From Monitoring to Sharing of Attention in Dyadic Interaction: The Affordances of Gaze Data to Better Understand Social Aspects of Remote Collaborative Problem Solving

Johanna PÖYSÄ-TARHONEN[a*], Nafisa AWWAL[b], Päivi HÄKKINEN[a] & Suzanne OTIENO[c]
[a]*Finnish Institute for Educational Research, FIER, University of Jyväskylä, Finland*
[b]*Assessment Research Centre, ARC, University of Melbourne, Australia*
[c]*Department of Psychology, University of Jyväskylä, Finland*
*johanna.poysa-tarhonen@jyu.fi

Abstract: This paper aims to better understand the social aspects of collaborative problem solving (CPS) through studying joint attention behaviour (JAB) in an online game–like environment. To capture these behaviours and exemplify how 'jointness' is achieved in CPS in remote dyadic interaction, event-related measures are utilised based on the following multiple interaction data: (1) individuals' gaze data from CPS task completion and (2) automatically generated log files (i.e. chats and actions) from dyadic interactions. The results give empirical evidence of the detached, individualistic attention experiences (i.e. monitoring and common attention) and of bidirectional relations (i.e. mutual and shared attention) in which partners adopt an engaged approach towards one another to solve the task together. It is also observed how lower level attention in CPS can be a precursor to a higher level; that is, during interaction, there is a move from monitoring the partner's actions towards common attention experience. In addition, it is noticed that richer second-person relations may come in degrees. In methodological terms, the gaze data can provide access to better uncover dyadic processes during remote CPS, but without the information embedded in the log data, they would not provide sufficient contextual details of the real interaction to fully understand social connotations related to CPS.

Keywords: collaborative problem solving, joint attention behaviour, live eye tracking, process-orientated research

1. Introduction

Collaborative problem solving (CPS), a key competence of 21st-century learners, is defined as a skill set required for solving problems in non-routine, collaborative situations in different domains (e.g. Funke, Fischer, & Holt, 2018; Graesser et al., 2018; Scoular & Care, 2020; Scoular, Care, & Hesse, 2017). In most frameworks, CPS is based on a socio-cognitive approach to learning, and it is seen to lie in a two-dimensional space of social and cognitive components that intermingle over the processes of problem solving (e.g. Funke et al., 2018; Graesser et al., 2018; Scoular et al., 2017; Zwiecki, Ruis, Farrell, & Williamson Shaffer, 2020). However, as the theorised CPS constructs are relatively new and rooted in individual problem-solving approaches, the social components have not yet been fully covered in the existing CPS models (Funke et al., 2018; Scoular et al., 2017). Therefore, endeavouring to understand how social components function in CPS is an essential step in moving from individual problem solving in a social context to a CPS construct in which the social and cognitive components are more amalgamated.

To better understand CPS and the qualities related to its social components in dyadic interaction, this study applies the concept of *joint attention behaviour* (JAB; e.g. Carpenter & Liebal, 2011; Korkiakangas, 2018; O'Madagain & Tomasello, 2019; Seemann, 2012; Siposova & Carpenter, 2019; Tomasello, 1995), the foundation of any interaction predicting productive collaboration (Barron, 2003). However, despite the growing interest in studying joint attention and its premises to understand dyadic interaction, no unified interpretation exists for what is considered joint attention and how 'jointness' in joint attention is achieved (Carpenter & Liebal, 2012; Seemann, 2012; Siposova & Carpenter, 2019). As a recent viewpoint to better understand the complexity related to joint attention, Siposova and Carpenter (2019) propose an approach to joint attention and social knowledge as a process of closely connected yet distinct phenomena of social attention. They argue that, instead of a single 'state' of joint attention as a binary event (i.e. there is jointness or there is not), jointness may come in degrees.

To capture behaviours and exemplify how 'jointness' and common knowledge in JAB is achieved in CPS in remote settings, *event-related measures* based on multiple interaction data (gaze data, logfiles) are utilised. To do this, the study takes the theorised CPS construct (Hesse, Care, Buder, Sassenberg, & Griffin, 2015) and the unique properties of JAB in dyadic interaction in a remote, game-like CPS assessment environment (Assessment and Teaching of 21st Century Skills [ATC21s], http://www.atc21s.org; e.g. Care, Griffin, & Wilson, 2018; Care, Scoular, & Griffin, 2016; Scoular et al., 2017) as its point of departure. Here, a student collaborates with another student, and the collaborative tasks aim to stimulate and elicit the social and cognitive elements and sub-elements that are part of the complex CPS framework by Hesse et al. (2015). In this environment, the level of participation can be compared with a real social situation with dynamic stimuli (i.e. actionable artefacts) and a chat property as the communication affordance for dyadic interaction. While the automatically generated logfiles as chat and actions of interacting dyads incorporate multiple information of joint processes, to make visible the typology of jointness as defined by Siposova and Carpenter (2019), the gaze patterns—recorded with remote eye trackers from the individual partners—are seen as significant. To better understand JAB in CPS, we expect that focussing on these gaze patterns in parallel with the interaction sequences of the communicating partners identified from the activity log data will help us 'go beyond' these sequences (Korkiakangas, 2018) and better understand CPS in this regard.

