to the adaptive regulatory response to achieve the negotiated sameness by joint revision of group task understanding and enactment of task plans (evidenced in Figure 4 as the students' joint enactment to replace dairy products with other ingredients after the cognitive trigger), which has been shown to decrease physiological synchrony (Mønster et al., 2016). It has also been found that higher linguistic alignment of task-description words is associated with success (Friedberg et al., 2012; Rahimi et al., 2017) and developing mutual understanding in a collaborative learning task (Hayashi, 2023).

When Buseyne et al. (2024) studied linguistic alignment in a CSCL task, they did not find significant differences in it between all the task phases. Similarly, we found the lack of significant changes in the linguistic alignment after both emotional trigger events (voice messages to the groups asking them to hurry up). These findings may relate to the content of these trigger events that might not necessitate the need to renegotiate and adapt the joint goals or (enactment of) task plans that would have become visible at the contextual level. Instead, these triggers might only have led the groups to hurry in accomplishing their current goals with already negotiated task plans, maintaining the negotiated sameness within a group (Dang et al., 2023). This interpretation is also evidenced in our excerpt after the second emotional trigger event as shown in Figure 5. In contrast, the emotional trigger events increased transitions to high arousal physiological synchrony in the groups. These events can be considered as feedback to the groups that they are not progressing as fast as they should according to the standards. This feedback could induce metacognitive monitoring with negative valence within groups (see Figure 5), contributing to the existing empirical evidence that if a group monitors it is not progressing in the task according to the standards, high arousal physiological synchrony increases (Haataja et al., 2022). These findings also align with those of Li et al. (2023), who found that physiological synchrony with high arousal was more common during emotional trigger events (and related regulatory responses) than during cognitive ones.

When interpreting our findings relating to the second emotional trigger event, the general progress of the task might also be associated with the changes in the physiological synchrony. Namely, it has been noted that both high physiological arousal (Blikstein et al., 2017) and physiological synchrony (Malmberg et al., 2023) are more frequent at the beginning and at the end of a task. Thus, when we found that especially the second emotional trigger stimulated transitions to high arousal physiological synchrony, this may be related to the shared anticipation of a need to complete the task, which might have been further amplified by the second emotional trigger event. Moreover, the accumulating role of the trigger events (cognitive trigger and the first emotional trigger) might also have affected the more frequent physiological synchrony with high arousal after the second emotional trigger event (Galy et al., 2012).

When interpreting the findings of this study, the following limitations should be considered. First, the sample size (six groups, $N = 18$) was small, and since EDA data was not available from all the participants, we could only use data from four groups ($n = 12$) for physiological data analysis to address RQ2. Despite the small sample size, we were able to find statistically significant associations between the presence of trigger events for SSRL and linguistic alignment and physiological synchrony. A larger sample size would probably have made more difference, especially in the linguistic alignment visible at the level of statistical significance. Moreover, a larger sample size would have allowed us to perform comparative analyses among the conditions that received external cognitive and emotional trigger events and those that did not. Future studies with a larger sample size could also explore the effects of individual external trigger events to avoid potential effects from the accumulation of such events (one cognitive trigger event and two emotional trigger events in our study). In the future, human–AI collaboration and the use of automated measures for studying linguistic alignment and physiological synchrony could mitigate this limitation by enabling the scaling of these investigations with larger sample sizes (Järvelä, Nguyen, & Hadwin, 2023). Second, the collaborative learning task and associated trigger events for SSRL might not have been engaging enough (i.e., too easy or too complex) to stimulate physiological responses among all the participants (Törmänen et al., 2023). As task difficulty may be related to student physiological responses (Malmberg et al., 2022), future studies could use tasks with varying difficulty levels to better capture the range of cognitive and emotional obstacles that different types of tasks elicit. Moreover, multimodal measures of group learning traces could be used to personalize these learning tasks and associated trigger events (Hayashi, 2023), mitigating this limitation. Third, we aggregated investigations of the linguistic alignment and physiological synchrony over three-minute time windows. We chose this time interval as it was the duration between the presence of the different trigger events. Even though we were able to find statistically significant associations over these three-minute time windows, in the future, sensitivity analysis of different window sizes and a temporally more fine-grained analysis of linguistic alignment and physiological synchrony could provide further insights into how the trigger events shaped the negotiated sameness at the contextual and physiological levels. For more fine-grained analysis of linguistic alignment, measures of alignment across other linguistic levels such as prosodic (e.g., pitch or tonal alignment) and syntactic (e.g., alignment of sentence structures) may help reveal how the negotiated sameness develops dynamically around trigger events. We operationalized physiological synchrony using three physiological states, which we determined based on the (non-)co-occurrence of SCR peaks. We considered those peaks to co-occur within the group if they were manifested within a time window of 1.5 seconds (Dindar et al., 2022). This operationalization allows for the study of the relative amount of time the groups spend in the three different physiological states over a time window of any size, thereby enabling a more fine-grained analysis. Fourth, spoken interactions and linguistic analysis were in Finnish. This presents limitations to both the methods available for analysis and to generalizability. Insofar as automatic tools for detection of shared expression are built for English, a language that follows stricter word-order rules than Finnish, tools like dialign (Dubuisson Duplessis et al., 2017) have not yet been developed and limited linguistic research exists that could support the development of such tools in Finnish. Likewise, although linguistic findings mirror those found in similar research performed in an English-language collaborative learning study (Sinclair & Schneider, 2021), further research across other language contexts is needed to understand how particular our linguistic findings are to Finnish.

