## Introduction

Studies on collaborative problem-solving in cognitive and learning science have revealed how concepts are understood or learned through social interaction. This type of learning is inspired by Vygotsky's socio-cognitive perspective (Vygotsky 1980). Many socio-constructivist researchers have analyzed students engaging in various kinds of social interactions and investigated the characteristics of successful and unsuccessful learners. Previous studies have shown that the integration of different knowledge and perspectives is a valuable learning experience (Greeno and de Sande 2007), and dialectal argumentation can lead peers to develop deeper levels of understanding (Asterhan and Schwarz 2009; Schwartz 1995). In such collaborative activities, learners are required to explain to others, which creates opportunities to integrate others' perspectives and develop a higher-level and more abstract representation of the content (Roschelle 1992). Researchers have shown that asking reflective questions for clarification to their conversational partners, who may have different perspectives, is an effective interaction strategy in order to better understand a problem or concept (Chi et al. 1994; Hayashi 2018b; Miyake 1986; Salomon 2001; Shirouzu et al. 2002).

In such collaborative situations, it is essential to focus on the cognitive learning process, considering aspects such as "how learners construct their knowledge based on others' perspectives" (Chi 2009). During these activities, the learners must successfully coordinate with each other and develop self-monitoring practices (Chi et al. 1994). Moreover, the learners must successfully establish common ground through conversations (Clark and Brennan 1991) on their differing knowledge and develop mutual understanding through integration of their respective perspectives. Knowledge integration tasks are used as an effective strategy to facilitate these interactions and require coordinating activities for joint practices such as establishing and maintaining shared understanding, as highlighted by Greiff et al. (2017). According to Rummel et al. (2009), achieving a successful collaboration has five dimensions: communication, information processing, coordination, interpersonal relationship, and motivation. Considering that collaborative activities are conducted in situations with low awareness, such as in computer-mediated environments that lack communication channels, difficulties may be encountered in achieving success with such a process. Moreover, verbal communication may produce misinterpretations, and this will be more difficult for those who do not have training or knowledge in the practice of good collaboration. Hence, several research questions arise: How can students best learn through social interactions under conditions of low awareness in online activities? What type of communication technology can support both explanation activities and higher levels of cognition during knowledge integration within these contexts?

Over the past several decades, research in human-computer interaction (HCI) has begun to investigate communication technologies enhancing social awareness, and these technologies have been introduced to the fields of computer-supported collaborative learning (CSCL) (Belenky et al. 2014; Schneider and Pea 2013). These show that social awareness tools may facilitate joint attention, enabling students to establish common ground and provide information/knowledge about their partners. These studies show that such tools improve the coordination process. However, not many studies clarify the impacts of these tools in regard to facilitating elaborated explanation activities. An example of such an activity would be complex knowledge integration tasks that require metacognitive processing in order to achieve success within the collaborative activity. Are these awareness tools capable of supporting cognitive processing, which involves successful coordination, or should other tools be used to complement them? Studies in CSCL have investigated several methods, such as scripts, prompts, orchestration, and representations, that describe how specific types of support can mediate participants' learning processes and outcomes (Rummel et al. 2009). Some of these studies use methods that provide interventions directly, such as teaching learners what to do or how to converse, and they may lack the independence of natural social interactive activities in collaboration.

In contrast, some studies use interventions that provide indirect metacognitive suggestions to facilitate self-regulated behaviors. One of the challenges in CSCL is to provide interventions dynamically, such as when it has been determined in real time that the learners' collaboration process needs to be supported. Recent studies have developed artificial intelligence systems to investigate the effects of providing metacognitive facilitation offered by pedagogical conversational agents (PCAs), which monitor the learners' behavior and intervene when external support is required (Hayashi 2019). These studies have examined how prompts and agents facilitate learner conversations, such as dialogues featuring explanations. However, it is not clear how these tools may have different effects depending on the context, especially with respect to how they impact the learning process and learning gains respectively. Furthermore, it is not fully understood how the combination of these two tools will play out, where each has advantages in connection with different aspects of collaboration, and how they may prompt the reflection on and reconsideration of issues of coordination at the meta-level needed in order for learners to succeed in gaining knowledge.

This study investigates the impacts of two facilitation methods, coordination support via learner gaze-awareness feedback and provision of metacognitive suggestions by a PCA, on the learning process and learning performance in peer collaborative learning. In the following subsection, the two methods considered in this study are discussed. These methods are designed to facilitate the coordination process and provide additional opportunities for individuals to expand their own and others' knowledge. Finally, the goal and hypothesis of this study are described in detail.



## Related work and study goal

### Two types of facilitation techniques for enhancing coordinative practice

As stated above, this study investigates the effects of learner gaze-awareness feedback and PCA-based feedback, which are two intervention paradigms that have been used to effectively facilitate coordination toward mutual understanding (Richardson and Dale 2005). These methods have recently been applied in educational contexts (Schneider and Pea 2013, 2014). This subsection first reviews and discusses studies on the utility of awareness tools, particularly visual gaze feedback via eye-tracking, in facilitating coordinative activities. Then, we review and discuss the use of the PCA (Heidig and Clarebout 2011), which is a useful tool for collaborative learning activities using Intelligent Tutoring systems (ITSs) as they facilitate metacognitive processes. Both the advantages and limitations of the PCA technology in terms of the provision of support for inter-learner coordinative activities are also examined.

## Social awareness: Visual gaze feedback via eye-tracking

Studies on computer-mediated communication have shown that, in distributed learning situations such as computer-mediated environments, individuals communicating through devices exhibit low awareness of one another and may form an incorrect understanding (Sproull and Kiesler 1991). Previous studies on computer-supported collaborative work, which were conducted over the last few decades, investigated the development of awareness tools that provide rich information on the ways individuals engage in activities through their experimental manipulation of behavior (Dourish and Bellotti 1992; Schmidt 2002). Hence, several types of awareness, such as social and cognitive awareness, were defined (Janssen and Bodemer 2013). In particular, social awareness refers a group member's awareness of the activities and the online states of others. Cognitive awareness refers to the awareness of information about group members' knowledge and expertise. According to Janssen and Bodemer (2013), these two types of awareness are important for learners performing social and communicative activities by establishing shared common knowledge and can further enable learners to acquire a deeper understanding of the domain knowledge associated with a given collaborative task.

Mutual gaze is commonly known as eye contact and is studied in the context of social relationships and interpersonal interactions (Goodwin 1981). In one of the primary studies in the field of HCI (Buxton and Moran 1990), systems known as "video tunnels" were developed. In these tunnels, half-silvered mirrors are used to set a camera angle as though the camera originates behind the eyes of the video image of a remote viewer. Similarly, a gaze awareness display inspired by ClearBoard (Ishii et al. 1993) was developed to enable speakers to establish full gaze awareness, including facial expressions. Such devices have been used to investigate how full gaze awareness can be an efficient resource for establishing grounding beyond that provided by a view of facial expressions with real mutual gaze (Monk and Gale 2002).

With advances in sensing technology, eye-tracking sensors were used in several studies to elucidate the nature of human-human coordination so as to facilitate the communication process (Jermann et al. 2011; Richardson and Dale 2005). Richardson et al. (2007) utilized two eye-trackers to investigate the relationships between speakers engaged in live, spontaneous dialog. Their analysis revealed a recurrence between the eye movements of the speaker and listener, which was established through shared common knowledge. A study showed that the degree of gaze recurrence (the portion of time for which the gazes are aligned) in speaker-listener dyads is correlated with the establishment of common ground; thus, building common ground positively influences visual attention coordination (Richardson and Dale 2005).

Various study results show that, in search tasks with different conditions, speakers can successfully communicate and coordinate their search efforts using shared gaze (Brennan et al. 2007; Keysar et al. 2000). In these studies, direct feedback on the visual gazes of collaborative partners was provided using eye-trackers (Jermann et al. 2011), which physically indicated the direction of the other collaborative learner's gaze on the same computer screen; hence, joint attention could be achieved. In another study, the learner sequence alignments were modeled (Khedher et al. 2017). Some well-known works (Schneider and Pea 2013, 2014) demonstrated that visibly representing a partner's gaze during a remote computer-based learning task can facilitate social collaboration and learning. In these studies, dyads collaborating remotely in a learning task learned about a neuro-science phenomenon by employing diagrams and tracing one another's gaze behaviors. In one experiment, the participants were provided with information on their partner's physical eye gaze on the screen. The control group lacked this information. Subsequent analysis revealed that real-time direct mutual gaze perception enables higher-quality collaboration for students.

Considering the results of these studies, gaze awareness tools can be taken to foster joint attention and take advantage of the collaboration process, which requires success in communication, as illustrated through examples such as perspective-taking in a knowledge integration task. This type of intervention is not invasive, so it has the benefit that it ensures learners' free engagement in social interactions. However, there are concerns that to generate effective collaboration, collaboration methods must be learned and thus require guidance, instruction, and training (Slavin 1992). These findings indicate that students who do not have experience or training in collaboration may encounter difficulty in complex explanation activities, specifically in regulating their own cognitive behaviors regarding decision making in particular circumstances. Awareness tools do not provide explicit guidance about how to foster the individual and collaborative learning process. Building on this point, the use of gaze awareness tools may be limited to only particular aspects of the collaborative process. Thus, it may be more effective to also use external support over and above what gaze awareness provides, in particular, interventions that facilitate the student's metacognition and self-reflection. The next section describes in detail the effective use of conversational agents to enhance collaboration support at the meta-level.

## Metacognitive suggestions: Pedagogical conversational agents

Previous studies in CSCL have investigated the use of external collaboration scripts for collaborative learning that are supportive of individual acquisition of knowledge. As discussed at the beginning of this paper, such methods take advantage of students building on each other's contributions within the knowledge integration tasks. Recent studies reveal that providing these external suggestions dynamically based on detecting learner states remains a challenge. In the context of ITS development, artificial intelligence in education has a long history (Aleven et al. 2006; Koedinger et al. 1997). Many of these tutoring systems provide adaptive and individual learner support, which would be difficult to achieve using human teachers alone. Other studies investigated the effects of teaching via tutoring systems (Biswas et al. 2005), the relative effectiveness of agent-provided facilitation prompts in self-regulated learning (Azevedo and Cromley 2004), and the development of systems employing advanced detectors to elucidate the learner state and generate facilitation prompts (D'Mello et al. 2012). Learning involving one-on-one dialog with dialog-based tutoring systems was shown to be more effective than simple reading and lecture attendance (VanLehn et al. 2007). When discussing the development of such dialog-based tutoring systems, it is important to mention the work of Graesser and studies on learners using AutoTutor (Nye et al. 2014). In those studies, conversational agents that provide hints, prompts, and motivate learners to meet expectations for answers to posed questions were developed based on student dialog analysis (Graesser et al. 2005). These emerging technologies for the design of PCAs as virtual teachers have been recognized as effective learner support methods.