A Typology of 'Jointness' in Joint Attention Behaviour

Generally, joint attention is defined as a capacity to focus together with another on an external source or object in the environment (e.g. Eilan, 2005; Korkiakangas, 2018; O'Madagain & Tomasello, 2019). According to Siposova and Carpenter (2019), the objects of joint attention can be diverse sensory inputs, such as visual or auditory stimuli, or they can be present, past or future events or even mental states (i.e. ideas, plans). Given that gaze following is viewed as a promising basis of JAB (Seemann, 2012), it is also seen to include the coordination aspect of joint attention and the sharing of attention (Carpenter & Liebal, 2012; Tomasello, 1995). In its richest definition, individuals must equally recognise that they are attending to the same thing (O'Madagain & Tomasello, 2019; Siposova & Carpenter, 2019; Tomasello, 1995). Thus, following Carpenter and Liebal (2012), it is only communication that 'turns mutually experienced event into interaction, into something joint' (p. 168).

To better understand the complexity and multiple definitions of JAB and how it is achieved, Siposova and Carpenter (2019) argue that joint attention should not be considered a single state but as a process comprising various closely connected but distinct phenomena that can be discovered in related literature. Accordingly, Siposova and Carpenter (2019) have developed a spectrum of 'jointness' described as 'a typology of social attention and social knowledge' that aims to cover the diversity of the existing definitions. The typology comprises four 'levels' (basic components of JAB)—monitoring, common, mutual and shared—that all include the notion of a triadic relationship between self, other and an object or their attention. As a precondition for each of the four levels of social attention is the individual's ability to engage in individual attention.

The first level in the spectrum of jointness is called *monitoring attention* (Siposova & Carpenter, 2019). This refers to a situation in which an individual is taking an observer's perspective on a second individual involved, and in this way, attending to the same matter to which the partner is attending. At this level, the participants have individual knowledge of the situation, and their attention levels are independent; at the same time, an individual has knowledge that the other participant is paying attention to the same object or situation. Nevertheless, although both individual participants are simultaneously monitoring each other's attention to the object or situation, they still assess the attention and knowledge states of the other participants individually. Often, monitoring behaviour is observable, such as turning one's gaze or bodily orientation, but such behaviour can also be present without easily noticeable actions.

At the second level, *common attention*, two individual participants take an observer's perspective, and nearly simultaneously, attend to what the other is focussed on (Siposova & Carpenter, 2019, p. 262). Here, individuals not only attend to the same object or situation but are also *attending to each other's attention* to the object or situation. Engaging in common attention requires that the object of attention is pronounced and marked; that is, the participants can both assume that they are attending to the same object or situation. In addition, they have a reason to consider other participants' attention; for example, they have *a predefined common goal to be achieved*, and in this respect, for both participants, the other individual's attention is relevant. Siposova and Carpenter (2019) point out that, 'under these conditions individuals could know they are attending to each other's attention without any contact or communication' (p. 262). The dependency of the other at this attention level is based on the awareness that they are both engaging in the same attention processes. Yet, notably, *the evaluation of whether they are in common attention is based on an individual's perspective*, and thus, it may not be correct.

According to Siposova and Carpenter (2019), at the third and fourth levels of social attention, *mutual* and *shared attention*, the observer's attitude toward the other and his/her attention no longer exists (i.e. a third-person experience), but the experience is based on direct commitment to the other, where the participants are both senders and receivers of the information (i.e. a second-person experience; see also Zahavi, 2015). Through direct social interaction, each participant becomes a 'constituent part' of the experience of the other (Zahavi, 2015), and attention to an object or situation is coloured by mutual awareness of each other's attention (Siposova & Carpenter, 2019). This bidirectional nature makes the experience different if compared with monitoring and common attention levels that are individualistic (Siposova & Carpenter, 2019). Thus, in *mutual attention*, the participants are more or less simultaneously attending to the same object or situation but not necessarily communicating intentionally (Siposova & Carpenter, 2019). If compared with common attention, at this level, their experience is co-created.