Despite these limitations, our study has important implications for supporting collaborative learning and SSRL with multimodal learning analytics. Our findings contribute to the existing empirical evidence that multimodal learning analytics can be used to interpret cognitive, emotional, and social processes in learning (Giannakos & Cukurova, 2023). However, the lack of learning theoretical underpinnings has limited the provision of plausible explanations for observed phenomena in the field of multimodal learning analytics (Giannakos & Cukurova, 2023). Our study and its research design were based on our theoretical understanding on SSRL (Hadwin et al., 2017) and empirical evidence on how trigger events stemming from collaborative learning situation may serve as catalysts, requiring regulatory responses within a group (e.g., Li et al., 2023). We showed how the contextual data (audio recordings from which the linguistic alignment was studied) and physiological data (EDA data from which the physiological arousal and synchrony were studied) may be sensitive to the different types of trigger events and group regulatory responses to those. Specifically, when looking at the negotiated sameness of the groups at the physiological level, we found that three different physiological states of the groups (synchrony with low arousal, out of synchrony, synchrony with high arousal) were needed to evidence the trigger events, instead of focusing only on the physiological synchrony with high arousal. Since the groups spent different amounts of time in physiological synchrony (with high or low arousal) and out of synchrony and their physiological responses, especially to the emotional triggers, varied (see Figure 2), this challenges between-group comparisons and the reliability and validity of physiological synchrony alone as an indicator of trigger events for SSRL. Our findings show that when combining physiological and contextual data about learning processes, linguistic alignment and physiological synchrony can together provide insights into the presence of cognitive and emotional trigger events for SSRL and group responses to those events. Audio and physiological data are also unobtrusive to collect in classrooms or other ecologically valid settings, which is crucial in practical implementations of multimodal (collaborative) learning analytics (Worsley et al., 2021). In terms of linguistic alignment analysis, online automatic speech recognition and alignment analysis would enable scaling and practical implementation, making alignment changes more visible in real time. Recent studies have also explored the potential of different natural language processing models, including large language models, to detect cognitive and emotional trigger events from student discourse (Suraworachet et al., 2024). For studying physiological synchrony, the development of cheaper, less intrusive, and wearable sensors (Ates et al., 2022), along with advancements of estimating heart rate from facial video recordings (Liu et al., 2024), could enable the scaling and practical implementation of physiological synchrony measurements.

6. Conclusion

This study shows how trigger events for SSRL shape groups' negotiated sameness at the contextual (linguistic alignment) and physiological (physiological synchrony) levels. In particular, we contribute by empirically verifying the trigger event conceptual framework (Järvelä, Nguyen, & Hadwin, 2023) as a theory-guided way to empirically evidence modes of regulation (self-regulation, co-regulation, and socially shared regulation) during learning and collaboration episodes. Our study also addresses the need for collaboration between the fields of CSCL and multimodal learning analytics not only by focusing on a core theoretical concept in collaborative learning, namely negotiated sameness, but also by searching for new methodological and analytical means to understand it.

It is still important to remember that any data stream or signal, such as physiological arousal or synchrony, is not an indicator of regulation in and of itself. However, the signal provides information about the internal, external, or shared contextual conditions that may trigger a purposeful strategic response. We were able to see this in our study, for example, in terms of the emotional triggers that increased a group's probability of moving physiologically to a high arousal synchrony state without significant differences in linguistic alignment. The lack of changes in linguistic alignment might indicate that groups strategically act not to renegotiate their task understanding or plans in response to the emotional trigger event but to focus on the enactment of current learning strategies to hurry up the task accomplishment, which has been shown to relate to increased high arousal physiological synchrony (Dindar et al., 2020; Sobocinski et al., 2020).

The findings of this study can be applied to how various data streams can be used in studying trigger events and eventually in developing metrics for AI-enabled real-time support for SSRL. For example, future AI agents could help students to become aware of the cognitive triggers (indicated, e.g., by an increase in linguistic alignment in task-description words) and emotional triggers (indicated, e.g., by physiologically decreased presence in out of synchrony), after which students can adaptively and strategically respond to these situations through joint negotiations. Moreover, with the help of multimodal measures, these AI agents could help students in taking adaptive regulatory response if they do not have enough metacognitive knowledge to take the necessary control themselves. However, more research is needed to better understand what factors show evidence of the impact of trigger events stemming from internal and external sources. As was seen in this study, multidisciplinary efforts are needed to make progress in this research. We had expertise from the learning sciences, educational psychology, learning analytics, information systems, and human–computer interaction fields collaborating in this study.

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 (a.k.a. Academy of Finland) Profi7 Hybrid Intelligence 352788 and Research Council of Finland projects 324381 and 350249. The data collection was carried out with the support of the LeaF Research Infrastructure, University of Oulu, Finland.