Moreover, in 2015, the Program for International Student Assessment (PISA) governing board, administered by the Organization for Economic Cooperation and Development (OECD), assessed collaborative problem-solving using PCAs (Greiff et al. 2017). As regards the scope of the present study, there are several works that focus on the use of PCAs in learner-learner collaborations and the conversational agents were found to leverage performance by facilitating goal achievement (Holmes 2007), prompting periodic initiation opportunities (Kumar and Rosé 2011), and collaboratively setting sub-goals (Harley et al. 2017). Several design methods were investigated, such as the provision of positive emotional feedback via both dialog and visual representations of metacognitive suggestions (Hayashi 2012), the use of multiple PCAs based on this feedback (Hayashi 2019), and the use of gaze gestures during learner-learner interactions (Hayashi 2016). In those studies, the PCA successfully facilitated learner self-explanation activities and metacognitive behaviors, such as reflections. These previous studies on PCAs have shown that such technologies are useful for providing external support for internal processes, such as self-regulation and metacognition, primarily for individual-level learning support.

Building on these considerations, this study focuses on the following three types of functions: (1) Metacognitive suggestions, (2) facilitation of knowledge integration, and (3) communication encouragement. For (1), this work draws from a past study (Hayashi 2012), and related studies such as Azevedo and Cromley (2004), which have shown that the use of indirect facilitation techniques can facilitate self-regulation and metacognition. For (2), this study offers facilitation in the form of questions to learners requiring them to give examples related to task achievement. In previous studies, Graesser et al. (2005) found a set of interaction components prevalent in normal tutoring situations, such as anchoring learning in specific examples. Additionally, as Chi et al. (1994) point out, for students to develop a deep understanding, it is important for them not only to understand each separate component but to explain to themselves the relationships within and among them. For (3), this study employs facilitation prompts to elicit the learner's motivation-related remarks on communication, such as compelling communication between the learners. This was accomplished by providing positive back-channel feedback when the learners were using words related to the task activity. Also, previous studies show that the embodied characteristics of the agent and its role in stimulating the learners by encouragement have the ability to foster motivation towards learning (Baylor and Kim 2005). We adopted this point by implementing an embodied agent that synchronized its movements and provided positive feedback when the learners were using sophisticated words during their interactions.

## Study goal and hypothesis

Combining and integrating different background knowledge across members of a group is an effective strategy for developing new knowledge. Collaborative learning is beneficial in that it offers learners the opportunity to generate explanations and be exposed to different opinions from others, which might provide the opportunity to elaborate their own internal representation of knowledge. During such tasks, learners must both coordinate with others and regulate themselves to think dialectically and construct a comprehensive perspective. However, as discussed previously, knowledge integration activities may fail for several reasons. In computer-mediated environments, students often persist with low awareness about the perspectives of others, such as the topics/opinions their partners refer to during the exercise. Moreover, most learners, especially those in the early years of college, do not have any training or knowledge of how to coordinate successfully or self-regulate their cognitive behaviors to adjust their conversations in an ideal way.

Based on these points, this study used a simple knowledge integration task and investigated the effects of using gaze awareness tools, which are interventions that can foster joint attention, and external facilitations from a PCA, which can foster metacognitive awareness. Previous studies have examined the effects of using gaze awareness tools and PCAs on some specific tasks; however, there is a gap in the literature specifically with respect to evaluations of which supportive technology within this space holds the greatest potential for impact on the collaborative process and performance on knowledge integration tasks. By probing into this space in particular, we may be able to design better online collaborative learning systems, especially for knowledge integration tasks under challenging conditions with respect to group awareness. Therefore, this study investigated (1) direct facilitation using partner gaze awareness and (2) indirect third-person facilitation via a PCA, following a 2 × 2 controlled experiment design; hence, the manner in which these two methods facilitate the collaborative process was examined. On investigating the collaborative process, a coding scheme (Meier et al. 2007) was employed, which captures the crucial collaborative coordination features: mutual understanding, dialog management, information pooling, consensus reaching, task division, time management, technical coordination, reciprocal interaction, and individual task orientation (details are provided in the Methods section). For learning performance, this study focuses on assessing how well the learners were able to gauge differences in each other's knowledge.

Figure 1 shows the research framework of this study, with the two targeted methods highlighted. These methods facilitate the collaboration process, influencing the learning performance during the task activities. Note that good learning performance is the byproduct of a successful coordination process.

We predicted that using gaze awareness interventions would enhance joint attention. Therefore, learners can coordinate better with their partner by understanding what their partner is paying attention to during the task. More specifically, the partner's gaze provides awareness of their focus of attention, which enables learners to perform more successfully in all the collaborative processes enumerated in the coding scheme (Meier et al. 2007). Furthermore, it is predicted that if the tool enables the learners to see where their partner is looking while they are producing their explanations, it will allow them to see if their partner is paying attention to their explanations or referring to the suggestions from the PCA's comments. This may make it easier to plan their next conversational move and influences their turn-taking behaviors. Therefore, it is predicted that gaze awareness will influence the collaborative process of communication, such as mutual understanding and dialogue management. Also, awareness of their partner's gaze patterns can help reduce conflict when they try to pool information and

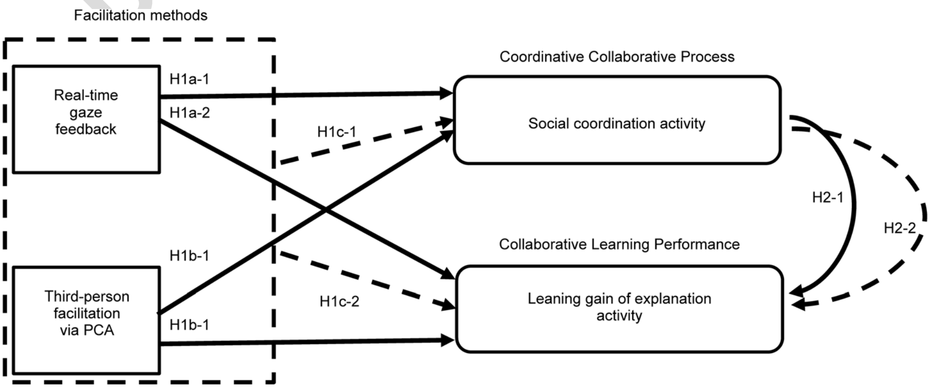


**Fig. 1** Study framework. (Left-hand side). Facilitation methods and (right-hand side) dependent variables investigated in this study. The hypotheses regarding synergetic use of the facilitation methods are indicated by dotted lines

reach a consensus because the gaze patterns show where their partner's areas of interest have been, such as their contributions or their partner's contributions or the PCA's comments. Therefore, it is predicted that gaze awareness will influence joint information processing, such as information pooling and consensus building. Moreover, if one can see that their partner's gaze is biased to a particular point, one might conclude that they should work on the task more efficiently, possibly through a change in roles. It is predicted that this will influence the collaborative process, such as coordination related to task divisions and time management. Also, if the learners are successful in developing such a process during the task, this may impact the learning performance's efficiency, deepening the understanding of each other's individual knowledge and their integrated knowledge. Therefore, in this study, it is hypothesized that learners using gaze feedback will achieve better results in the collaborative learning process compared to those who do not use such a method (H1a-1). Moreover, if learners can achieve successful collaborations during their explanations, this may improve their understanding of each other's different knowledge and therefore, influence learning performance. Consequently, it is expected that this will also affect the learning performance in the explanation task (H1a-2). However, H1a-2 may not produce a strong effect because gaze awareness does not directly scaffold metacognition and knowledge integration.

As mentioned in the previous section, the PCA will provide interventions such as (1) Metacognitive suggestions, (2) facilitation of knowledge integration, and (3) communication encouragement. Therefore, the PCA was expected to provide direct facilitation about coordination and metacognition, such as encouraging their activities and self-regulating their behaviors when making explanations to meet the task goal. Such direct verbal information should help the learners to think about the task goal, what to do, and what to talk about. The hops is that this process will lead them to more effectively adjust their behaviors to coordinate with each other more successfully. Moreover, the PCA's comments are expected to elevate their level of motivation, thus encouraging task orientation during their collaborative process. Considering these points, it was hypothesized that learners receiving these suggestions from a PCA would achieve superior performance in terms of the collaborative process as compared to those who do not have access to such support (H1b-1). However, metacognitive suggestions may have limited effect on facilitating the collaborative process. This form of support lacks in providing information about the partner's awareness, which may play an important role in establishing successful communication. In contrast, it was expected that the metacognitive suggestions would impact learning gains through better understanding of the task knowledge, which requires reflective cognitive processing (H1b-2). Therefore, PCA intervention might be more effective than the mutual gaze intervention with respect to collaborative performance, while gaze might be more effective at improving the collaborative learning process.

Upon review, each tool has its advantages and disadvantages in supporting collaborative process and performance in this knowledge integration task. Therefore, metacognitive suggestions about the collaboration process from the PCA is expected to complement the collaborative process for joint attention using gaze awareness tools. Conversely, metacognitive suggestions should provide learners with ways to think about coordinating by providing more visibility into the reasons behind their partner's gaze. Therefore, a combination of the two methods can be expected to facilitate coordinated activity (H1c-1) and influence learning (H1c-2). Overall, H1 pertains to the synergetic use of the targeted facilitation methods. In Fig. 1, this aspect is indicated by a dotted line.

The next aspect considered in this study is the relationship between the learning process and learning performance, and how this relationship is improved when the two facilitation methods (gaze feedback and PCA) are used. As discussed previously, successfully coordinated activities are essential for completing the task considered in this study, which is for learners to understand each other's different perspectives and to integrate these perspectives to develop new knowledge. Therefore, it can be predicted that successful coordination and explanation will yield higher performance in terms of learners' understanding of their own and others' knowledge (H2–1). Learners who receive both gaze feedback and PCA suggestions can exploit both facilitation methods. Accordingly, it is hypothesized that learners who employ both facilitation methods will achieve higher performance in terms of coordination and explanation than learners who do not use either or both methods (H2–2). The study and findings are reported below.