The fourth level of social attention, *shared attention*, meets the qualifications of mutual attention, but this level also requires the participants to *deliberately communicate* with each other about the object or situation and/or the fact that they are sharing attention to it (Siposova & Carpenter, 2019). Thus, what makes shared attention different if compared with mutual attention is its intentional nature. Shared attention is characterised by behaviours in which individual participants verify to each other they are attending to the same object or situation; such behaviours are not necessarily verbal actions, and they can also be 'communicative' and sharing looks (Carpenter & Liebal, 2012) or gestures, such as pointing and showing (Siposova & Carpenter, 2019). To conclude, the lower attention levels (monitoring and common) include third-person perspectives; that is, the participants are *individually attending to the same thing*. The two higher attention levels (mutual and shared attention), in turn, include a second-person relation, which means that the participants are *jointly attending to the same thing*. (For an overview of the sliding scale of jointness, as Siposova and Carpenter [2019] call it, see Figure 1.)

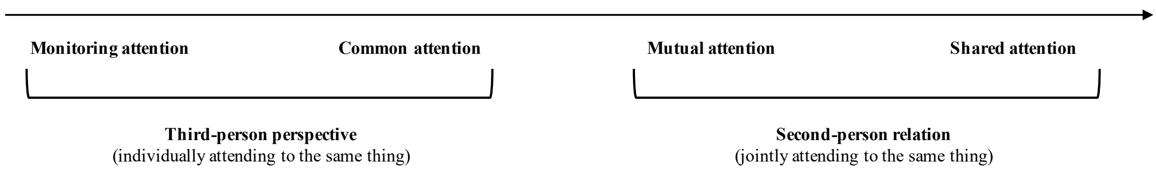


Figure 1. A scale of jointness in joint attention behaviour (JAB). Modified from Siposova and Carpenter (2019).

2. Operationalisation of Joint Attention Behaviour During Collaborative Problem Solving

Basically, JAB can be divided into two types of behaviour—*initiating* joint attention and *responding* to joint attention (e.g. Mundy & Newell, 2007). In social interaction, *immediacy* is the key: normally, there is a tight timeframe within which communication occurs (or does not occur), such as a 2-second window, defined as a sequence consisting of recognisable initiating and responsive actions (Korkiakangas, 2018). Accordingly, in productive CPS, it is expected that a well-performing dyad would organise their efforts in serial sequences of discussion (chat) and actions (manipulating artefacts), consisting of initiating and responsive actions with their partner. However, in CPS situations when students first tackle a new problem-solving task, they are not only expected to represent and refer to 'relevant' parts of the problem but also to somehow ensure that the partner understands what has been referred to (Zemel & Koschmann, 2013). Thus, the placement of gaze in relation to these other activities is an important measure to better understand the properties of JAB during CPS, especially how the 'jointness' and common knowledge is developed and materialised here.

3. Research Questions

In this paper, we ask the following questions:

1) How are 'jointness' and common knowledge in JAB materialised in dyadic interaction in a remote CPS context? What is the meaning of gaze in understanding JAB here?
2) Do the unique properties of remote dyadic interaction restrict/require JAB to arise, and are some 'degrees' of jointness more evident and valuable in CPS?

4. Methods

5.1 Dual-Space Interaction Environment as the Context of Study

The environment, as a 'dual-space' interaction space (Zemel & Koschmann, 2013), encompasses a chat property as a free form, synchronous interface, and a space with actionable artefacts that have either a symmetrical or asymmetrical outlook for the individuals. In a symmetrical task, stimulus content and actionable artefacts are equal for the partners, whereas in an asymmetrical task, the dyad is given a unique subset of resources for problem solving; alternatively, the screen view can be identical, but the permit to move certain objects or scroll the bars is divided between the partners. The success of one student depends on the behaviour of the other and the reactions offered (Care et al., 2016).

5.2 Participants, Study Setup, Data

The data were collected in a live eye-tracking situation (e.g. Korkiakangas, 2018; Dindar, Korkiakangas, Laitila, & Kärnä, 2017) from two student dyads from an initial teacher education programme in a Finnish university. In the study, dyads were physically located in separate cognitive labs. While completing the tasks in dyads, their eye movements were recorded with desktop eye trackers (screen-based; SensoMotoric Instruments [SMI] RED 250 Mobile), and a chin rest at a 60-cm viewing distance was used. A (13-point) calibration was conducted prior to the experiment and before each task. The dataset includes recorded observational data as activity logs from the online environment (consisting of time-stamped information of the movement of artefacts [i.e. actions], as well as the dyadic interactions via free-form chat interface; see Care, Griffin, Scoular, Awwal, & Zoanetti, 2015), combined with gaze data. In the paper, the focus is on a symmetrical task titled 'Laughing Clowns' (e.g. Care et al., 2015), in which, unknown to them, the students are presented each with a clown machine and 12 balls to be shared (see Figure 2 for a screen view of Student B). The screen views of Students A and B are mirror images of each other's, where both can view which balls are being used by the partner but are unable to see how it is being used (i.e., the drop position of the ball in the clown's mouth or the exit point when it comes out). In other words, the visual information that gets transmitted in real-time is the number of the balls used by partners and the location of the ball being used. However, the trajectory of the ball when in use by their partner is not visible to the other student. The students must place the balls into their clown's mouth while the mouth is moving to determine the rule governing the direction the balls will go (entry: left, middle, right; and exit: positions 1, 2, 3).The goal is to determine whether their clown machines work in the same way. To do this, the dyad needs to share information and discuss the rules, negotiating how many balls they should each use. In this regard, communication via the chat interface is central to succeed in this task.

