



Student visual attention during group instruction phases in collaborative geometry problem solving

Eeva S. H. Haataja¹ · Anniina Koskinen-Salmia¹ · Visajaani Salonen¹ · Miika Toivanen² · Markku S. Hannula¹

Accepted: 19 May 2024
© The Author(s) 2024

Abstract

Teacher gestures support mathematics learning and promote student collaboration. Aligned with speech, gestures can help students to notice the important visual information of geometry tasks. However, students' visual attention to the teacher's gestural cues during collaborative problem solving remains a largely unexplored field in mathematics education research. This mixed-method case study investigated relations between students' visual attention, teacher gestures, and students' collaborative problem-solving process on a geometry task. The data were collected with video cameras and mobile gaze trackers on four students simultaneously in two Finnish 9th-grade mathematics lessons with the same teacher. The findings show that the students attended to their own papers most of the time during the teacher's gestures, but differences in student attention between the gesture types emerged. The qualitative analysis showed that the teacher's tracing, pointing, and representational gestures helped in directing student attention to targets relevant to the situational learning process. We conclude that teacher gestures can both convey mathematical contents and direct student attention, which intermediates mathematical thinking in problem solving.

Keywords Gesture · Attention · Mathematical thinking · Eye movements · Teacher-student interaction

1 Introduction

Gesturing during instruction can be highly effective and can lead to greater retention of instructional material (Congdon et al., 2017), if the students notice the gestures. Teachers' simultaneous use of speech and gestures can enhance the students' mathematical learning, as well as transferring (Cook et al., 2017), generalizing (Congdon et al., 2017), and memorizing what has been learned (Cook et al., 2013). Dynamic gestures simulate expert mathematicians' embodied cognitive processes and help students access insights and proving

✉ Eeva S. H. Haataja
eva.haataja@helsinki.fi

¹ Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

² SeeTrue Technologies, Joensuu, Finland

practices in geometry (Nathan et al., 2021). Gestures are also significant in ensuring a common understanding between the teacher and the student (Koskinen et al., 2015) and engaging the students in the task (Svensson & Johansen, 2019).

Gestures are a tool for the teacher to facilitate shared mathematical cognition: with gestures, the teachers can guide students' attention to new ideas and create a common understanding of the learning content (Nathan & Walkington, 2017). However, the situational mechanisms behind this effect require more investigation (e.g., Schindler & Lilienthal, 2019). There are alternative methods to study students' situational visual attention in the classroom. The first would be to observe student behavior and the direction of their gaze. However, even with good video footage, it would be possible to know only the general direction of the student gaze. The second method would be for the student to report their attention thinking aloud or in a video-stimulated recall interview. However, as the target of attention can shift multiple times over the course of 1 min, this would be disturbing for the learning and difficult for the student. The results would be highly unreliable, especially regarding the timing of attentional shifts. To cover these issues, this study uses the method of eye tracking (cf. Strohmaier et al., 2020).

Eye tracking is a method used to explore a participant's cognitive processes through visual attention (Hannula et al., 2022; Just & Carpenter, 1980), whereas gestures can be seen as a window to a person's mathematical thinking beyond the limits of verbal understanding (Edwards, 2009). The direction of the teacher's visual attention both gathers and communicates information to the students in the instructional interaction (Böckler et al., 2016). Therefore, both teacher attention and gestures are modes of communication that are intuitive and even unconscious but communicate the teacher's pedagogical intentions. A recent review has indicated that eye tracking is particularly suitable for studying the subconscious processes of mathematical thinking (Strohmaier et al., 2020).

Combining eye tracking with video data on the teacher's gestures enabled us to examine precisely the role of a teacher's gestural cues as facilitators of students' attention in geometry problem solving. As geometry combines visual skills with abstract thinking (Levav-Waynberg & Leikin, 2009), a complex geometry problem relates to students' increased use of multimodality and representational gestures (Chen & Herbst, 2013). Specifically, as embodied cognition is central in mathematical thinking, teachers can support students with simultaneous gestures and speech that both simulate the dynamics of the geometry objects (Nathan et al., 2021).

This study investigates group-level teacher-student interaction in authentic mathematics classrooms with mobile eye tracking. Schindler and Lilienthal (2019) suggest that when learning geometry, students' actual presentations of thinking (e.g., speech or drawings) may appear delayed compared to the cognitive processes that only happen mentally but may be visible to eye-tracking technology. In this paper, we argue that these processes can be seen in the interplay of the students' visual attention and teacher gestures.