## Materials and methods

Q8

### Participants and experiment design

The study was conducted after an institutional ethical review and approval by the ethical reviewing committee of the author's university. There were 80 study participants (Female: 48, Male: 32, Average age: 19.85), and all were Japanese students majoring in psychology. From here, these participants are referred to as "learners." These learners participated in the experiment through a participant pool in exchange for course credit. Only freshmen and sophomore students majoring in psychology participated, and they were randomly grouped into same-gender dyads. The experimenter confirmed students within a dyad had not interacted with one another previously and that they did not possess technical knowledge related to debating.

When the participants arrived in the experiment room, the experimenter thanked them for their participation. The two participants briefly introduced themselves to each other. Following this procedure, the experimenter provided the task instructions, informing the participants that they would perform a scientific explanation task. They were told that they would use two different technical concepts from cognitive science to explain human language processing. Before the main task was initiated, they performed a free-recall test on these two concepts to ensure that the related knowledge was new to them. Next, each participant was given a detailed description of one of the concepts to study before they engaged in the task. In this learning phase, they were given information on only one of the concepts so that they would have to coordinate with their partner. Thus, the participants were required to explain the different concepts to each other and further discuss the unfamiliar concepts during the task. After the learning phase, the participants proceeded to the main explanation task, which had a 10-min duration. After the main task, the learners performed another free-recall test as a post-test. As mentioned earlier, a 2 (gaze: no gaze vs. visible gaze) × 2 (PCA: no agent vs. visible agent) experiment design was adopted to investigate the two factors of gaze feedback and PCA use.

### Task

The task was designed to investigate how learners explain to each other a topic the other partner is not familiar with, and to develop a comprehensive understanding of the said topic through their discussions. The participants are required to cooperate and understand each other's perspectives to complete the task. This part of the process is in common with the process of "jigsaw" methods studied in the learning science (Aronson and Patnoe 1997). The experimenter provided information on one of the two concepts separately to each learner. The learners did not know each other's concepts and, therefore, produced an explanation based on different knowledge than their partner was familiar with. To achieve the goal of explaining the topic using the two different conceptual frames, the learners needed to exchange knowledge via their respective explanations.

In detail, the learners' goal was to explain a topic (e.g., human information processing in language perception) using two sub-technical concepts (e.g., "top-down processing" and "bottom-up processing"). In the main phase of this task, each participant had to explain their assigned concept to their partner. Prior to this main phase (before the task began), a detailed description of the concept was provided to each learner. The learners each received a different technical concept, e.g., either "top-down processing" or "bottom-up processing," and worked on this assignment individually. During the main collaborative explanation task, a brief description of the participant's original concept was provided on their screen. Note that the participants could not see the brief description of their partner's task. Thus, the only way for each participant to gain an understanding of their partner's technical concept was from their partner's explanation. When the main task began, one learner was asked to first read their brief description and explain it to their partner. This was repeated for the other partner and, thus, the two different technical concepts were presented and explained. The learners were instructed that to complete this jigsaw-like task; each learner would need to explain their partner's technical concept to explain the overall topic using the two concepts successfully. The total time for the experiment, including the time for instructions and debriefing, was approximately 1 h.



### Experiment system

As shown in Fig. 2, each learner sat before a computer display. The learners could not see each other but could communicate orally, and they were instructed to look at the display while conversing with each other. For the PCA, a redeveloped version of a system designed in previous studies was used (Hayashi 2012, 2014, 2016). The system was programmed in Java for a server-client network platform and designed only for this experiment. The system used multi-threaded information processing for delivering messages during the network processing from the client programs installed in each of the participants' computers. The PCA was installed on the server and analyzed the conversations, sending signals to the client programs to provide metacognitive suggestions to facilitate the explanation activities (See the section describing the PCA for more specific functional information). Real-time direct visual gaze feedback about the partner was also presented on the display. The server received the gaze locations from the program client, inserted the logs into the database, and sent those to the client on the other learner's computer.

### Participant screens and gaze feedback

A brief explanation of the assigned concept was presented on either the right- or left-hand side of each learner's display (see Fig. 3). The explanation of the other learner's concept was presented on the opposite side but hidden so that the learner could not simply read the information on their partner's technical sub-concept and gain understanding in that way.

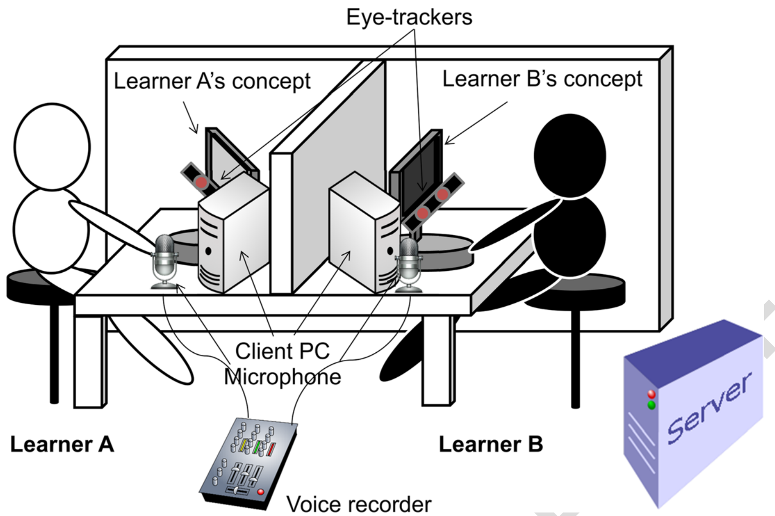


**Fig. 2** Experiment setup. The learners sit at the same table but cannot see each other

The only way a learner could fully understand their partner's concept was to ask questions and receive explanations.

To produce gaze feedback, two eye-trackers (X2–30, Tobii, Sweden) were used. A software visualization program was developed to track the partner's shifting visual gaze during the task in real-time, projecting it as a small moving square superimposed on the display (see Fig. 3). This real-time gaze feedback system was developed in C# and runs on a Windows 10 computer. Semi-transparent colored squares were used in the display so that the partner student's gaze pattern would not be too distracting, and the learners would still be able to read the text underneath the projected gave pattern. After the task was completed, the experimenter asked the learners if the indicator was distracting, and no learners claimed that it was.

As mentioned above, the learners were instructed to begin explaining by reading the text on their respective screens; this was a simplified version of the text they read before the main task. It was expected that, while one partner (learner A) explained their concept by looking at the

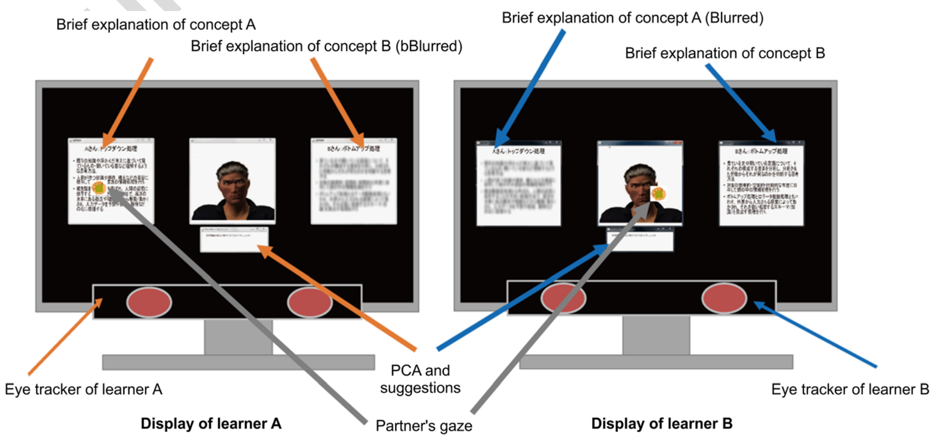


**Fig. 3** Sample participant screens. Learner A's gaze is shown on Learner B's screen (right-hand side) in the PCA area. Learner B's gaze is presented on Learner A's screen (left-hand side) in the concept description area

area with the concept explanation, the listener (learner B) can see their partner's gaze on the mosaic area where the other partner's brief explanation is blurred. Because of this, the text of the mosaic area is difficult to read. However, the listener (learner B) could both hear and trace their partner's gaze in the blurred area, enabling them to better understand what is written there.

## PCA

The PCA was originally developed in a previous study (Hayashi 2012, 2014, 2016) for use in text-based interaction. Since interaction in this study was conducted in speech, the experimenter and an assistant acted as intermediaries between the participants and the PCA software to enable the conversion between text and speech. For the detection of the learner's keywords and generation on the prompts from the PCA, the sentences were sent to the PCA on the server-side, and the PCA automatically analyzed the types of words to determine if the learner was providing effective explanations. Then, a rule-based generator determined the type of metacognitive suggestion to be offered. These were selected from five types, based on a previous study (Hayashi 2019), and involved facilitations that included clarifications of the learners' goal to encourage them to achieve efficient communication. The suggestion types are listed below.

- Type A: Facilitations to help learners consider the assignment purpose (e.g., "Please remember that the task is to explain the topic using the two concepts.") [Metacognitive suggestion].
- Type B: Facilitations to aid interpretation from a different perspective (e.g., "Try to consider the concept you are now explaining by using other examples.") [Facilitation of knowledge integration].
- Type C: Facilitations to urge learners to focus on concepts from both students (e.g., "When you have finished explaining one concept, switch turns.") [Facilitation of knowledge integration].
- Type D: Motivational remarks (e.g., "Good job! Keep going!") [Communication encouragement].
- Type E: Facilitations to aid focus on tasks and collaboration (e.g., "Stay focused on the topic," and "Pay attention to your partner's perspective") [Communication encouragement].

Based on the types of keywords used, the module defines whether the learners are (a) efficiently providing explanations or (b) not efficiently providing explanations. For (a), the system will look for words such as "schema" and "data-driven." While for (b) the system will look for words/phrases such as "don't know" and "give-up." If the system detects keywords defined in (a), it was programmed to present Type D facilitations, such as encouragements. For (b), the PCA will provide Type E facilitations. Next, if more than one minute has elapsed and there was no detection of keywords, the PCA randomly generates prompts of Types D and C (See Fig. 4). Moreover, the keyword detections in the figure were disabled if there were no prompts generated from Types A to C during the last four minutes. Additionally, the prompts generated automatically on the server-side are executed by the signal from the experimenter waiting for a momentary gap, because we did not want the PCA to distract the learners while they were talking.