5.3 Data Analysis: Focussing on Event-Related Measures

In eye-tracking studies, the predominant focus has been on the overall looking times at predefined areas of interest (AOIs; 'where' questions; Falck-Ytter, Bölte, & Gredebäck, 2013). Yet, when studying social interaction, this may not be sufficient (Falck-Ytter et al., 2013). Combining *'where'* people look with *'when'* they look at the AOIs, that is, the timing of gazing, is critical to understand JAB as sequential interaction at the pair level during CPS processes. These event-related gaze measures, focussing on the interactional *organisation of gaze*, are more informative about what makes some instances of gazing 'social' (e.g. Dindar et al., 2017; Korkiakangas, 2018). In the current study, with the challenging dynamic scene of the dual-space remote environment, there were multiple behaviours of interest linked to JAB, such as gazing at the chat window, the actionable artefacts and the instructions. For analysing the gaze data, video exports with fixations and scan path data views using BeGaze software were produced. Fixation means maintaining the gaze on a single location, whereas a scan path data view shows gaze positions and eye events plotted on the stimulus video. (For an example of a video export as a scan path view and the AOIs of the symmetrical Laughing Clowns task, see Figure 2.)

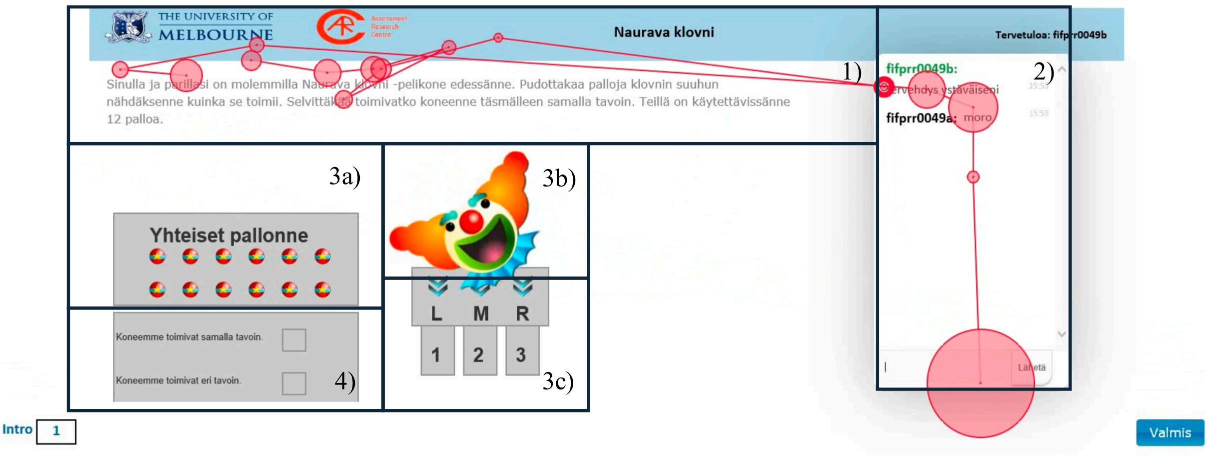


Figure 2. Image from a video export (scan path data view) of the Laughing Clowns task with areas of interest (AOIs).

Note. The example of the AOIs include the following: 1) the instructions; 2) the chat property; 3) actionable artefacts, such as (a) the shared balls, (b) the clown's head/mouth and (c) the 'issue' of the dropped balls as letter-number combinations; and 4) the solution.

To better understand JAB in terms of the typology of 'jointness' (Carpenter & Siposova, 2019) and to search for related behaviour during CPS, a manual qualitative coding procedure was applied. *First*, the focus was on the pair-level activity log data to code for initiating and responsive activities (i.e. chat and actions of Student A and B) during the CPS process. The aim was to search for meaningful *interaction events* in dyadic interaction with regard to JAB in CPS. *Second*, to better understand the triadic interaction (i.e. relationship between self, other and an object or their attention) during problem solving, these interaction events were further analysed in parallel with the eye-tracking video exports that show the fixations and scan path data views of the individuals while completing he tasks. These videos make visible the location and the order of the gaze cursor at specific AOIs during these selected events. At this point, the levels related to the spectrum of 'jointness', described as 'a typology of social attention and social knowledge' by Siposova and Carpenter (2019, p. 261), were identified from the data. It was assumed that these two data types (log files, gaze data), when analysed for consistency, would allow for better understanding of the interaction events in this regard. In addition, it was assumed that, in the remote environment, gaze data would show important moment(s) related to JAB during CPS, composed without writing or moving artefacts, comprising the following: (a) what the log file can show and (b) how the gaze data may reinforce or modify the initial interpretation of the interaction, based on the log file only. (For the different phases of analysis, see Table 1.)