References

Ates, H. C., Nguyen, P. Q., Gonzalez-Macia, L., Morales-Narváez, E., Güder, F., Collins, J. J., & Dincer, C. (2022). End-to-end design of wearable sensors. *Nature Reviews Materials*, *7*(11), 887–907. https://doi.org/10.1038/s41578-022-00460-x

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Blikstein, P., Gomes, J. S., Akiba, H. T., & Schneider, B. (2017). The effect of highly scaffolded versus general instruction on students' exploratory behavior and arousal. *Technology, Knowledge and Learning*, *22*(1), 105–128. https://doi.org/10.1007/s10758-016-9291-y

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, *75*(2), B13–B25. https://doi.org/10.1016/S0010-0277(99)00081-5
Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology*, *22*(6), 1482–1493. https://doi.org/10.1037/0278-7393.22.6.1482
Buseyne, S., Rajagopal, K., Danquigny, T., Depaepe, F., Heutte, J., & Raes, A. (2024). Assessing verbal interaction of adult learners in computer-supported collaborative problem solving. *British Journal of Educational Technology*, *55*(4), 1465–1485. https://doi.org/10.1111/bjet.13391
Dang, B., Nguyen, A., & Järvelä, S. (2023). Clustering deliberation sequences through regulatory triggers in collaborative learning. *2023 IEEE International Conference on Advanced Learning Technologies* (ICALT 2023), 10–13 July 2023, Orem, UT, USA (pp. 158–160). https://doi.org/10.1109/ICALT58122.2023.00052
Dawson, M. E., Schell, A. M., & Filion, D. L. (2016). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (4th ed., pp. 217–243). Cambridge University Press. https://doi.org/10.1017/9781107415782.010
Dich, Y., Reilly, J., & Schneider, B. (2018). Using physiological synchrony as an indicator of collaboration quality, task performance and learning. *Proceedings of the 19th International Conference on Artificial Intelligence in Education* (AIED 2018), 27–30 July 2018, London, United Kingdom (pp. 98–110). Springer Cham. https://doi.org/10.1007/978-3-319-93843-1_8
Dideriksen, C., Christiansen, M. H., Tylén, K., Dingemanse, M., & Fusaroli, R. (2023). Quantifying the interplay of conversational devices in building mutual understanding. *Journal of Experimental Psychology*, *152*(3), 864–889. https://doi.org/10.1037/xge0001301
Dindar, M., Järvelä, S., & Haataja, E. (2020). What does physiological synchrony reveal about metacognitive experiences and group performance? *British Journal of Educational Technology*, *51*(5), 1577–1596. https://doi.org/10.1111/bjet.12981
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
Dubuisson Duplessis, G., Clavel, C., & Landragin, F. (2017). Automatic measures to characterise verbal alignment in human-agent interaction. *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue* (SIGDIAL 2017), 15–17 August 2017, Saarbrücken, Germany (pp. 71–81). Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-5510
Emara, M., Hutchins, N., Grover, S., Snyder, C., & Biswas, G. (2021). Examining student regulation of collaborative, computational, problem-solving processes in open-ended learning environments. *Journal of Learning Analytics*, *8*(1), 49–74. https://doi.org/10.18608/jla.2021.7230
Febriantoro, W., Gauthier, A., & Cukurova, M. (2023). The promise of physiological data in collaborative learning: A systematic literature review. *Proceedings of the 18th European Conference on Technology Enhanced Learning* (EC-TEL 2023), 4–8 September 2023, Aveiro, Portual (pp. 75–88). Springer Cham. https://doi.org/10.1007/978-3-031-42682-7_6
Friedberg, H., Litman, D., & Paletz, S. B. F. (2012). Lexical entrainment and success in student engineering groups. *2012 IEEE Spoken Language Technology Workshop* (SLT 2012), 2–5 December 2012, Miami, FL, USA (pp. 404–409). IEEE. https://doi.org/10.1109/SLT.2012.6424258
Galy, E., Cariou, M., & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types? *International Journal of Psychophysiology*, *83*(3), 269–275. https://doi.org/10.1016/j.ijpsycho.2011.09.023
Giannakos, M., & Cukurova, M. (2023). The role of learning theory in multimodal learning analytics. *British Journal of Educational Technology*, *54*(5), 1246–1267. https://doi.org/10.1111/bjet.13320
Giles, H., Coupland, J., & Coupland, N. (Eds.). (1991). *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge University Press. https://doi.org/10.1017/CBO9780511663673
Haataja, E., Malmberg, J., Dindar, M., & Järvelä, S. (2022). The pivotal role of monitoring for collaborative problem solving seen in interaction, performance, and interpersonal physiology. *Metacognition and Learning*, *17*(1), 241–268. https://doi.org/10.1007/s11409-021-09279-3
Hadwin, A., Järvelä, S., & Miller, M. (2017). 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. https://doi.org/10.4324/9781315697048-6
Hayashi, Y. (2023). Modeling synchronization for detecting collaborative learning process using a pedagogical conversational agent: Investigation using recurrent indicators of gaze, language, and facial expression. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-023-00381-y
Hollink, V., Kamps, J., Monz, C., & de Rijke, M. (2004). Monolingual document retrieval for European languages. *Information Retrieval*, *7*(1), 33–52. https://doi.org/10.1023/B:INRT.0000009439.19151.4c

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., 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

Järvelä, S., Nguyen, A., Vuorenmaa, E., Malmberg, J., & Järvenoja, H. (2023). Predicting regulatory activities for socially shared regulation to optimize collaborative learning. *Computers in Human Behavior*, *144*, 107737. https://doi.org/10.1016/j.chb.2023.107737

Lämsä, J., Uribe, P., Jiménez, A., Caballero, D., Hämäläinen, R., & Araya, R. (2021). Deep networks for collaboration analytics: Promoting automatic analysis of face-to-face interaction in the context of inquiry-based learning. *Journal of Learning Analytics*, *8*(1), 113–125. https://doi.org/10.18608/jla.2021.7118

Li, X., Hu, W., Li, Y., & Mao, Z. (2023). Exploring what synchronized physiological arousal can reveal about the social regulatory process in a collaborative argumentation activity. *Frontiers in Psychology*, *13*, 1042970. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.1042970