The timing of the presentation of metacognitive suggestions was decided by the experimenter who sat on one side of the experiment room. The experimenter intervened

 



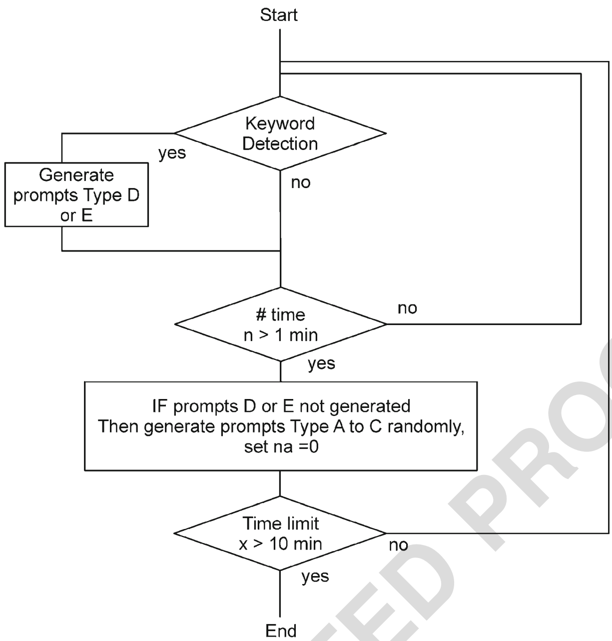


**Fig. 4** Flow chart of detection of keywords and types of facilitation presented



whenever there was a momentary gap in the dyad's conversation. No more than one signal was executed within a 1-min period, and the suggestions were controlled so that each would only be presented a maximum of 10 times during the task. The PCA was presented in the middle of the screen, and communicated through speech composition and physical movements (see Fig. 3). The length of the speech composition took an average of three seconds. Moreover, for each speech utterance by the PCA, a text version of the content was presented under the image box showing the physical movements. This enabled the participants to check whether their partner was paying attention to the PCA's comments during facilitation.

## Measures

As discussed in the Introduction, this study focused on performance components such as success regarding coordination and communication during the task. Dialog analysis was performed, for which the author transcribed all conversations into text and coded the dialogs based on the coding scheme explained below. In addition, the extent to which the learners were able to explain their own and others' concepts during and after the task was investigated. Thus, the analysis included dialog analysis and evaluation of the learning gains. The following subsections first explain the measures used for assessment of the collaborative process and then explain the learning gain evaluation.

### Collaborative process

To investigate the collaborative process, eight of the nine rating schemes from Meier et al. (2007) were used. Note that the "technical coordination" dimension was excluded, as there were no technical issues during the task, and it was not appropriate to annotate this point for this study. Based on the same principle, the definitions of some codes were also adjusted (Table 1).

For the analysis, the first procedure was to annotate the conversations using the coding scheme. The analysis was conducted at the end of multiple turns when there was a momentary gap during turn-taking in conversations. Two annotators discussed and coded the data using the definitions and the examples shown in Table 1. Then the annotators independently rated the conversations on a five-point scale. This procedure was followed by the method used in Schneider and Pea (2013). The inner reliability between the two coders was Kappa = 0.78. We used the average rating across the two annotators when there was a discrepancy. Kappa was also calculated for each separate code and shown in Table 1.

## Learning performance

To investigate the manner in which the independent variables (the two facilitation methods) influenced the learning performance, the author calculated the gain score from the pre- and post-task free-recall test scores. In both tests, the participants were asked to explain (1) their concept and (2) their partner's concept, as well as (3) the integrated conceptualization incorporating the two concepts. For each type of answer (1) to (3), the answers were coded as was done in a previous study (Hayashi 2016), where the explanations consisted of three different levels: (1) naive explanations that were made based on an individual's reasoning, (2) concrete explanations that were made based on the materials presented, and (3) further in-depth explanations that included analogies with knowledge transformations. The score for each code/category was based on the number of dimensions that comprise the category. The grading was performed by two coders (including the author and a volunteer) using the grading scheme presented in Table 2.

The inter-annotator agreement in terms of kappa for the grading was 0.73, after which the coders discussed their disagreements regarding the code and came to a consensus. The total score (for (1) self, (2) other, and (3) integrated) was taken as the dependent variable for the pre- and post-task test scores used for analysis. The gain score was calculated using the following equation:

$$\text{score} = \text{self} + \text{other} + \text{integrated} \tag{1}$$

$$\text{gain} = (\text{post-task test score} - \text{pre-task test score})/(1 - \text{pre-task test score}) \tag{2}$$

Thus, this proportional learning gain was used as an estimate of learning between the pre and post tests.

Note that the performance index calculated in this study differs from that reported by Hayashi (2018a), which considers self-explanation only and neglects the gain score determined from Eq. (2) of this study. This preliminary study focused on the effects of self-explanation only and did not consider the collaborative learning process or how this setting would affect individuals' understanding of their own and others' perspectives.

The average and standard deviation for the pre- and post-test raw scores for each condition were also calculated to confirm that pre-test scores were rather low, given the potential unfamiliarity of the topic. For the no visible gaze/no agent condition, the pre-test raw score

 



**Table 1** List of codes for the collaborative process, used in this study. Modified from Meier et al. (2007)

| | Definition | Example | Kappa |
|---|---|---|---|
| 1: Mutual understandings | The speaker avoids jargon and paraphrases it into a coherent statement. The listener provides their understanding through back channels and paraphrases, or they understand or asks for an explanation in an appropriate response. | A: Did you get it?<br>B: Yes, you mean that processing is based on the experience, right?<br>A: Yes, Exactly | 0.77 |
| 2: Dialogue management | The smooth flow of communication is maintained, and there is little duplication of conversation or confusion about who speaks. Turn taking is facilitated by questions or the delivery of explicit conversations. At that time avoid redundant phrases and fillers. Before starting a new conversation, get your partner's attention by calling a name or using a meta-statement. | A: Are you ready to hear about my concept next?<br>B: Mhm…. Yea.<br>A: Okay so I will start | 0.82 |
| 3: Information pooling | Partners try to collect relevant information for as many solutions as possible. New information is introduced in sophisticated ways, such as by associating with established facts and pointing out their relevance to solutions. It also draws expertise from its partners by using its expertise as a resource. | A: Wait, can you explain more about the definition schema in your text?<br>B: Sure, I will. | 0.77 |
| 4: Reaching consensus | Determining the options leading up to the final solution is at the end of a critical discussion that collects and evaluates the pros and cons. If partners prefer different options, they will discuss it until a factual agreement is reached. In addition, even if they agree, they are critically evaluated, and not only the facts they support but also the facts they disagree with are searched. | A: So, for your point it seems that information process comes first? Which do you think is correct?<br>B: Well I think that both could happen simultaneously. | 0.78 |
| 5: Task division | Tasks are divided into subtasks. Individual work is established either by planning or by short-term arrangements initially established by the partner. The work is divided evenly so you do not have to waste time waiting for subtasks to complete. | A: What about you going first and I go next?<br>B: That's a good way. After that we will discuss about how to combine these two. | 0.79 |
| 6: Time management | Monitor the time remaining and ensure that the task is completed in sufficient time to complete the remaining tasks. See if you need to spend a lot of time and remind each other of the time remaining. Set and monitor a realistic time. | A: So, we don't have much time for explaining our concepts shall be move on to the next phase?<br>B: Yes, let's do so. | 0.72 |
| 7: Technical coordination | Excluded from this study | | |
| 8: Reciprocal interaction | Treat each other with respect and provide opinions and views. Critical remarks are constructive, not personal. That is, they contribute equally to the task and are not dominated by one person. | A: I think I am talking too much.<br>B: No, no, please continue and I will explain after you finish. | 0.79 |
| 9: Individual task orientation | Each actively engages in finding solutions to problems and concentrates on their knowledge and skills. Pay attention to the information associated with the task, avoid distractions, and bring together your skills and resources, not just your partner. Show general interest in the subject or interest in fun or tasks. | A: So, this is interesting because it can be applied to other terms I learned in another psychology class.<br>B: Yea I agree. | 0.76 |



was SD = 0.3(0.732), and the post-test raw score was 1.7(1.417). For the visible gaze/no agent condition, the pre-test raw score was 0.6(1.382), and the post-test raw score was 4.55(1.972). For the no visible gaze/agent condition, the pre-test raw score was 0.11(0.345), and the post-test raw score was 3.34(1.330). For the visible gaze/ agent condition, the pre-test raw score was 0.25(0.524), and the post-test raw score was 4.45(0.385).

t2.1 **Table 2** Grading scheme for learner descriptions of their own and their partner's concepts, and of the relationship between the two concepts

| | Grade (points) | Considered dimensions | Examples (self-explanations) |
|---|---|---|---|
| t2.2 | | | |
| t2.3 | 0 | non-answer/incorrect | "I don't know how to explain this." |
| t2.4 | 1 | naive explanation, but correct | "Top-down processing is a kind of process that is oriented from a particular view." |
| t2.5 | 2 | concrete explanation based on materials presented | "Top-down processing is information processing that is driven from individual knowledge. This knowledge, which is a type of schema, is used to match the visual characteristics of the words, and if the fit is good, it will be a candidate of the meaning of the word." |
| t2.6 | 3 | concrete explanation based on materials presented and using examples | "Top-down processing is information processing that is driven from individual knowledge. This knowledge, which is a type of schema, is used to match the visual characteristics of the words, and if the fit is good, it will be a candidate of the meaning of the word. For instance, if there is a particular letter that is upside down or unreadable, we will use our schema, which can detect that word by complementation." |

## Results

### Collaborative learning process

This section presents the results of the social-collaboration conversational analysis. Table 3 lists the average ratio for each code under each condition and shows the statistical analysis results, where an asterisk (*) indicates a statistical significance of 5%.

To investigate the effects of each factor on each code, a 2 × 2 between-subject analysis of variance (ANOVA) was conducted for each code. For the "dialogue management" code, main effects were found for the use of gaze feedback. Thus, the learners tended to manage their dialogs when they used the visible gaze ($F\ (1,76) = 27.000$, $p = 0.0000$, $\eta^2_p = 0.2621$). For the "information pooling" code, a main effect was found for the use of the gaze feedback. That is, the learners tended to pool more information when they used the visible gaze ($F\ (1,76) = 93.957$, $p = 0.0000$, $\eta^2_p = 0.5528$). For the "consensus reaching" code, main effects were found for the use of both gaze feedback and PCA. Thus, the learners tended to reach consensus when they used the visible gaze ($F\ (1,76) = 29.277$, $p = 0.0000$, $\eta^2_p = 0.2781$) and the PCA ($F\ (1,76) = 11.244$, $p = 0.0012$, $\eta^2_p = 0.1289$). For the "task division" code, a main effect was found for the use of both the gaze feedback and the PCA. That is, the learners tended to effectively divide tasks when they used the visible gaze ($F\ (1,76) = 36.538$, $p = 0.0000$, $\eta^2_p = 0.3247$) and the PCA ($F\ (1,76) = 4.060$, $p = 0.0475$, $\eta^2_p = 0.0507$). For the "time management" code, a main effect was found for the use of gaze feedback. Thus, the learners tended to manage their time when they used the visible gaze ($F\ (1,76) = 55.583$, $p = 0.0000$, $\eta^2_p = 0.4224$). For the "reciprocal interaction" code, a main effect was found for the use of the PCA. That is, the learners tended to interact reciprocally when they used the PCA ($F\ (1,76) = 4.734$, $p = 0.0327$, $\eta^2_p = 0.0586$).