1.1 Gestures in geometry instruction

We define gestures as movements of hands in space or on objects (McNeill, 1992). Gestures bring the aspect of embodied cognition to learning of abstract mathematical contents (Arzarello et al., 2015) and the connections between them (Alibali et al., 2014). Gestures are a particularly effective tool for communicating visual, imagistic, and dynamic features, such as geometric representations that would be laborious to describe in speech (Sabena, 2008). Teachers use gestures to capture and maintain the students' attention and ground

the instruction in the physical environment (Alibali & Nathan, 2012). The use of precise dynamic gestures enhances the students' abilities of mathematical reasoning (Nathan & Walkington, 2017), as they simulate motor actions that implement the embodied cognition through which the mathematical concepts and physical environment relate to each other (Pier et al., 2019).

McNeill's (1992) categorization of different kinds of gestures is widely used in educational studies (Alibali & Nathan, 2012). According to McNeill (1992), beat gestures are simple, rhythmic, and non-semantic gestures that align with the prosody of speech. Pointing gestures, which are mathematics teachers' most common gesture type during instruction (Alibali & Nathan, 2012), serve to indicate objects and their locations (McNeill, 1992). Pointing gestures are a natural way of capturing the attention of the listener (Kendon, 2000). For mathematics education, Alibali and Nathan (2012) combined McNeill's (1992) original categories of iconic and metaphoric gestures to a new category of representational gestures. Representational gestures depict the semantic contents of the instructional interaction either via a metaphor or directly via the shape or motion trajectory of the hand(s). Conveying the dynamic aspects of mathematical structures, representational gestures are associated with mathematical expertise (Nathan et al., 2021). The representational gestures tracing a concrete shape are called tracing gestures (Alibali & Nathan, 2012). Additionally, Salminen-Saari et al. (2021) separate attention to specific targets (e.g., a solution drawing) from gestures on imaginary shapes (e.g., drawing in air) in the problem-solving process.

The multimodal use of gestures and speech can enhance students' understanding of mathematical knowledge (Congdon et al., 2017), as the development of abstract mathematical thinking takes place in the shared space of verbal and sensorial channels (Radford, 2009). Even young students use pointing gestures to direct each other's attention and tracing gestures to justify their thinking in collaborative problem solving (Wathne & Carlsen, 2022). In the context of learning geometry, gestures can complement speech to overcome the limitations of static, drawn geometric representations (Chen & Herbst, 2013). In a virtual learning environment with dynamic geometry representations, the tracing gestures complemented with simultaneous speech can simulate the operations on the mathematical software and, therefore, serve as a communicational bridge between the participants (Ng, 2016).

Wakefield et al. (2018) suggest that gestures' beneficial effect on learning comes from their ability to synchronize with speech and thus affect what learners glean from the speech. Additionally, gestural expressions of mathematical contents can also serve as an independent cognitive source in geometry modeling or as a tool to overcome students' limited abilities to express themselves verbally (Nathan et al., 2021; Radford, 2009). When working on geometrical representations, the cognitive process seems to develop from mental processing (visible in visual attention), through gestural presenting to written or drawn form of the solution (Schindler & Lilienthal, 2019).

1.2 Visual attention in collaborative geometry problem solving

Mathematical problem solving is a process of carrying out a task unfamiliar and nonroutine to the solver (Schoenfeld, 1985). This process is cyclic and can be operationalized through division into phases. A recent eye-tracking study (Salminen-Saari et al., 2021) synthesized theories by Pólya (1945), Schoenfeld (1985), Artzt and Armour-Thomas (1992), and Carlson and Bloom (2005) from the perspective of collaboration in geometry problem solving that requires joint visual attention to shared targets. The initial phases of problem solving

(*Understanding the problem, Planning and exploring, and Implementing*) include mainly attention to one's own task sheet rather than collaboration that would also be implemented in students' visual attention (Salminen-Saari et al., 2021). Salminen-Saari et al. (2021) argue that during collaboration, the students tend to pay attention to other participants and their geometric representations mainly during *Verifying* and *Watching and listening* phases. *Verifying* refers to the evaluation of one's solutions and watching and listening means attending to other participants' presentations. In these phases, the students display and confirm their understanding of the problem task in interaction with each other and this underlines the importance of sharing the attention.