Liu, L., Xia, Z., Zhang, X., Peng, J., Feng, X., & Zhao, G. (2024). Information-enhanced network for noncontact heart rate estimation from facial videos. *IEEE Transactions on Circuits and Systems for Video Technology*, *34*(4), 2136–2150. https://doi.org/10.1109/TCSVT.2023.3301962

Liu, Y., Wang, T., Wang, K., & Zhang, Y. (2021). Collaborative learning quality classification through physiological synchrony recorded by wearable biosensors. *Frontiers in Psychology*, *12*, 674369. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.674369

Malmberg, J., Haataja, E., & Järvelä, S. (2022). Exploring the connection between task difficulty, task perceptions, physiological arousal and learning outcomes in collaborative learning situations. *Metacognition and Learning*, *17*(3), 793–811. https://doi.org/10.1007/s11409-022-09320-z

Malmberg, J., Haataja, E., Törmänen, T., Järvenoja, H., Zabolotna, K., & Järvelä, S. (2023). Multimodal measures characterizing collaborative groups' interaction and engagement in learning. In V. Kovanovic, R. Azevedo, D. C. Gibson, & D. lfenthaler (Eds.), *Unobtrusive observations of learning in digital environments: Examining behavior, cognition, emotion, metacognition and social processes using learning analytics* (pp. 197–216). Springer Cham. https://doi.org/10.1007/978-3-031-30992-2_12

Malmberg, J., Järvelä, S., Holappa, J., Haataja, E., Huang, X., & Siipo, A. (2019). Going beyond what is visible: What multichannel data can reveal about interaction in the context of collaborative learning? *Computers in Human Behavior*, *96*, 235–245. https://doi.org/10.1016/j.chb.2018.06.030

Martinez-Maldonado, R., Gašević, D., Echeverria, V., Fernandez Nieto, G., Swiecki, Z., & Buckingham Shum, S. (2021). What do you mean by collaboration analytics? A conceptual model. *Journal of Learning Analytics*, *8*(1), 126–153. https://doi.org/10.18608/jla.2021.7227

Miller, M., & Hadwin, A. (2015). Scripting and awareness tools for regulating collaborative learning: Changing the landscape of support in CSCL. *Computers in Human Behavior*, *52*, 573–588. https://doi.org/10.1016/j.chb.2015.01.050

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

Mønster, D., Håkonsson, D. D., Eskildsen, J. K., & Wallot, S. (2016). Physiological evidence of interpersonal dynamics in a cooperative production task. *Physiology & Behavior*, *156*, 24–34. https://doi.org/10.1016/j.physbeh.2016.01.004

Mooney, C. Z., & Duval, R. D. (1993). *Bootstrapping: A nonparametric approach to statistical inference*. Sage. https://doi.org/10.4135/9781412983532

Näykki, P., Isohätälä, J., Järvelä, S., Pöysä-Tarhonen, J., & Häkkinen, P. (2017). Facilitating socio-cognitive and socio-emotional monitoring in collaborative learning with a regulation macro script: An exploratory study. *International Journal of Computer-Supported Collaborative Learning*, *12*(3), 251–279. https://doi.org/10.1007/s11412-017-9259-5

Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., & Goodwin, M. S. (2017). Interpersonal autonomic physiology: A systematic review of the literature. *Personality and Social Psychology Review*, *21*(2), 99–141. https://doi.org/10.1177/1088868316628405

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–190. https://doi.org/10.1017/S0140525X04000056

Quadt, L., Critchley, H., & Nagai, Y. (2022). Cognition, emotion, and the central autonomic network. *Autonomic Neuroscience*, *238*, 102948. https://doi.org/10.1016/j.autneu.2022.102948

Rahimi, Z., Kumar, A., Litman, D., Paletz, S., & Yu, M. (2017). Entrainment in multi-party spoken dialogues at multiple linguistic levels. *Interspeech 2017*, 20–24 August 2017, Stockholm, Sweden (pp. 1696–1700). International Speech Communication Association. https://doi.org/10.21437/Interspeech.2017-1568

Roos, A.-L., Goetz, T., Voracek, M., Krannich, M., Bieg, M., Jarrell, A., & Pekrun, R. (2021). Test anxiety and physiological arousal: A systematic review and meta-analysis. *Educational Psychology Review*, *33*(2), 579–618. https://doi.org/10.1007/s10648-020-09543-z

Schneider, B., Dich, Y., & Radu, I. (2020). Unpacking the relationship between existing and new measures of physiological synchrony and collaborative learning: A mixed methods study. *International Journal of Computer-Supported Collaborative Learning*, *15*(1), 89–113. https://doi.org/10.1007/s11412-020-09318-2

Schneider, B., Sung, G., Chng, E., & Yang, S. (2021). How can high-frequency sensors capture collaboration? A review of the empirical links between multimodal metrics and collaborative constructs. *Sensors*, *21*(24), 8185. https://doi.org/10.3390/s21248185

Sharma, K., Pappas, I., Papavlasopoulou, S., & Giannakos, M. (2019). Towards automatic and pervasive physiological sensing of collaborative learning. *Proceedings of the 13th International Conference on Computer Supported Collaborative Learning 2019* (CSCL 2019), 17–21 June 2019, Lyon, France (pp. 684–687). International Society of the Learning Sciences. https://repository.isls.org//handle/1/1652