These results indicate that H1a-1 and H1b-1 are supported. However, the statistical analysis results show that there were no interactions; therefore, H1c-1 is not supported. The next subsection reports the gain scores for the learners' explanations of their partners' concepts and investigates how the two factors (i.e., the facilitation methods) influenced this performance.

**Table 3** Average ratio of each code by condition and significant main effects

| Code | no gaze/no agent | no gaze/ visible agent | visible gaze/ no agent | visible gaze/ visible agent | Main effect: Gaze | Main effect: Agent |
|---|---|---|---|---|---|---|
| 1 Mutual understanding | 3.6(0.663) | 3.6(0.800) | 3.7(0.781) | 4.11(0.537) | | |
| 2 Dialog management | 3.4(0.490) | 3.58(0.490) | 3.9(0.300) | 4(0.000) | *($p$ = .0000) | |
| 3 Information pooling | 2.7(0.640) | 3.23(0.700) | 4.4(0.663) | 4.33(0.632) | *($p$ = .0000) | |
| 4 Consensus reaching | 2.5(0.671) | 3.17(0.600) | 3.5(0.671) | 3.77(0.596) | *($p$ = .0000) | *($p$ = .0012) |
| 5 Task division | 3(0.775) | 3.47(0.640) | 3.8(0.400) | 4(0.000) | *($p$ = .0000) | *($p$ = .0475) |
| 6 Time management | 2.6(0.663) | 2.82(0.640) | 3.7(0.458) | 3.66(0.632) | *($p$ = .0000) | |
| 7 Technical coordination | Excluded from this study | | | | | |
| 8 Reciprocal interaction | 4(0.000) | 3.82(0.300) | 4(0.000) | 3.88(0.298) | | |
| 9 Individual task orientation | 3.5(0.806) | 3.58(0.663) | 3.8(0.400) | 3.55(1.274) | | |

## Gain score analysis: Concept-understanding performance

Using the gain score as the dependent variable, a 2 × 2 between-subject ANOVA was conducted. Figure 5 shows the average gain score results. There was a significant interaction between the two factors ($F$ (1,76) = 6.460, $p$ = 0.013, $\eta^2_p = 0.078$). Further analysis conducted for the simple main effects revealed that the score for the visible-gaze condition was higher than that for the no-gaze condition when there was no PCA ($F$ (1,76) = 13.627, $p$ = 0.00, $\eta^2_p$ = 0.152). Moreover, the score for the visible-agent condition was higher than that for the no-agent condition when the learners did not/did receive visible feedback on their partners' gaze ($F$ (1,76) = 33.880, $p$ = 0.000, $\eta^2_p$ = 0.308; $F$ (1,76) = 4.956, $p$ = 0.029, $\eta^2_p$ = 0.061, respectively).




**Fig. 5** Average Gain Score. The error bars indicate the standard deviations

This result shows, overall, that the use of the PCA improved the gain score for concept understanding during the task. Moreover, the use of the gaze feedback was efficient only when the PCA was not used (the score was higher for the visible-gaze/no agent condition than the no gaze/no agent condition), which is consistent with the results of related studies such as Hayashi (2018a). In addition, the use of the gaze feedback yielded higher performance when the PCA was used, which supports H1c-2. This implies that combining these two technologies is advantageous for facilitating performance of learner activities. To further investigate H2–1 and H2–2, the influence of the collaboration process on these results is discussed in the following subsection.

## Correlation between process and performance

Previous studies have shown that the quality of the collaborative learning process is correlated with the collaborative learning performance (Hayashi 2019). In this study, participants were required to perform a task (understanding each other's different perspectives by establishing common ground) in which a social coordination process positively affected their learning performance. Hence, it was hypothesized that learners who used both provided facilitation methods would have greater opportunities to exploit those methods and, thus, exhibit a good correlation between the interaction process and performance. Considering the results reported in the previous section, in which learners using both facilitation interventions (visible gaze/visible agent) exhibited higher learning performance, it is assumed that success in the collaborative process played an important role in this performance improvement. To investigate this point and test H2–1 and H2–2, it was predicted that learners who successfully achieved collaborative coordination would exhibit better learning performance and that this tendency would appear strongly for the learners who experienced the visible-gaze/visible-agent condition. Fig. 6 shows the correlation between the learning gain and collaboration process.

To investigate H2–1, i.e., to determine whether the coordinating process facilitated learning performance, a multiple regression analysis was conducted. In this analysis, the two variables of the collaborative process (consensus reaching and task division) were used, which were found as important variables that influenced the two types of interventions (see Table 3). This analysis also explains what type of variables in the collaborative process strongly influence learning performance. The regression coefficient $R^2$ was 0.109 and the ANOVA F-value was 4.720, indicating statistical significance ($p = 0.012$.). The equation used for the regression analysis was as follows (Eq. 3).

$$y = 0.053 + (0.071 * cr_i) + (0.025 * tv_i) \quad (3)$$

where cr indicates "Consensus Reaching" and tv indicates "Task Division." The results support H2–1.

Next, further investigation was conducted to determine whether successful collaboration strongly facilitated the learning gain performance for learners who were able to take advantage of both facilitation methods. To investigate H2–2, i.e., to determine whether the coordinating process strongly facilitated the learning performance, especially for the visible-gaze/visible-agent condition, multiple regression analysis was conducted for all conditions.

For the visible-gaze/visible-agent condition, $R^2$ was 0.322, and the ANOVA F-value was 4.035, indicating statistical significance ($p = 0.037$). The equation used for the regression analysis was as follows (Eq. 4):

$$y = -53.653 + (0.009 * cr_i) + (13.453 * tv_i) \tag{4}$$

The results differed with the other conditions. For the no-gaze/no-agent condition, $R^2$ was 0.085, and the ANOVA F-value was 0.786, indicating no statistical significance ($p = 0.472$). Next, for the no-gaze/visible-agent condition, $R^2$ was 0.125, and the ANOVA F-value was 1.217, indicating no statistical significance ($p = 0.321$). Finally, for the visible-gaze/no-agent

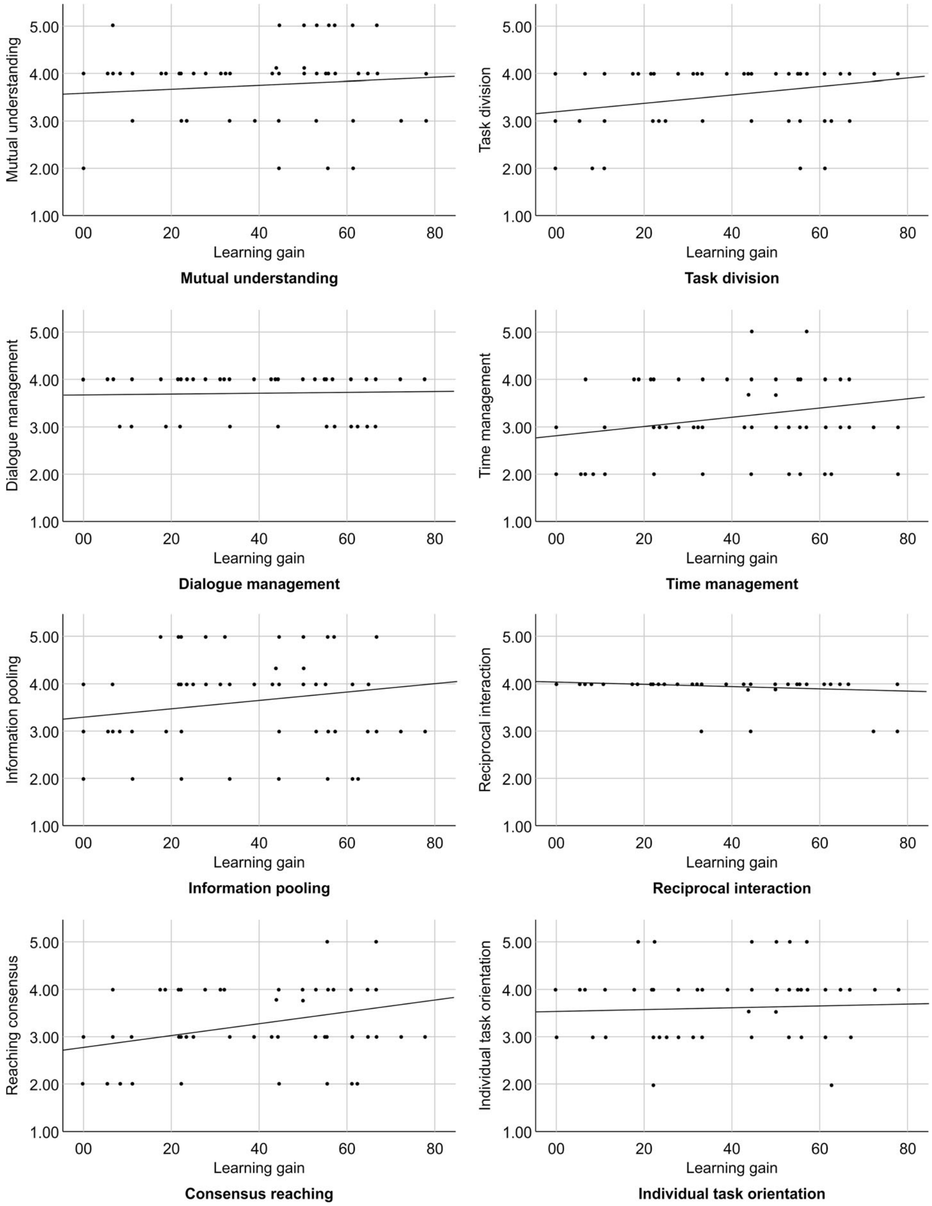


**Fig. 6** Correlation between collaborative process and explanation learning gain for each dependent variable

condition, $R^2$ was 0.030, and the ANOVA F-value was 0.266, again indicating no statistical significance ($p = 0.769$).

The regression analysis results indicate that a significant learning gain was found only for the participants who experienced the visible-gaze/visible-agent condition, in contrast to the other conditions, which did not show significant learning gains. This result supports H2–2 and shows that, when both facilitations are used and the learners successfully implement the collaborative process, they are able to acquire more knowledge on the target concepts.