Teachers' role in problem solving is to provide opportunities for the students to interact with the learning environment and each other (Svensson & Johansen, 2019). In addition to the participants' individual problem-solving skills, the success of collaborative mathematics learning requires social skills, such as the ability to maintain meaningful attention (Barron, 2003). Research with mobile eye tracking has specified this idea: there is a strong connection between students' high-quality interaction and the amount of visual attention to shared targets during collaborative problem solving (Schneider et al., 2018). We measure visual attention through the direction of the central area of sharp vision in the middle of the visual field (Holmqvist & Andersson, 2017). Persons' visual attention lies on those areas of a target that they see with sharp vision (Buswell, 1935; Land, 2006). Teachers' momentary scaffolding can help the students in directing their attention toward meaningful targets (Haataja et al., 2019a). The teacher can reflect the student's momentary attention and adapt the instructions accordingly by complementing verbal communication with embodied gestures (Shvarts, 2018). The students interpret the meanings of the learning contents and instructional intentions by simultaneously following teachers' verbal instruction, gestures, and gaze cues (Shvarts, 2018; Tomasello, 1995).

However, studies on students' momentary visual attention in authentic learning environments with mobile eye tracking are scarce, especially at the social level of more than two participants. Heyd-Metzuyanim et al. (2023) analyzed processes behind failed problem-solving combining discourse analysis with multiple mobile eye tracking in a mathematics classroom. Salminen-Saari et al. (2021) analyzed students' shared visual attention in relation to the phases of a mathematical problem-solving process. Haataja et al. (2021) and Haataja et al. (2019b) have examined teacher's eye contact initiatives and student responses to them in mathematics classrooms. The present study explores the students' visual attention in an authentic classroom environment from the perspective of attention to teacher gestures. Our multi-person gaze data enables analyzing precisely where the students look at when the teacher wishes to direct their visual attention towards learning materials.

2 Research questions

Teacher gestures and meaningful situational attention can be significant for mathematical learning, and these processes can be captured with mobile eye tracking in real learning contexts. This study with one teacher and two student groups was an investigation of student gaze during teacher gestures in authentic contexts of geometry problem solving. Our research questions were:

1. What do the students pay visual attention to during the teacher's gestures in the context of mathematical problem solving?

2. How are the teacher's gestural cues associated with the students' geometry problem solving processes?

3 Methods

For this study, we used a mixed-method approach to eye tracking. For educational interpretations on the eye-tracking recordings, the mixed-method approach is fruitful, as quantitative methods can provide information on general patterns of visual attention, while qualitative analyses can deepen this understanding with detailed reflections (Beach & McConnel, 2019; Haataja, 2021).

3.1 Participants

In this study, we used data from two Finnish ninth grade mathematics lessons with the same mathematics teacher (students 15–16 years of age). Our participant teacher volunteered in this research twice (2017 and 2018) with different student groups. The participation of the students was voluntary and confirmed with written consent forms. At the time of the first data collection, she had 14 years of teaching experience. The first class included 19 students and the second 26 students. Among students in both classes, we selected four volunteers (target students) to wear eye-tracking glasses. In the first lesson, the target student group included four girls (student 1–student 4), and in the second lesson, three girls and one boy (student 5–student 8).

3.2 Setting and eye-tracking device

With data triangulation, we aimed to reach an in-depth understanding (Denzin, 2012) of teacher-student momentary multimodal interaction. We combined four sources of data for a comprehensive and detailed picture of the students' problem-solving processes: *gaze recordings* of students' visual attention, *video recordings* of classroom interaction, *Smart-Pen recordings* of the drawing process of the students, and *audio recordings* of students' verbal interaction. Previous studies on data from the first lesson indicate that the teacher-student eye contact interaction was relative to the teacher's pedagogical intentions (Haataja et al., 2019b) and her behaviors of friendliness and control (Haataja et al., 2021). Also, the distribution of the teacher's visual attention was found to vary in relation to the level and phase of instruction (Määttä et al., 2021). A study in the second lesson indicated that eye tracking can provide detailed information on the challenges of sharing and developing solution ideas in collaboration (Hannula & Toivanen, 2019).

The teacher and the target students wore eye-tracking glasses, which recorded their gaze throughout one mathematics lesson. The device was calibrated for each participant. The self-made eye-tracking devices consisted of two eye cameras, a scene camera, and simple electronics attached to plastic goggles (cf. Lukander et al., 2013; Toivanen et al., 2017, Hannula et al., 2022). The device provides an accuracy of approximately 1.5 degrees of the visual angle, which was adequate for valid interpretations in this study, as the areas of interest (AOIs) were clearly distinct from each other (cf. Strohmaier et al., 2020). The device has been used in several studies, and the data have been valid for fine-grained analyses (e.g. Haataja, 2021; Salminen-Saari et al., 2021). Each

of the five eye-tracking devices was connected to a PC laptop that was placed in a backpack. On the laptops, software recorded and processed the streams, computing the user's estimated gaze point in the scene camera in the glasses.