Sinclair, A. J., & Schneider, B. (2021). Linguistic and gestural coordination: Do learners converge in collaborative dialogue? *Proceedings of the 14th International Conference on Educational Data Mining* (EDM21), 29 June–2 July 2021, Paris, France (pp. 431–438). International Educational Data Mining Society. https://eric.ed.gov/?id=ED615547

Sobocinski, M., Järvelä, S., Malmberg, J., Dindar, M., Isosalo, A., & Noponen, K. (2020). How does monitoring set the stage for adaptive regulation or maladaptive behavior in collaborative learning? *Metacognition and Learning*, *15*(2), 99–127. https://doi.org/10.1007/s11409-020-09224-w

Society for Psychophysiological Research Ad Hoc Committee on Electrodermal Measures. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology*, *49*(8), 1017–1034. https://doi.org/10.1111/j.1469-8986.2012.01384.x

Sung, G., Bhinder, H., Feng, T., & Schneider, B. (2023). Stressed or engaged? Addressing the mixed significance of physiological activity during constructivist learning. *Computers & Education*, *199*, 104784. https://doi.org/10.1016/j.compedu.2023.104784

Suraworachet, W., Seon, J., & Cukurova, M. (2024). Predicting challenge moments from students’ discourse: A comparison of GPT-4 to two traditional natural language processing approaches. *Proceedings of the 14th Learning Analytics and Knowledge Conference* (LAK ’24), 18–22 March 2024, Kyoto, Japan (pp. 473–485). https://doi.org/10.1145/3636555.3636905

Törmänen, T., Järvenoja, H., Saqr, M., Malmberg, J., & Järvelä, S. (2023). Affective states and regulation of learning during socio-emotional interactions in secondary school collaborative groups. *British Journal of Educational Psychology*, *93*(S1), 48–70. https://doi.org/10.1111/bjep.12525

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, Z., Ng, J. T. D., Que, Y., & Hu, X. (2024). Unveiling synchrony of learners’ multimodal data in collaborative maker activities. *Proceedings of the 14th Learning Analytics and Knowledge Conference* (LAK ’24), 18–22 March 2024, Kyoto, Japan (pp. 922–928). https://doi.org/10.1145/3636555.3636935

Worsley, M., Martinez-Maldonado, R., & D’Angelo, C. (2021). A new era in multimodal learning analytics: Twelve core commitments to ground and grow MMLA. *Journal of Learning Analytics*, *8*(3), 10–27. https://doi.org/10.18608/jla.2021.7361

Yan, L., Martinez-Maldonado, R., Zhao, L., Li, X., & Gašević, D. (2023). Physiological synchrony and arousal as indicators of stress and learning performance in embodied collaborative learning. *Proceedings of the 24th International Conference on Artificial Intelligence in Education* (AIED 2023), 3–7 July 2023, Tokyo, Japan (pp. 602–614). Springer Cham. https://doi.org/10.1007/978-3-031-36272-9_49