## Qualitative analysis of the learning process

The results show that interventions of gaze feedback and metacognitive facilitation from the agents facilitated the learning process, especially with respect to "Consensus reaching" and "Task division." Based on these points, a further investigation was conducted to investigate how the two interventions influenced the collaborative discussions. Table 4 shows some of the discussions by condition, and the excerpts that illustrate the process variables. The quotes were selected based on the score of the codes and the author's decision of the dyads that clarity about the collaborative process that was examined, which was about "consensus reaching" and "task division" processes.

As seen in the examples, participants receiving gaze feedback used their partner's gaze information as an integral part of the communication media, which enabled them to abbreviate phrases when they referred to entities on the shared visual space. For example, participants in the visible gaze/no agent condition [task division] exhibited a pattern in which one learner was able to understand where their partner was looking at the time they talked to their partner. Therefore, it can be interpreted that the gaze enabled the learner's to discuss efficiently what to do next, which eventually enabled them to work efficiently within their task divisions. In contrast, learners in the no visible gaze/no agent condition had some momentary times where there was a pause in the interaction. This could have benefited from interventions of gaze awareness allowing them to know if their partner had finished reading their part so they could have proceeded more efficiently. In the consensus reaching dialogues, the visible gaze helped learners understand what was written within each other's blurred area, which helped their inference in understanding what should be treated as part of the common ground (visible gaze/no agent condition).

As seen from the examples of the visible gaze/visible agent in the [task division], the use of PCA interventions has been demonstrated to help learners at the meta-level to further reconsider how to coordinate successfully. This shows that the external support offered by the PCA helped learners to think and realize how to make efficient use of the gaze information to further understand what their partner was referring to. In the visible gaze/agent conditions, there was compelling evidence on the synergy of the use of the two technologies. Participants in this condition were using their gaze as "pointers" to gesture to some of the words mentioned from the PCA. In the [task division] case, they referred to the word "switch turns" to clarify some of the statements that the PCA had mentioned. Such activity reduces ambiguity by directly pointing at their reference point. This kind of interaction strategy only occurs when the two tools are used together, allowing an increase in the accuracy with which the meaning of utterances can be understood. This pattern was never observed in the no visible/no agent condition.

t4.1 **Table 4** Qualitative analysis on learning process of all four conditions

| | | | | |
|---|---|---|---|---|
| t4.2 | no visible gaze/no agent | visible gaze/no agent | no visible gaze/visible agent | visible gaze/ visible agent |
| t4.3 | [task division] | | | |
| t4.4 | A:I think language processing is mostly influenced not only by the knowledge but also the expectations acquired from past experience. | A: So, I see you are now reading your description now. Should we first read by our self and then explain about it after we finish **[gaze moving on left hand side to partner]** | PCA: Try to **change roles and switch turns** after explaining one concept. [TypeB] | PCA: Try to **change roles and switch turns** after explaining one concept. [TypeB] |
| t4.5 | B:Are you reading your concept description is that your interpretation or guess? | B: Good idea. | A: Should we read the description first and then switch turns one by one? | A: As the agent says here **[gaze moving toward the PCA's text box] shall we follow this?** [gaze moving on PCA] |
| t4.6 | A: Oh, I was just reading my concept and telling you what is written here | ...[pause. Each are reading] | B: I agree, let's do so and let's start reading it aloud. | B: I see that part and agree to do this. **[gaze pointing at the word "switch turns"]** |
| t4.7 | B: So maybe we just read and tell me when you finished reading it? | A: So, **it seems that you are not reading it anymore**, it seems you finished could you start explaining? **[gaze moving on left hand side to partner]** | A: Yes, please do so and **let's switch turns** and I will when you finish | A: Yea you are already reading yours, **why don't you start first**, and I will try to decode what is written in this blurred area. [gaze moving on left hand side to partner] |
| t4.8 | A: yes | B: Absolutely, I was finished and was curious about what is written on your part, so I was looking there. [gaze moving on right hand side to partner] | ... [A and B explaining] | ... [A and B explaining] |
| t4.9 | ...[pause. Each are reading] | | PCA: Good job! [TypeD] | PCA: Keep going! [TypeD] |
| t4.10 | B: So, are you finished? | | | |
| t4.11 | A: Oh, yea quite a long time ago… Sorry about that. | | | |
| t4.12 | [consensus reaching] | | | |
| t4.13 | A: So, do you know what we should do about integrated perspective? | A: So, **did you say** that there is a description about finding a schema is part of the processing in bottom-up? [gaze moving on right hand side to partner] | PCA: **Think about the relations** between the two different concepts. [Type C] | PCA: Try to **consider the concept you are now explaining by using other examples**. [Type B] |
| t4.14 | B: As far as what I understood from what you have said, your concept seems to have a stronger relation with the topic? | B: **Yes, it says** that ......**[gaze moving on right hand side on the word schema]** | A: I remember you were mentioning that the information is processed by data driven? | A: So, I think we now know what the two concepts are but what do you think about this **[gaze moving on PCA's comment ``examples"]** |

t4.15 **Table 4** (continued)

| | no visible gaze/no agent | visible gaze/no agent | no visible gaze/visible agent | visible gaze/ visible agent |
|---|---|---|---|---|
| t4.16 | A: I don't know, did you think so? Should I repeat what I read again? | A: Oh **I see that word is located on the last sentence** [gaze moving on the same area], so it seems that the use of schema is in the final stage of processing. | B: Yes, perceptual information comes first and then tries to find what it means. | B: So, **examples**. Well, what about when we read a word that is handwritten and there are letters that are hard to read. **[gaze moving on PCA's comment and left-hand side to partner]** |
| t4.17 | B: Maybe... There is no way these two concepts can merge. Let's just work on this individually it might be better work on our own. | B: Yes, that right. And you said that human memory was related to top down processing at **the beginning of yours [gaze moving on left hand side on the word schema]** | A: So, this may be opposite, because top down processing is based on knowledge from past experience and it tries to look at the word using the knowledge. | A: Yea, that is a good one. Bottom-up process starts from data analysis so dirty handwritten words cannot be decomposed? Was this decomposition mechanism written in the begging on your description? [gaze moving on right hand side to partner] |
| t4.18 | | A: **It says** that ... **[gaze moving on left hand side on the word memory]** | B: I see. So, **these two concepts relate** in terms of both looking at the data and knowledge but process them differently? | B: Yes, it was around here [gaze moving on the right-hand side and on the words composed]. People cannot understand the letter but can somehow infer from the other letters what words is written by the context of the other letters. |
| t4.19 | | B: **So, there might be a connection between these two concepts**, in terms of using knowledge from experience? [gaze moving on left to right] | A: Yea, I think that is **the relations between the two concepts**. | A: **I agree with that example.** This is a good explanation of how both concepts are process all together. [gaze moving on right hand side to partner] |
| t4.20 | | A: I agree and with that point and let's continue to think on that more. [gaze moving on right hand side to partner] | PCA: Yes, discussing about relations and concepts are good. Keep on working on it. [TypeD] | PCA: Good job you are using examples. Keep explaining! [TypeD] |

A and B refer to each participant. The bold words indicate the important phrases and keywords and gaze movement. The detected keywords from the agent are underlined and the type of suggestions are in []



## Discussion

This study investigated the effects of two types of collaboration support technology, namely gaze feedback and PCA-based metacognitive suggestions, on the learning process and learning performance of peer collaborative learners. The effective use of awareness technology that directly supports speakers in remote environments has been investigated in the field of HCI, along with the use of third-person support, such as conversational agents, to facilitate human-human interaction. CSCL studies have used these technologies to support learner-learner collaborative learning activities; however, no study has investigated the aspects or their effects investigated in this paper. Only a few studies have investigated the effects of gaze awareness technology and pedagogical agents on activity coordination.

In the present study, a conversational analysis was performed regarding coordination processes and their influence on learning performance. It was hypothesized that gaze feedback effectively facilitates coordination activities (H1a-1) and learning gains (H1a-2). It was also predicted that the use of PCAs and metacognitive suggestions facilitates these processes (H1b-1) and learning gains (H1b-2). Synergetic effects on these processes (H1c-1) and performance (H1c-2) were also predicted. The results showed that both gaze feedback and PCA use effectively facilitate the collaborative process, supporting H1a-1 and H1b-1. However, gaze feedback had a greater influence on the collaborative process compared to PCA use. In terms of learning gains, gaze feedback and PCA use are more advantageous than the use of gaze feedback alone (supporting H1c-2).

Further investigation of the relationship between the learning process and learning gains showed that pairs that successfully performed coordination processes also exhibited superior performance in terms of the learning gain (H2–1). Moreover, the results indicated that learners who successfully used both gaze feedback and the PCA exhibited a stronger relationship between the learning process and learning performance. Thus, the combination of both considered facilitation methods may produce better learning opportunities (H2–2). As expected, it was found that verbal suggestions from the PCA complement the collaborative process, which the gaze awareness tools do not support.

## Collaboration process

This subsection discusses the results obtained for the two facilitation methods, which were expected to yield a successful coordination process. The conversational analysis results show that learners who received gaze feedback as a byproduct of their peers' behavior successfully coordinated with their partners compared to pairs who did not have the benefit of this method. More specifically, significantly higher scores were obtained for activities such as "dialog management," "information pooling," "consensus reaching," "task division," and "time management" when gaze feedback was used. This indicates that the reference to their partner's gaze was useful for the learner in terms of successful communication (consensus-building and pooling information). Moreover, gaze feedback may produce social awareness of the partner. This can be expected to cause participants to become more responsible in terms of task participation and, thus, influence dependent variables such as task management, task division, and time management. The effects of PCA use to facilitate the collaborative process were only apparent in the context of conversations, where the representation of the "consensus reaching" and "task division" codes displayed the impact. Thus, third-person facilitations has the ability to aid the collaborative process. However, comparing the number of significant results for PCA use with those for gaze feedback use, as detailed in Table 3, the value for the latter was found to be more than twice as high as that for the former. This indicates that the gaze feedback method was in some sense more effective in facilitating coordination activities than the PCA approach. Why was the effect of PCA interventions relatively limited compared to gaze awareness? One of the reasons might be the type of facilitation prompts that were used. This study used metacognitive suggestions to activate self-regulation of communicative behaviors. However, these metacognitive suggestions might have been too abstract for some of the learners to realize what kind of conversations they should have. Learners may encounter difficulty with metacognition on communication since they do not have a good model of what kind of conversation will be useful in such a situation. This should be considered in the future by providing examples or providing scripted models of conversations, such as in Rummel et al. (2009).