Table 1. *The Different Phases of Analysis of the Event-Related Measures of Joint Attention Behaviour (JAB) During Collaborative Problem Solving*

Phase of the analysis	Data source/level	Manual coding procedure	Target, expected output
Phase 1	Pair-level activity log data	To code the initiating and responsive actions on the activity log file	To search for sequential characteristics of the log data (chat and actions) as meaningful events related to joint attention behaviour (JAB) in collaborative problem solving (CPS)
Phase 2	Gaze data, individual level (Students A and B), meaningful events (pair level) related to JAB in CPS, chosen at the previous phase	To code the location and order of the gaze cursor at specific areas of interest (AOIs) during the selected events (to analyse for congruence; frame-by-frame analysis)	To identify whether and how data reinforce or modify the initial interpretation of the interaction, based on the log file only, to better understand the interactional event at the pair level

Note. The table includes the different phases of the analysis, description of the manual coding procedure and the target or expected outcome(s).

5. Results

The first phase of the analysis provided a general overview of the sequential characteristics of dyadic interaction (i.e. the individual partner's initiating and responding actions; actions, e.g. as the manipulations of the artefacts) and gave evidence of diverse orientations of the two dyads when completing the following tasks: (a) synchronised activity (Pair 1) and (b) parallel actions (Pair 2) (e.g. Pöysä-Tarhonen, Care, Awwal, & Häkkinen, 2017, 2018). In short, when focussing on the log stream of Pair 1, their interaction was well coordinated (Miles, Lumsden, Flannigan, Allsop, & Marie, 2017), and the contents and the general interactional organisation reflected the designed, task-specific CPS elements of the Laughing Clowns task, such as interaction, audience awareness, reciprocity and systematicity (see Care et al., 2015). In turn, when focussing on the log stream of Pair 2, the contents and the general interactional organisation lacked evidence of the majority of the designed, task-specific CPS elements. Dyadic interaction was parallel, encompassing autonomous actions of individuals, without systematicity or coordination, or as 'trial–error' actions (e.g. Davis et al., 2015). At the second phase, the identified patterns were analysed in greater depth with the gaze data, and resulted in illustrative examples that exemplify the spectrum of jointness (Siposova & Carpenter, 2019) in CPS, presented as follows: from individual and monitoring attention (Figure 3) to common attention (Figure 4); and from mutual attention to shared attention (Figure 5). The examples (see Figures 3–5) rely on both gaze and log stream data and are derived from two dyads accomplishing the task.

6.1 Individual and Monitoring Attention During Collaborative Problem Solving

Figure 3 exemplifies an individual (Student B) *monitoring attention* in the CPS situation, observed from the onset of the Laughing Clowns task. As typical of a monitoring situation (Siposova & Carpenter, 2019), Student B has individual knowledge of the situation and evidence of the partner, and via the screen, Student B attends to what the partner (Student A) is attending to. While dragging and dropping a ball, Student B takes an observer's perspective on the actions of Student A; there are no communications yet, but Student B frequently monitors the screen and the interaction property. In this case, Student A is simultaneously in *individual attention* (Siposova & Carpenter, 2019*)*, concentrating on reviewing the instructions and testing the machine individually without any monitoring or communication via the chat property.

(a) Log stream data view (a dyad level):

B	**action**	**startDrag:ball1:70:150**	6.3.2019 23:34
B	**action**	**stopDrag:ball1:519:142**	6.3.2019 23:34
B	**action**	**dropShuteL:ball1:519:142**	6.3.2019 23:34
A	action	startDrag:ball2:396:150	6.3.2019 23:34
A	action	stopDrag:ball2:148:138	6.3.2019 23:34
A	action	dropShuteR:ball2:148:138	6.3.2019 23:34

(b) Gaze data view (an individual level, **Student B**): **(c) Gaze data view** (an individual level, **Student A**):

Figure 3. Illustration of a monitoring attention experience (Student B perspective) and individual attention experience (Student A perspective) during completing the 'Laughing Clowns' task.

Note. The example includes simultaneous moment from (a) the log data, combined with (b) and (c) individual level screen captures from the eye-tracking video exports.