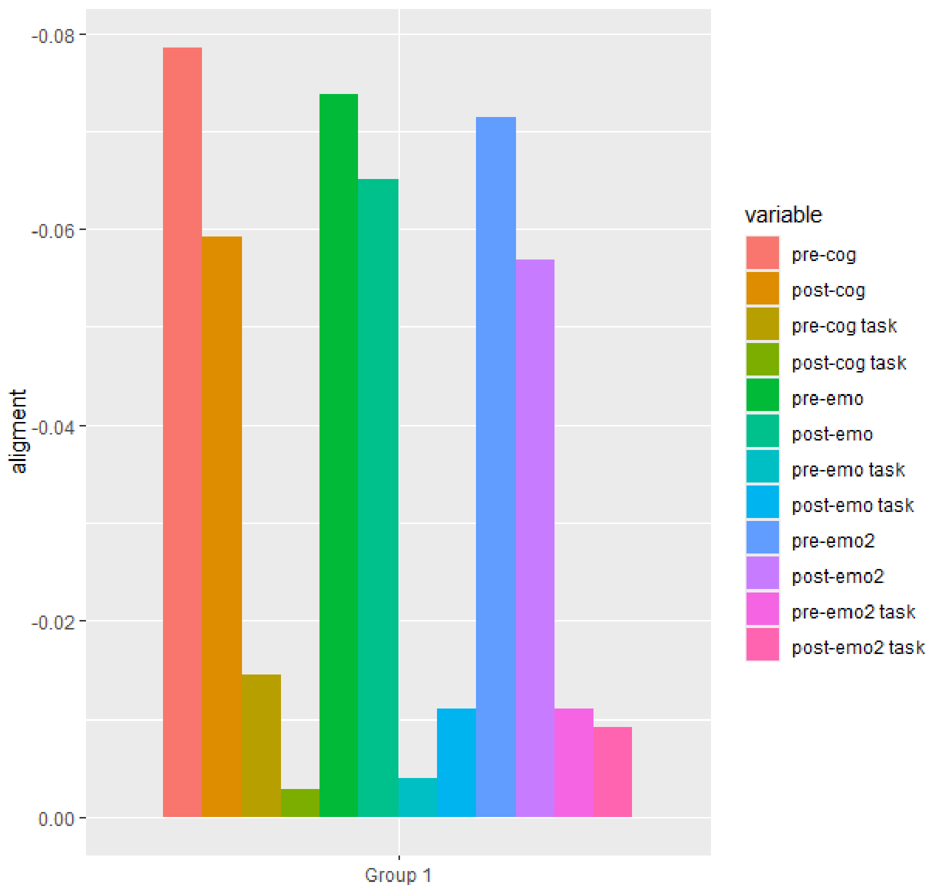
Appendix

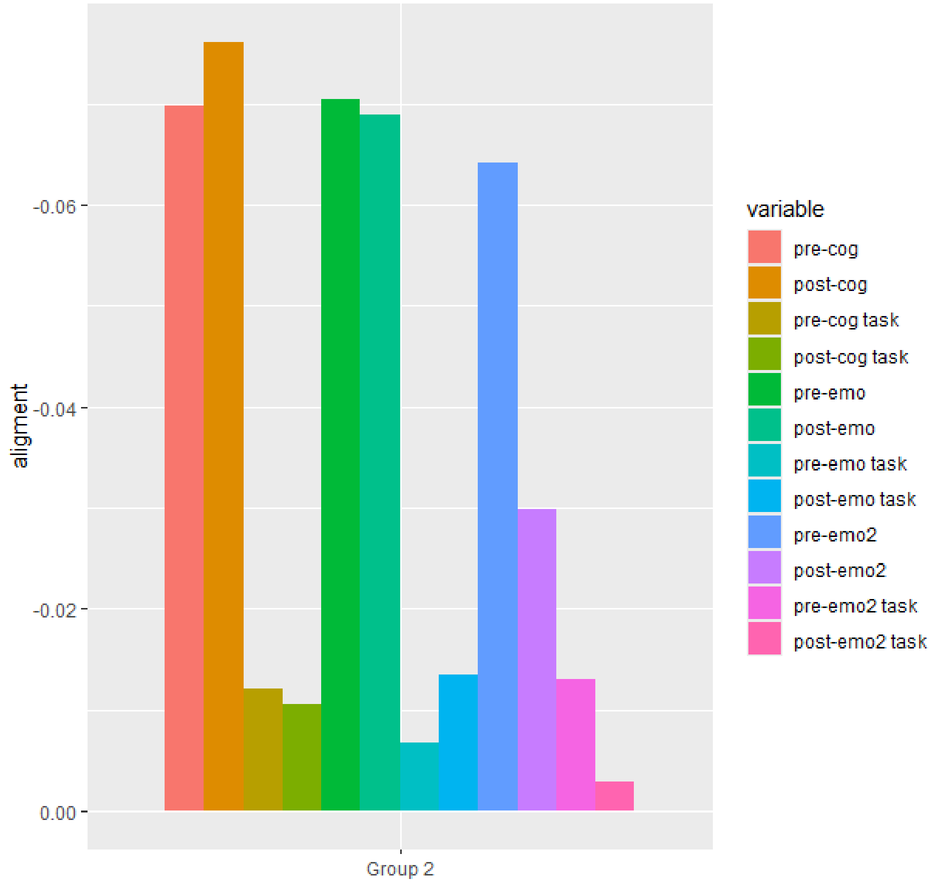
Group Linguistic Alignment

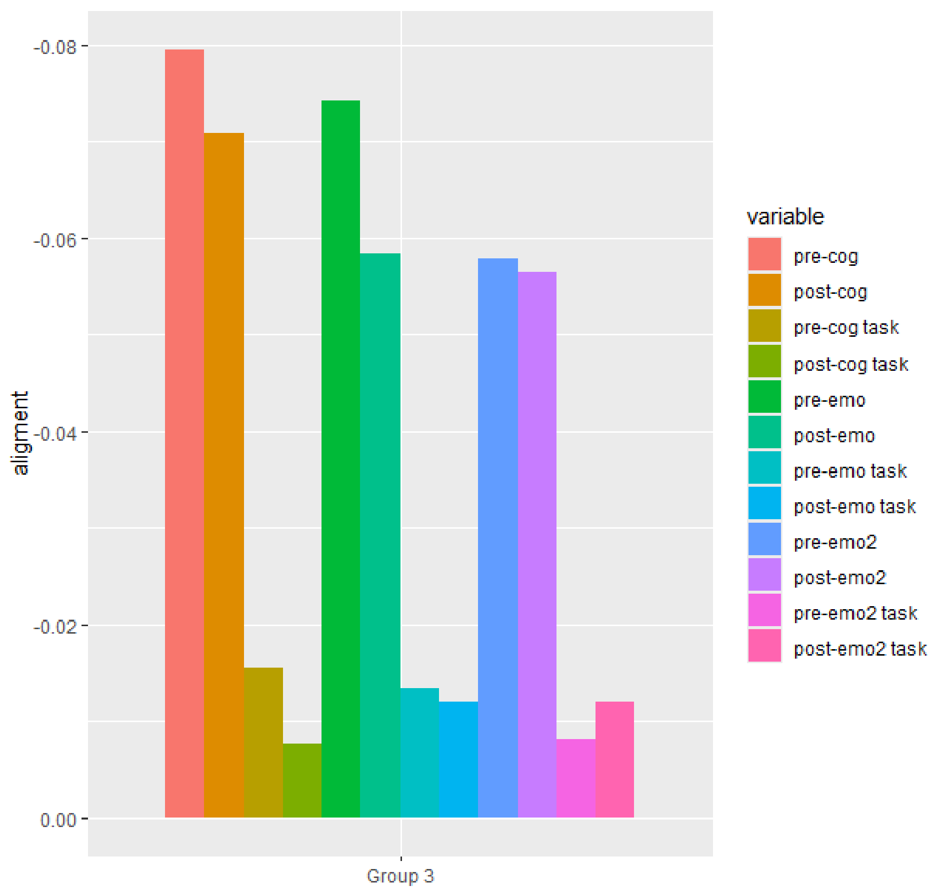
This appendix shows the data— in one table and six figures—for group linguistic alignment in the most common words and task description words before and after the cognitive trigger event and two emotional trigger events.