As mentioned above, awareness tools provide comparatively rich information on individuals engaging in activities by altering the actions of others (Dourish and Bellotti 1992; Schmidt 2002). As seen in the descriptive analysis, shown in Table 4, through gaze feedback use, the learners are able to see and locate the text being read by their partners during their partners' explanations. Thus, listeners are able to trace their partners' gaze and use it as a clue to aid them in building a representative image of the blurred text. Moreover, conversational strategies, such as referring to the same area during conversational conventions, have been examined in communication studies using referential tasks (Richardson and Dale 2005; Richardson et al. 2007). Speakers and listeners tend to refer to the same area to establish common ground; however, there is also an egocentric bias through which people refer to different areas (Keysar et al. 2000). The results of the communication dialog analysis are consistent with the above theoretical implications.

## Influence on learning gain performance

The results of the analysis of the learning gains are discussed below, where the interaction between the two factors was considered (H1c-2). First, it became clear that gaze feedback use has a greater effect only when the PCA is not used. This can be interpreted from the evidence that better performance was obtained for the visible-gaze/no-agent condition than the no-gaze/no-agent condition. Comparisons with previous studies featuring the same experiment conditions show similarities to the works of Schneider and Pea (2013, 2014), in which the effects of using gaze feedback were shown to differ in terms of the learning gains associated with the task. Those researchers used an inference task in which the correct answers were selected from various options, which differs fundamentally from the task considered in this study. The findings of the present study offer new insights, indicating that gaze feedback is even effective in tasks involving explanation-based knowledge integration. In the assigned task, learners were required to change roles and explain their concepts to each other to develop a mutual understanding of each other's perspectives. In addition, they were required to further self-regulate in order to critically think and generate new hypotheses (concepts) to achieve the task goal. Therefore, the results of this study provide further evidence on the effectiveness of gaze feedback in different collaborative learning tasks.

In addition, the results showed that the use of PCA influenced knowledge acquisition regarding concept explanations under gaze feedback conditions. This can be interpreted from the comparison of the gain analysis results for the visible-gaze/no-agent and the visible-gaze/visible-agent conditions. Hence, PCA use appears to add value over not having a PCA. This finding reveals additional advantages compared with the works of Schneider and Pea (2013, 2014), which did not consider a PCA and examined real-time gaze feedback only. The findings of the current study have new implications for gaze feedback studies, particularly that PCA use may have a synergetic effect. The PCA used in this study provided feedback with metacognitive suggestions; thus, it is assumed that this verbal information enabled the learners to abstract their thoughts and, hence, influenced the pre- and post-explanation-task tests.

 



However, the gaze and PCA combination was not completely superior, as apparent from the comparison of the results for the no-gaze/visible-agent and the visible-gaze/visible-agent conditions, which reveal no difference between these conditions. One interpretation of this finding is that the PCA use had a strong influence on the dependent variable. Why was this? It may be because the PCA will allow them to think again about each other's different knowledge and the integrated knowledge. This is not supported by the gaze awareness tool, and only learners receiving such metacognitive suggestions can take advantage of such facilitation and therefore, produce better quality on explanations. Moreover, it is interpreted that reconsidering each other's perspective strongly influences the knowledge integration activity because it requires reflections about each other's knowledge and perspective. Taking this into consideration, it is predicted that the effects of the interventions will appear much stronger on the evaluations on the explanations of the integrated knowledge.

## Limitations and future work

This subsection discusses the limitations of this study and explores directions for future work. The results of this study indicate that the use of gaze feedback facilitates better coordination and communication processes. Although awareness technology using gaze feedback only was considered in this work (Ishii et al. 1993; Monk and Gale 2002), there are many methods of producing awareness that are studied in ubiquitous computing. Many different types of technologies can facilitate different aspects of interactions, enhancing not only social awareness as examined in this study but also cognitive awareness as studied by Janssen and Bodemer (2013). Those researchers defined cognitive group awareness (e.g., acquisition of information on group members' knowledge and expertise) and social group awareness (e.g., acquisition of information on group members' contributions to the group process). Cognitive awareness can be promoted by providing each learner with prior knowledge of the task and of the partners' need for knowledge. In studies using jigsaw-based environments, members are expected to be more decisive. In that sense, such awareness can be expected to enable learners to acquire the knowledge of others, enabling them to quickly establish common ground, as they would not have to ask each other questions to acquire this information. Relevant alerts could be generated automatically during the task by monitoring the learner's utterances and/or providing the learners with information on their partners' knowledge before the task starts. Such awareness could be exploited in a task similar to that performed in this study. However, further investigation is required in order to understand the effect of this cognitive awareness.

This study also found that third-person suggestions are effective in improving learning gains, i.e., a PCA is more useful than gaze feedback in facilitating cognitive information processing. PCA use did not have a significant effect on the coordination process; however, it did facilitate the "consensus reaching" and "task division" processes. Moreover, the results may change if the PCAs are designed such that they provide feedback based on the learner's degree of success in developing common ground. To this end, one approach is for the system to detect the learner's coordination status, for example, by understanding the learner's conversations. Gaze recurrence (Schneider and Pea 2013, 2014) is a potential means of detecting a learner's coordination success. Therefore, the development of a PCA based on these metrics may be useful.

Also, as found in the qualitative analysis results on the learning process, some interesting communication strategies emerged, such as gaze gestures, that were not predicted. Some learners who had a correlational relationship between the learning process and performance showed that they were using the gaze awareness tool as a pointer to refer to the PCA's suggestions. This is efficient as learners were able to deduce some of the phrases, which may be a type of sign system that was developed to aid communication (Galantucci 2005). From an instructional perspective, it is expected that if the participants have experience/training in using these types of strategies, using the gaze as pointers, and presenting multiple suggestions on their screen, learners might experience stronger performance in general and further outperform other conditions. A more focused experiment developing interfaces that enable easier use of this strategy or adding a condition that includes training the learners how to use these two tools most effectively may influence the synergistic effect on the learning gain.

Finally, the generality of this study must be discussed. It may be pointed out that it is difficult to generalize the results owing to the nature of the task and the situation that was proposed in this study. As mentioned earlier in this paper, this study focused on collaborative activities in a computer-mediated scenario where communication channels are limited and members have low awareness, and thus encounter difficulties in understanding each other's different perspectives. This type of collaborative activity may take place in a virtual space, such as in online settings (e.g., learners working on concept learning with members having different backgrounds using video-conferencing systems). Members may be located in different locations and time zones and have varying levels of network availability, rendering smooth communication more difficult. The purpose of the experimental task in this work was to manipulate such a situation with low awareness and with a divergence of perspectives among members. Therefore, as a result, the task used in this study was specific to this situation. It should be acknowledged that not all collaborative learning situations encounter such constraints in communication, and therefore, the effects of the two types of tools could change in different situations. Therefore, it should be stressed that the situation examined was not a general case of collaborative learning, and thus, the results may be limited to this situation where awareness is low and visibility is obscured. Consequently, one important component for future studies is the examination of the effectiveness of the proposed method in different situations and tasks. Nevertheless, this study clearly demonstrates the positive outcomes of combining the two tools—gaze awareness and PCA—for the learning process and learning performance, which suggests that it would be productive for the proposed method to be applied more broadly.



## Conclusion

The convergence of different types of knowledge and perspectives are effective strategies, and dialectal argumentation in peer collaboration helps develop higher levels of understanding between peers. These types of collaborations taking place in a computer-mediated environment with learners with little knowledge on how to collaborate can take advantage of using gaze awareness tools and metacognitive suggestions from a PCA. Past studies using gaze awareness have shown that such tools are effective in facilitating joint attention and thus facilitate the coordinative process. However, providing such information does not support learners on how to regulate appropriate behaviors. Studies in CSCL have investigated the use of scripts and prompts for effective interventions. In these studies, providing facilitation adaptively based on the learner's state is a challenging issue, and recently AI agents have been considered as a potential technological solution. From previous studies, it was unclear if these agents have the ability to support the same quality of collaborative process as technologies that enhance joint attention directly by gaze awareness. This study investigated how metacognitive suggestions from PCAs are able to provide interventions that gaze awareness tools cannot support during knowledge integration activities. The results from the analysis indicate that PCA intervention can be more effective than mutual gaze intervention in terms of learning gains. In contrast, the gaze is relatively more effective at improving the collaborative learning process. This provides implications for how the combinations of the two types of tools can be incorporated in a jigsaw-like task. The findings from this study not only contribute to design principles related to the use of knowledge integration tasks but also for collaborative learning tasks that require both coordination and individual cognitive processing. These findings thus contribute insights towards designing collaborative learning experiences for distant learning technologies and e-learning environments.

Studies in CSCL have worked on developing infrastructures and tools to aid learners working remotely (Ludvigsen and Steier 2019). It has been pointed out that without detailed process-oriented studies that focus on specific features of supporting collaborative learning, it will be difficult to make further progress in this field. This study provides empirical evidence from a laboratory experiment in a controlled setting and provides specific results on the use of each technology. Combining the use of scripts and group awareness tools has been considered an important topic in the Learning Science and CSCL community (Schnaubert et al. 2020). The author believes that future studies can build upon this study by designing new systems, and can solve the challenges in ITS, such as designing systems that can adaptively enable the users to use different types of interventions at the necessary time and task type. This study extends its contributions to studies in CSCL to studies on AI and communication technologies, providing new research questions on how these types of facilitations could be developed for new applications that can be used practically in classrooms. Further investigation can be conducted on designing PCAs that can play the role of a simulated student and provide gaze awareness for this simulated learner to facilitate better quality of collaborations. The example of this hybrid-facilitation method integrating techniques from artificial intelligence is meant as a challenge for the field of CSCL, with the idea of prompting more investigation into how such technologies interact with and affect individuals as well as how they interact and learn together.



**Acknowledgements** This work was supported by JSPS KAKENHI Grant No. 20H04299. I would like to thank the editors, and especially Crolyn P Rosé for all the work. I also appreciate all the reviewers for valuable feedback during the review process.

**Availability of data and material** Not applicable.

**Authors' contributions** Not applicable.

**Funding** This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant-in-Aid for Scientific Research (B), Grant Number JP20H04299.

## Compliance with ethical standards

**Conflict of interest** No potential conflict of interest was reported by the authors.

**Code availability** Not applicable.

**Ethical approval** The study was conducted after institutional ethical review and approval by the ethical reviewing committee of Ritsumeikan University.