6.2 Common Attention During Collaborative Problem Solving

As in monitoring attention, in *common attention* (see Figure 4), the experience is primarily *individual* (Siposova & Carpenter, 2019). Although working in parallel during CPS task completion and *without systematically communicating over their related goals*, Student A and B have the following characteristics: (a) they have an established joint objective, acquired via task instructions, and thus, it can be assumed that (b) their attention is relevant to the partner (Siposova & Carpenter, 2019). Here, the dyad is engaging in the same CPS situation; while they depend on the attention of the partner, their evaluation of the situation (common attention or not) is individual. In the log stream, the partners share their first notions as follows: Student B writes, *'The first ball went into L'*, and continues to test another ball without any further negotiation; immediately afterward, Student A writes, *'The same and the head was left'*. The communication is based on reporting parallel efforts that rely on individual partners testing the task-specific properties.

(a) Log stream data view (a dyad level):

A	action	startDrag:ball3:442:150	6.3.2019 23:34
A	action	stopDrag:ball3:142:141	6.3.2019 23:34
A	action	dropShuteL:ball3:142:141	6.3.2019 23:34
B	**chat**	**The first ball went into L**	6.3.2019 23:34
B	action	startDrag:ball4:208:150	6.3.2019 23:34
B	action	stopDrag:ball4:510:138	6.3.2019 23:34
B	action	dropShuteR:ball4:510:138	6.3.2019 23:34
A	**chat**	**The same and the head was on left**	6.3.2019 23:34

(b) Gaze data view (an individual level, **Student B**): **(c) Gaze data view** (an individual level, **Student A**):

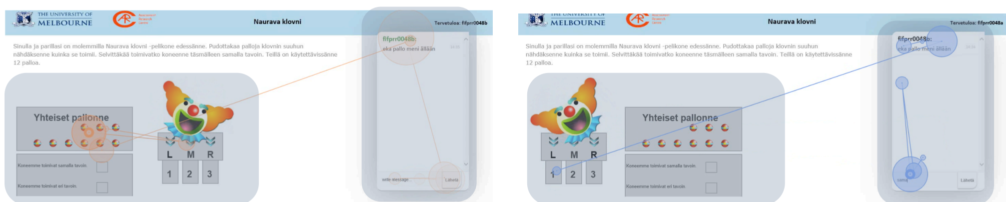


Figure 4. Illustration of a common attention experience (Student B and A perspectives).

Note. The example includes simultaneous moment from (a) the log data, combined with (b) and (c) individual level screen captures from the eye-tracking video exports.

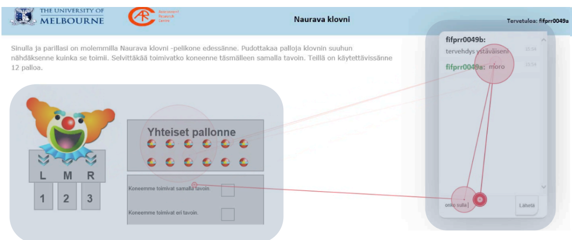
6.3 From Mutual Attention to Shared Attention During Collaborative Problem Solving

As in the third-person perspectives of individual and common attention, two individuals 'meet in the middle' (Siposova & Carpenter, 2019, p. 262; see also Carpenter & Liebal, 2011), in the second-person relations of mutual and shared attention a 'meeting of minds' occurs, and the partners are truly attending *with* each other towards the shared goal or object of attention (Siposova & Carpenter, 2019). In the example (see Figure 5), at the onset of the task, the partner's presence is verbally acknowledged—an attention contact is made. This can also be referred to here as a sign of *mutual attention* experience: In the remote environment, only verbal signs are available; the eye contacts typical of mutual attention in a physical environment are not possible. When proceeding with the task, both students' communicative exchanges about the task and the task properties are intentional and bidirectional. In addition, over the course of CPS (except the first ball thrown by Student B without first consulting the partner, but Student B coming back to the issue later; see Figure 5), the partners co-create their experiences by constantly sending and receiving information and negotiating how to solve the problem together: The partners are seemingly engaged in 'doing together' as *shared attention* experience (Siposova & Carpenter, 2019; Zahavi, 2015). (The gaze data examples represented in Figure 5 are from the moment when the dyad explores the available artefacts and communicates on whether they both have two rows of six balls.)

(a) Log stream data view (a dyad level):