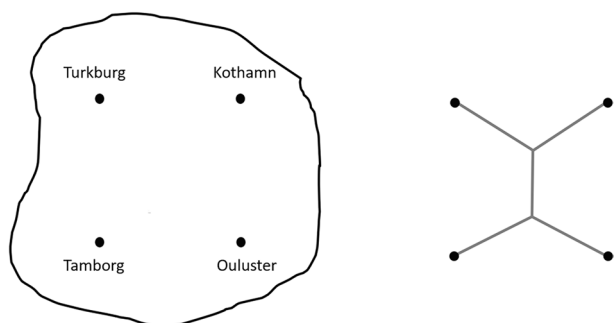
Additionally, stationary video cameras and personal microphones recorded the classroom activities and interaction. Two cameras were pointed at the target group. In the first lesson, students worked on the problem with SmartPens and paper. In the second lesson, they primarily used computers with Geogebra software, and SmartPens and paper served only as support. At the beginning of both lessons, we ensured the synchronizing of the data with a clapperboard. Attaching timestamps to the recorded video frames and synchronizing the computers' clocks prior to the recording ensured that the processed gaze videos were synchronized with each other and with the video and audio recordings.

3.3 Research setting

During data collection sessions (each of which was 75 min), both classes solved the same geometry task. The task was a Euclidian four-point Steiner Tree problem, formulated as a task to find the shortest connection between four imaginary cities located in the vertices of a square (Fig. 1). There were no significant disturbances in either lesson. Solving the task required both problem-solving skills (drafting various solutions, evaluating and comparing their affordances, and combining information to improve the solutions) and mathematical content knowledge on geometry (e.g., diagonals of a square, right-angled triangle, and Pythagoras' theorem).

Both lessons started with short whole-class instructions. Next, the students started solving the task individually and, after a short while, continued collaboratively in groups of three to five students (approx. 20 min). The teacher roamed the classroom advising and encouraging the students and asking questions but refraining from giving hints on the optimal solution. This report focuses on group instruction events with the target group ($N=17$) during the collaborative phases of the two lessons. We chose to focus on these parts of the data due to the interactional fruitfulness of the group-level teacher-student interaction and because teacher gestures during whole-class mathematics instruction are a relatively well-researched topic (e.g., Alibali et al., 2014). During the group instruction, the teacher advised the target students on the mathematical contents of the problem or affective aspects of the collaboration (cf. Haataja et al., 2019a). After the collaborative phase, the groups presented their solutions to the class, and the optimal solution was selected among them.

Fig. 1 The geometry problem task (left) and the optimal solution (right)



3.4 Analyses

In this study, we employed an inductive approach to the analysis of the video data (Derry et al., 2010). To gain an understanding of the nature of events in the classroom, we first observed the stationary video and the eye-tracking video data as a whole. We coded all the data with ELAN software (2019), which allowed us to annotate multiple features of the ongoing interaction (i.e., speech, gesture, and gaze) simultaneously.

3.4.1 The coding procedure

The coding of the data included four phases. First, we segmented the continuous stream of classroom activity based on the level of instruction to identify the group-instruction phases. Second, we transcribed the student and teacher speech verbatim.

Third, we identified all teacher gestures with the sound off. The segments of teacher gestures were coded from the moment when the teacher's hand(s) started moving until the moment when they were still again (cf. McNeill, 1992; Pier et al., 2019). With the sound on, all the gesture segments were coded as a gesture of beat (motorically simple, rhythmic gestures), pointing (indicate objects or locations), representational (iconic and metaphoric gesture in the air), or tracing (a moving iconic gesture on the student paper) (Alibali & Nathan, 2012; McNeill, 1992). The categorization of each gesture segment was directed by observations on the form of the gesture (i.e., hand shape and trajectory) and the words that accompanied the gesture (function). The second author coded all the gestures ($N=189$) and the first author coded 25% of the gestures the teacher produced during group instruction. We coded the gestures independently, and the inter-rater reliability was excellent (Cohen's Kappa=0.85). We also coded the targets for all teacher gestures, i.e., what the teacher refers to with the gesture. Gestures were either targeted on a student's paper or displayed or took place in the air (representational and beat gestures). In the findings, we have underlined the verbal utterance to which the stroke phase of each gesture was assigned (cf. McNeill, 1992).

In the fourth phase, we coded the gaze of each target student to find out where they were looking during the teacher's gestures. For the student gaze, we defined a "dwell" as a coding unit. Dwells refer to one glance at a researcher-defined area of interest (AOI) from the beginning until the end of the glance. A dwell can consist of several fixations on the AOI and saccades between them (Holmqvist & Andersson, 2017). The number of the dwells at certain AOIs resonate its significance for the activity at hand (Glöckner & Herbold, 2011).