## References

Aleven, V., Mclaren, B., Roll, I., & Koedinger, K. (2006). Toward meta-cognitive tutoring: A model of help seeking with a cognitive tutor. *International Journal of Artificial Intelligence in Education, 16*(2), 101–128 Retrieved from http://dl.acm.org/citation.cfm?id=1435344.1435346.

Aronson, E. & Patnoe, S. (1997). The jigsaw classroom: Building cooperation in the classroom *(2nd ed.).* Addison Wesley Longman.

Asterhan, C. S. C., & Schwarz, B. B. (2009). Argumentation and explanation in conceptual change: Indications from protocol analyses of peer-to-peer dialog. *Cognitive Science, 33*(3), 374–400. https://doi.org/10.1111/j.1551-6709.2009.01017.x.

Azevedo, R., & Cromley, J. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*(3), 523–535. https://doi.org/10.1037/0022-0663.96.3.523.

Baylor, A. L., & Kim, Y. (2005). Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education, 15*(1), 95–115.

Belenky, D., Ringenberg, M., Olsen, J., Aleven, V. & Rummel, N. (2014). Using dual eye-tracking to evaluate students' collaboration with an intelligent tutoring system for elementary-level fractions. *Proceedings of the 36th Annual Meeting of the Cognitive Science Society (COGSCI2014),* 176–181.

Biswas, G., Leelawong, K., Schwartz, D., & Vye, N. (2005). Learning by teaching: A new paradigm for educational software. *Applied Artificial Intelligence, 19*(3), 363–392. https://doi.org/10.1080/08839510590910200.

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2007). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition, 106*(3), 1465–1477. https://doi.org/10.1016/j.cognition.2007.05.012.

Buxton, W. A. S., & Moran, T. (1990). EuroPARC's integrated interactive intermedia facility. In S. Gibbs & A. A. Verrijn-Stuart (Eds.), *Multi-user interfaces and applications (pp. 11–34)*. Elsevier Science Publishers.

Chi, M. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*(1), 73–105. https://doi.org/10.1111/j.1756-8765.2008.01005.x.

Chi, M., Leeuw, N., Chiu, M., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439–477. https://doi.org/10.1207/s15516709cog1803_3.

Clark, H, H. & Brennan, S, E. (1991). Grounding in communication. In B. L. Resnick

D'Mello, S., Olney, A., Williams, C., & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of Human-Computer Studies, 70*(5), 377–398. https://doi.org/10.1016/j.ijhcs.2012.01.004.

Dourish, P. & Bellotti, V. (1992). Awareness and coordination in shared work spaces. *ACM Conference on Computer Supported Cooperative Work (CSCW '92).*

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science, 29*, 737–767.

Goodwin, C. (1981). Conversational organization: Interaction between speakers and hearers. Academic Press.

Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2005). Autotutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education, 48*(4), 612–618. https://doi.org/10.1109/TE.2005.856149.

Greeno, G. J., & de Sande, C. (2007). Perspectival understanding of conceptions and conceptual growth in interaction. *Educational Psychologist, 42*(1), 9–23. https://doi.org/10.1080/00461520709336915.

Greiff, S., Martin, R., Koenig, V., Mustafic, M., Herborn, K. & Schweitzer, N. (2017). PISA 2015 collaborative problem solving (CPS) validation study final report (tech. Rep.). *PISA Governing Board (PGB) and Organization for Economic Co-operation and Development (OECD).*

Harley, J. M., Taub, M., Azevedo, R., & Bouchet, F. (2017). “Let’s set up some subgoals”: Understanding human-pedagogical agent collaborations and their implications for learning and prompt and feedback compliance. *IEEE Transactions on Learning Technologies, 11*(1), 54–66. https://doi.org/10.1109/TLT.2017.2756629.

Hayashi, Y. (2012). *On pedagogical effects of learner-support agents in collaborative interaction. Proceedings of the 11th international conference on intelligent tutoring systems (ITS 2012) (pp. 22–32)*. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-30950-2_3.

Hayashi, Y. (2014). *Togetherness: Multiple pedagogical conversational agents as companions in collaborative learning. Proceedings of the 12th international conference on intelligent tutoring systems (ITS 2014) (pp. 114–123)*. Cham: Springer. https://doi.org/10.1007/978-3-319-07221-0_14.

Hayashi, Y. (2016). *Coordinating knowledge integration with pedagogical agents: Effects of agent gaze gestures and dyad synchronization. Proceedings of the 13th international conference on intelligent tutoring systems (ITS 2016) (pp. 254–259)*. Cham: Springer. https://doi.org/10.1007/978-3-319-39583-8_26.

Hayashi, Y. (2018a). Gaze feedback and pedagogical suggestions in collaborative learning: Investigation of explanation performance on self’s concept in a knowledge integration task. Proceedings of the 14th international conference on intelligent tutoring systems (ITS 2018) (lecture notes in computer science) (pp. 78–87). Springer, Cham. https://doi.org/10.1007/978-3-319-91464-0_8

Hayashi, Y. (2018b). The power of a “maverick” in collaborative problem solving: An experimental investigation of individual perspective-taking within a group. *Cognitive Science, 42*(S1), 69–104. https://doi.org/10.1111/cogs.12587.

Hayashi, Y. (2019). Multiple pedagogical conversational agents to support learner-learner collaborative learning: Effects of splitting suggestion types. *Cognitive Systems Research, 54*, 246–257. https://doi.org/10.1016/j.cogsys.2018.04.005.

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review, 6*(1), 27–54. https://doi.org/10.1016/j.edurev.2010.07.004.

Holmes, J. (2007). Designing agents to support learning by explaining. *Computers and Education, 48*(4), 523–547. https://doi.org/10.1016/j.compedu.2005.02.007.

Ishii, H., Kobayashi, M., & Grudin, J. (1993). Integration of interpersonal space and shared workspace: Clearboard design and experiments. *ACM Transactions on Information Systems, 11*(4), 349–375. https://doi.org/10.1145/159764.159762.

Janssen, J., & Bodemer, D. (2013). Coordinated computer-supported collaborative learning: Awareness and awareness tools. *Educational Psychologist, 48*(1), 40–55. https://doi.org/10.1080/00461520.2012.749153.

Jermann, P., Mullins, D., Nuessli, M, A. & Dillenbourg, P. (2011). Collaborative gaze footprints: Correlates of interaction quality. *Proceedings of CSCL 2011* (pp. 184–191).

Keysar, B., Barr, J. D., Balin, A. J., & Brauner, S. J. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science, 11*(1), 32–38. https://doi.org/10.1111/1467-9280.00211.

Khedher, A. B., Jraidi, I., & Frasson, C. (2017). *Assessing learners’ reasoning using eye tracking and a sequence alignment method. Proceedings of the 13th International Conference on Intelligent Computing (ICIC 2017) (pp. 47–57)*. Cham: Springer. https://doi.org/10.1007/978-3-319-63312-1_5.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8*, 30–43.

Kumar, R., & Rosé, C. (2011). Architecture for building conversational architecture for building conversational agents that support collaborative learning. *IEEE Transactions on Learning Technologies, 4*(1), 21–34. https://doi.org/10.1109/TLT.2010.41.

Levine, M, R. & Teasley, D, S. (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). APA Press. https://doi.org/10.1037/10096-006, 1991.

Ludvigsen, S., & Steier, R. (2019). Reflections and looking ahead for CSCL: Digital infrastructures, digital tools, and collaborative learning. *International Journal of Computer-Supported Collaborative Learning, 14*, 415–423.

Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *International Journal of Computer-Supported Collaborative Learning, 2*(1), 63–86. https://doi.org/10.1007/s11412-006-9005-x.

Miyake, N. (1986). Constructive interaction and the iterative process of understanding. *Cognitive Science, 10*, 151–177.

Monk, A. F., & Gale, C. (2002). A look is worth a thousand words: Full gaze awareness in video mediated conversation. *Discourse Processes, 33*(3), 257–278. https://doi.org/10.1207/S15326950DP3303_4.

Nye, B. D., Graesser, A. C., & Hu, X. (2014). Autotutor and family: A review of 17 years of natural language tutoring. *International Journal of Artificial Intelligence in Education, 24*(4), 427–469. https://doi.org/10.1007/s40593-014-0029-5.

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.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination. *Psychological Science, 18*(5), 407–413. https://doi.org/10.1111/j.1467-9280.2007.01914.x.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *Journal of the Learning Sciences, 2*(3), 235–276. https://doi.org/10.1207/s15327809jls0203_1.

Rummel, N., Spada, H., & Hauser, S. (2009). Learning to collaborate while being scripted or by observing a model. *International Journal of Computer-Supported Collaborative Learning, 4*(1), 69–92. https://doi.org/10.1007/s11412-008-9054-4.

Salomon, G. (2001). *Distributed cognition: Psychological and educational considerations*. New York: Cambridge University Press.

Schmidt, K. (2002). The problem with 'awareness': Introductory remarks on 'awareness in CSCW'. *Computer Supported Cooperative Work, 11*(3–4), 285–298. https://doi.org/10.1023/A:1021272909573.

Schnaubert, L., Vogel, F., Bodemer, D., Fischer, F., Radkowitsch, A., Schmidmaier, R., Fischer, M., Tsovaltzi, D., Puhl, T., & Azevedo, R. (2020). Combining scripts, group awareness tools and self-regulated learning-theoretical implications and predictions and practical implementations. *Proceedings of CSCL, 2020*, 350–357.

Schneider, B., & Pea, R. (2013). Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *International Journal of Computer-Supported Collaborative Learning, 8*(4), 375–397. https://doi.org/10.1007/s11412-013-9181-4.

Schneider, B., & Pea, R. (2014). Toward collaboration sensing. *International Journal of Computer-Supported Collaborative Learning, 4*(9), 5–17. https://doi.org/10.1007/s11412-014-9202-y.

Schwartz, D. L. (1995). The emergence of abstract representation in dyad problem solving. *Journal of the Learning Science, 4*, 321–354. https://doi.org/10.1207/s15327809jls0403_3.

Shirouzu, H., Miyake, N., & Masukawa, H. (2002). Cognitively active externalization for situated reflection. *Cognitive Science, 26*, 469–501.

Sproull, L., and Kiesler, S. (1991). Connections: New ways of working in the networked organization. MIT Press.

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rosé, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science, 31*(1), 3–62. https://doi.org/10.1080/03640210709336984.

Vygotsky, L. S. (1980). *The development of higher psychological processes*. Harvard University Press.

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

## Affiliations

**Yugo Hayashi**[1]

✉ Yugo Hayashi
yhayashi@fc.ritsumei.ac.jp

[1] College of Comprehensive Psychology, Ritsumeikan University, Ibaraki, Japan