B	chat	**Greetings my friend**	7.3.2019 0:53
A	chat	**Howdy**	7.3.2019 0:53
B	action	startDrag:ball1:70:150	7.3.2019 0:53
B	action		7.3.2019 0:53
B	action	dropShuteR:ball1:535:124	7.3.2019 0:53
A	chat	**Do you have two rows of balls 6 pcs?**	7.3.2019 0:54
B	chat	**Yes I have**	7.3.2019 0:54
B	chat	**I dropped the first ball**	7.3.2019 0:54
A	action	startDrag:ball7:350:196	7.3.2019 0:54
A	action	stopDrag:ball7:152:131	7.3.2019 0:54
A	action	dropShuteR:ball7:152:131	7.3.2019 0:54
B	chat	**I had right**	7.3.2019 0:54
A	chat	**I did so, too**	7.3.2019 0:54
A	chat	**I had left**	7.3.2019 0:54
B	chat	**I guess we have to try if that head posture makes any difference**	7.3.2019 0:54
B	chat	**Does the clown's head move horizontally for you**	7.3.2019 0:54
B	chat	**so like tilting**	7.3.2019 0:55
A	chat	**I'll drop a new ball so that the head is on left**	7.3.2019 0:55
A	chat	**and yes it moves all right**	7.3.2019 0:55
A	chat	**as a pendulum**	7.3.2019 0:55
A	action	startDrag:ball2:396:150	7.3.2019 0:55
A	action	stopDrag:ball2:147:134	7.3.2019 0:55
A	action	dropShuteR:ball2:147:134	7.3.2019 0:55
B	chat	**aa'**	7.3.2019 0:55
B	chat	**We have the same balls**	7.3.2019 0:55
A	chat	**Yes we have**	7.3.2019 0:55

(b) Gaze data view (an individual level, **Student A**):



(c) Gaze data view (an individual level, **Student B**):

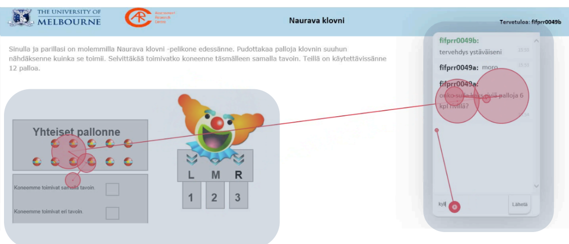


Figure 5. Illustration of sliding from a mutual to shared attention experience (Student A and B perspectives).

Note. The example includes simultaneous moment from (a) the log data, combined with (b) and (c) individual level screen captures from the eye-tracking video exports.

6. Discussion and Conclusions

The results gave empirical evidence of the theorised, third-person, individualistic attention experience (monitoring and common attention), as well as second-person relations (mutual and shared attention; see Siposova & Carpenter, 2019), in which both Students A and B adopted an engaged approach toward each other to solve the task together. In addition, it was observed how a lower level attention experience in CPS can be a precursor to a higher level of attention, such as moving from monitoring the partner's actions towards common attention experience between partners. Furthermore, it was noticed that richer, second-person relations may come in degrees. Siposova and Carpenter (2019) point out that generating shifts in the scale of jointness requires *salient stimulus* (here, the moving head of the clown), a relevant *shared goal* between the partners (to solve the problem of whether their machines work similarly) and a common ground (previous experiences and knowledge of the partner as a co-student). Moreover, the short *time interval* (immediate nature of sharing the appearance of the object and communication of the objects) and the limited *perceptual space* of the structured CPS environment can 'push' forward the situation in the scale of jointness. In our study, the two dyads were also rather distinctive in their orientations: The second dyad was seemingly more motivated to share ideas and thoughts than the first, which may reflect *individual differences* in this regard, helping to slide right on the scale of jointness (Siposova & Carpenter, 2019). Individual commitments can trigger social obligations beyond one's own, and through communication, create joint goals and joint commitments to it (Siposova & Carpenter, 2019; Siposova, Tomasello, & Carpenter, 2018). This issue is closely connected to the quality of the *behaviours*—for example, how explicit or detailed the communication is. In our cases, there were also evident differences in this regard. Moreover, to fully understand social connotations in CPS, gaze data did not provide sufficient contextual details of the actual interactions alone, but, in the lower attention levels such as monitoring attention conditions, for example, the gaze data view was beneficial. This is because it not only reinforced the individual orientation, visible in the log stream data view, but also gave evidence of the 'social' behind these data. Due to the unique properties of remote social interaction, the 'richness' of attention experience or the indicative behaviours may not be fully identical if compared with a face-to-face situation as defined in Siposova & Carpenter (2019). Also, the attention experience and the indicative behaviours of JAB in remote dyadic interaction may not be fully equal with younger population of students, the authentic target audience of the designed CPS tasks (11-15-year-olds). As limitations of this study also include the low number of cases, as a next step, the focus is on *larger population* of middle school students' CPS processes (40 dyads), and, particularly, on *longer behavioural sequences* of CPS interaction in dyads. That is, to search for more evidence by analysing what behaviours preceded and what followed after identified social attention levels in longer tasks with multiple student dyads in school context (see Siposova & Carpenter, 2019).

Acknowledgements

This research is funded by the Academy of Finland (Grant no. 316836).

References

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, *12*(3), 307–359.

Care, E., Scoular, C., & Griffin, P. (2016). Assessment of collaborative problem solving in education environments. *Applied Measurement in Education*, *29*(4), 250–264.