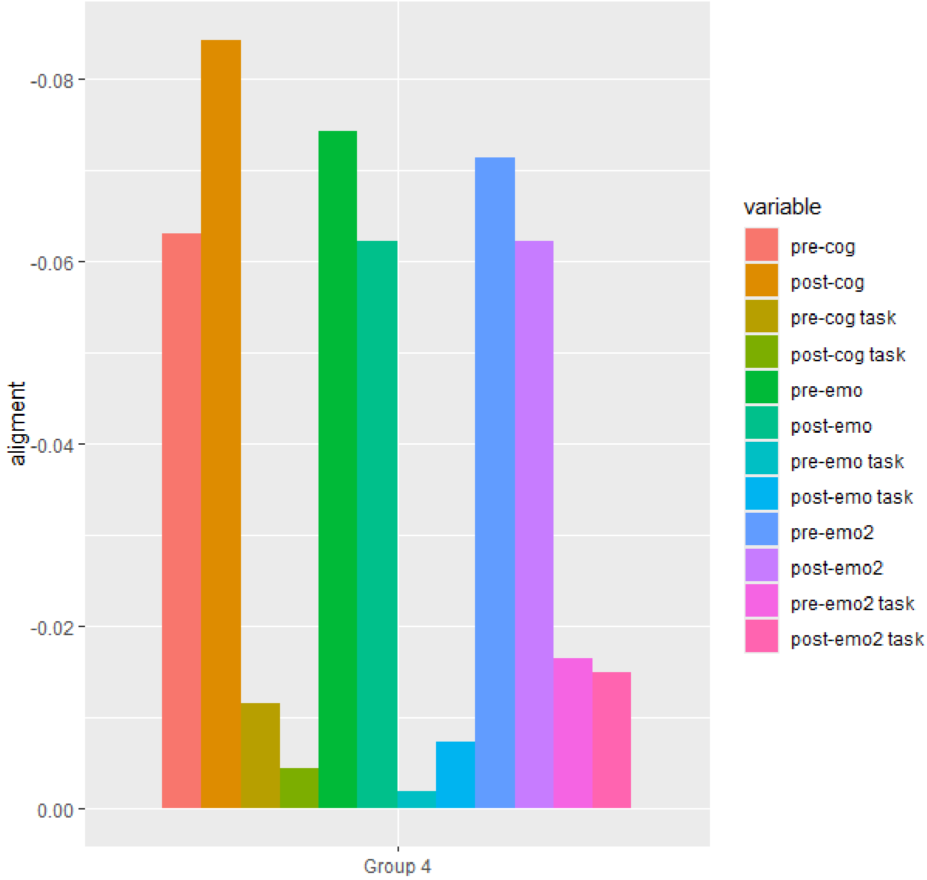
Table A.1 Linguistic Alignment Variables by Group

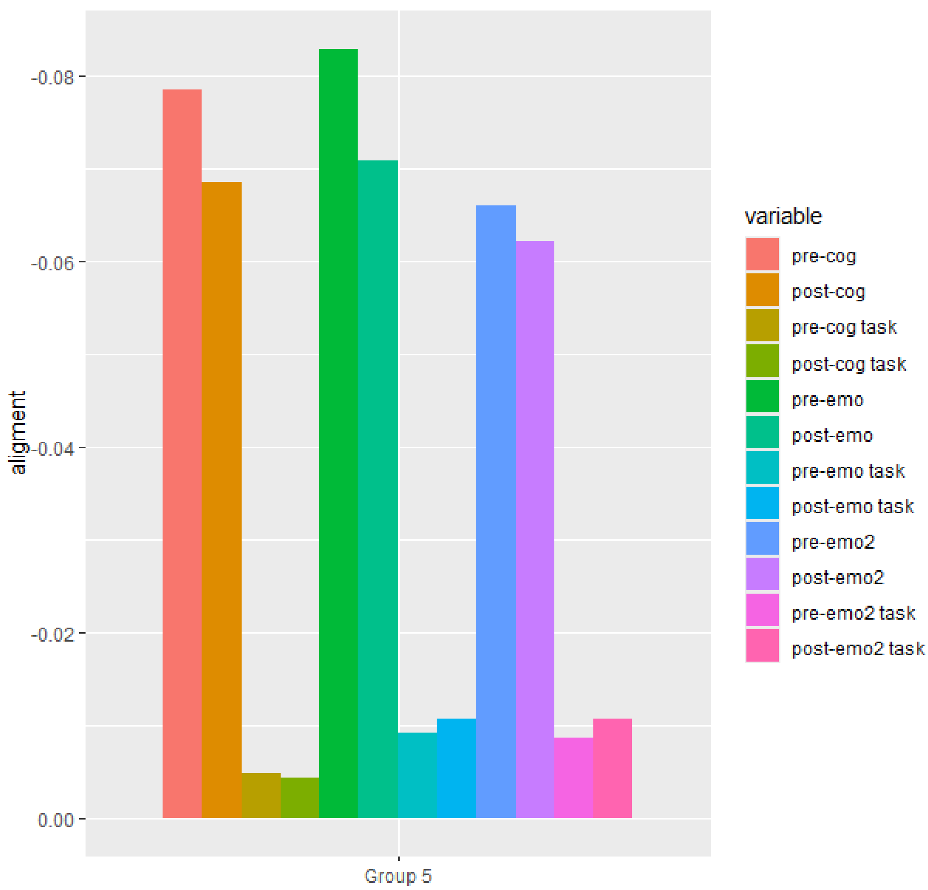
Group	pre-cog	post-cog	pre-cog task	post-cog task	pre-emo	post-emo	pre-emo task	post-emo task	pre-emo2	post-emo2	pre-emo2 task	post-emo 2 task
1	-0.079	-0.059	-0.014	-0.003	-0.074	-0.065	-0.004	-0.011	-0.071	-0.057	-0.011	-0.009
2	-0.070	-0.076	-0.012	-0.011	-0.07	-0.069	-0.007	-0.013	-0.064	-0.030	-0.013	-0.003
3	-0.080	-0.071	-0.015	-0.008	-0.074	-0.058	-0.013	-0.012	-0.058	-0.056	-0.008	-0.012
4	-0.063	-0.084	-0.012	-0.004	-0.074	-0.062	-0.002	-0.007	-0.071	-0.062	-0.016	-0.015
5	-0.079	-0.068	-0.005	-0.004	-0.083	-0.071	-0.009	-0.011	-0.066	-0.062	-0.009	-0.011
6	-0.095	-0.081	-0.024	-0.004	-0.062	-0.077	-0.005	-0.012	-0.081	-0.079	-0.013	-0.011