Care, E., Griffin, P., & Wilson, M. (2018) (Eds.). *Assessment and teaching of 21st century skills: Research and applications*. Dordrecht, Netherlands: Springer.

Care, E., Griffin, P., Scoular, C., Awwal, N., & Zoanetti, N. (2015). Collaborative problem solving tasks. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st century skills: Methods and approach* (pp. 85–104). Dordrecht, Netherlands: Springer.

Carpenter, M., & Liebal, K. (2012). Joint attention, communication and knowing together in infancy. In: A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 159–182). Cambridge, MA: MIT Press.

Davis, P., Horn, M., Block, F., Phillips, B., Evans, M. E., Diamond, J., & Shen, C. (2015). 'Whoa! We're going deep in the trees!' Patterns of collaboration around an interactive information visualization exhibit. *International Journal of Computer-Supported Collaborative Learning*, 10, 53–76.

Dindar, K., Korkiakangas, T., Laitila, A., & Kärnä, E. (2017). An interactional 'live eye tracking' study in autism spectrum disorder: Combining qualitative and quantitative approaches in the study of gaze. *Qualitative Research in Psychology.*

Eilan, N. (2005). *Joint attention: Communication and other minds: Issues in philosophy and psychology*. Oxford, United Kingdom: Oxford University Press.

Falck-Ytter T., Bölte, S., & Gredebäck, G. (2013). Eye tracking in early autism research. *Journal of Neurodevelopmental Disorders*, *5*(28), 1–13.

Funke, J., Fischer, A., & Holt, D. V. (2018). Competencies for complexity: Problem solving in the twenty-first century. In E. Care, P. Griffin, & M. Wilson (Eds.), *Assessment and teaching of 21st century skills: Research and applications* (pp. 41–53). Dordrecht, Netherlands: Springer.

Graesser, A., Fiore, S. M., Greiff, S., Andrews-Todd, J., Foltz, P. W., & Hesse, F. W. (2018). Advancing the science of collaborative problem solving. *Psychological Science in the Public Interest*, *19*(2), 59–92.

Hesse, F. M., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). Framework for teachable collaborative problem solving skills. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st century skills. Methods and approach* (pp. 37–56). Dordrecht, Netherlands: Springer.

Korkiakangas, T. (2018). *Communication, gaze and autism: A multimodal interaction perspective*. London, United Kingdom: Routledge.

Miles. L. K., Lumsden, J., Flannigan, N., Allsop, J. S., & Marie, D. (2017). Coordination matters: Interpersonal synchrony influences collaborative problem-solving. *Psychology*, *8*(11), 1857–1878.

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science*, *16*(5), 269–274).

O'Madagain, C., & Tomasello, M. (2019). Joint attention to mental content and the social origin of reasoning. *Synthese*, doi:10.1007/s11229-019-02327-1

Pöysä-Tarhonen, J., Care, E., Awwal, N., & Häkkinen, P. (2017). Case-based portraits of contrasting micro-interaction processes during online assessment of collaborative problem solving. In W. Chen et al. (Eds.), *Proceedings of the 25th International Conference on Computers in Education,* (pp.146-130). Christchurch: Asia-Pacific Society for Computers in Education.

Pöysä-Tarhonen, J., Care, E., Awwal, N., & Häkkinen, P. (2018). Pair interactions in online assessments of collaborative problem solving: Case-based portraits. *Research and Practice in Technology Enhanced Learning*, *13*(12), 1-29.

Scoular, C., & Care, E. (2020). Monitoring patterns of social and cognitive student behaviors in online collaborative problem solving assessments. *Computers in Human Behavior, 104.*

Scoular, C., Care, E., & Hesse, F. (2017). Designs for operationalizing collaborative problem solving for automated assessment. *Journal of Educational Measurement*, *54*(1), 12–35.

Seemann, A. (2012). Introduction. In A. Seemann (Ed.) *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp.1-19). Cambridge, Massachusetts: MIT Press.

Siposova, B., & Carpenter, M. (2019). A new look at joint attention and common knowledge. *Cognition*, *89*, 260–274.

Siposova, B., Tomasello, M., & Carpenter, M. (2018). *Cognition, 179*, 192–201.

Tomasello, M. (1995). Joint attention as a social cognition. In: C. Moore and P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Lawrence Erlbaum.

Zahavi, D. (2015). Self and other: From pure ego to co-constituted we. *Continental Philosophy Review*, *48*(2), 143–160.

Zemel, A., & Koschmann, T. (2013). Recalibrating reference within a dual-space interaction environment. *International Journal of Computer-Supported Collaborative Learning*, *8*(1), 65–87.

Zwiecki, Z., Ruis, A. R., Farrell, C., & Williamson Shaffer, D. (2020). Assessing individual contributions to collaborative problem solving: A network analysis. *Computers in Human Behavior, 104*.