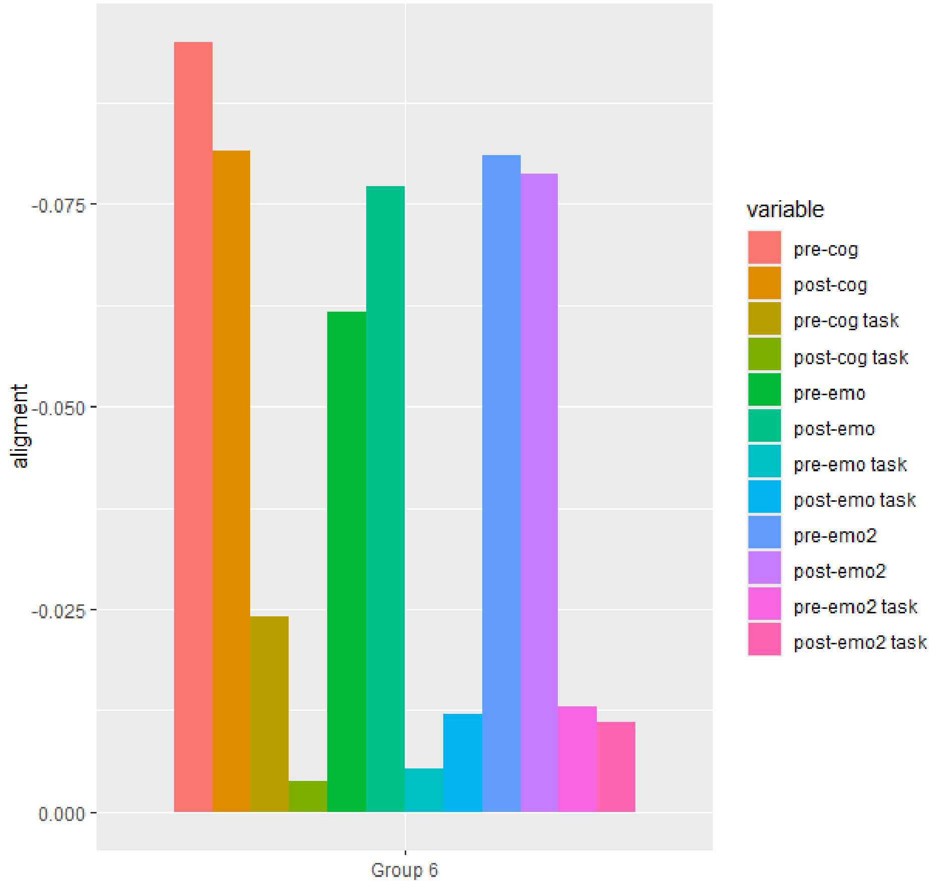



alignment
-0.06
-0.04
-0.02
0.00
Group 2
variable
pre-cog
post-cog
pre-cog task
post-cog task
pre-emo
post-emo
pre-emo task
post-emo task
pre-emo2
post-emo2
pre-emo2 task
post-emo2 task


alignment
-0.08
-0.06
-0.04
-0.02
0.00
Group 3
variable
pre-cog
post-cog
pre-cog task
post-cog task
pre-emo
post-emo
pre-emo task
post-emo task
pre-emo2
post-emo2
pre-emo2 task
post-emo2 task


-0.08
-0.06
-0.04
-0.02
0.00
aligment
Group 4
variable
pre-cog
post-cog
pre-cog task
post-cog task
pre-emo
post-emo
pre-emo task
post-emo task
pre-emo2
post-emo2
pre-emo2 task
post-emo2 task


-0.08
-0.06
-0.04
-0.02
0.00
aligment
Group 5
variable
pre-cog
post-cog
pre-cog task
post-cog task
pre-emo
post-emo
pre-emo task
post-emo task
pre-emo2
post-emo2
pre-emo2 task
post-emo2 task



-0.075
-0.050
-0.025
0.000
aligment
Group 6
variable
pre-cog
post-cog
pre-cog task
post-cog task
pre-emo
post-emo
pre-emo task
post-emo task
pre-emo2
post-emo2
pre-emo2 task
post-emo2 task