We annotated all the dwells during teacher gestures. In this report, we examined those student dwells that were targeted towards the AOIs that were relevant for the instructive interaction. AOIs that were not relevant from the viewpoint of collaborative problem solving (e.g., walls or cameras) were excluded from the analysis. The AOIs included in the analyses were (1) students' own displays and papers that included the solution drafts and calculations, (2) other student displays and papers, (3) students' belongings (phones, pencils, rulers, etc.) on the desk, (4) student faces, (5) teacher face, and (6) teacher gestures. We coded the AOI as a teacher gesture when the student directed their attention at the teacher's hand(s) or followed the spot pointed at by the teacher's finger. The display and paper AOIs refer to other parts of the displays and papers.

3.4.2 The analysis procedure

We first analyzed quantitatively the students' AOIs during teacher gestures. For the overview on the students' gaze behavior during teacher gestures, we used crosstabulation and the Pearson Chi-square test and adjusted residuals for cellwise analyses with Bonferroni correction (cf. Garcia-Perez & Nunez-Anton, 2003) to compare the distribution of the count of gaze dwells between gesture categories.

After this, we focused on qualitative analyses of students' visual attention during teacher gestures. We analyzed the students' gaze data and the videos from the perspective of each student's problem-solving process in relation to the teacher's visual and gestural cues during the scaffolding interventions. Furthermore, we explored in detail two segments in which the teacher intervention made a significant contribution to the students' problem-solving process: the groups were stuck before the intervention and were able to continue towards the optimal solution after the intervention. The selection was based on a comprehensive qualitative analysis of all data sources, especially the combination of the student interaction auditive from the video data and their process of drawing the solutions recorded by the SmartPens.

To illustrate the momentary formation of attention during teacher gestures, we created heatmaps on student gaze. After identifying a period of teacher gesture, we selected the moment on the synchronized gaze recordings that showed the visual attention to the gestural target most clearly and took a screen capture of that moment for the heatmap. At the top of a screenshot on the selected frame, we created a heatmap on all the registered gaze locations within the period with MATLAB 2019. Figure 2 exemplifies this method: the heatmap shows the density of the student's visual attention in the environment during a teacher gesture on the peer student's notebook.

4 Findings

We start our description of the findings with an overview of the teacher's gestures over the two lessons. Table 1 presents the distribution and the frequencies of all teacher gestures during the group-instruction phases of the two lessons.

We found a clear increase in the number of teacher's gestures from the first lesson with only pen-and-paper notebooks ($n=41$) to the second lesson with laptops and notebooks ($n=148$). The distribution of the gesture types was similar in both lessons. Half of the

Fig. 2 Example of a heatmap on student visual attention to teacher gestures

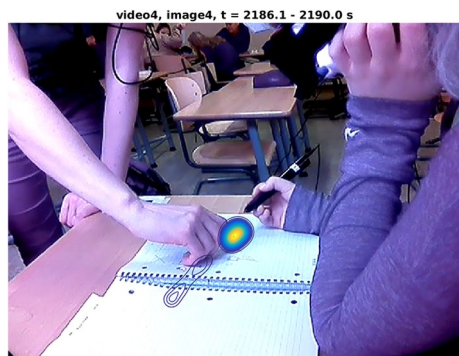


Table 1 The distribution of teacher gesture types in two lessons

Teacher gesture	Lesson 1		Lesson 2		Total	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Pointing	21	51%	86	58%	107	57%
Beat	6	15%	24	16%	30	16%
Representational	9	22%	23	16%	32	17%
Tracing	5	12%	15	10%	20	10%
Total	41	100%	148	100%	189	100%

gestures were pointing gestures, followed by representational gestures, beat gestures, and tracing gestures.

The following two sections focus on the role of teacher gestures in the students' problem-solving process. First, we present our findings on the total direction of student visual attention during teacher gestures. Second, we examine how the teacher's gestures and student attention to them related to the students' geometry problem-solving processes.

4.1 Student visual attention during teacher gestures

The Chi-square test with Bonferroni-adjusted pairwise comparisons yielded statistically significant differences between the AOIs of student visual attention when the teacher gestured during group instruction ($\chi^2(15) = 114.1, p < 0.001$). We combined the findings from the Chi-square test with observations on the video recording to provide a picture of the situational role of the teacher gestures in the teacher-student interaction. Figure 3 shows the overall distribution of students' visual attention to the teacher gestures.

During the teacher's *beat gestures* in the group-instruction phases of the lessons ($n = 30$), the students looked at the teacher's gestures only twice (1%). They focused on their own solution papers and displays ($n = 94, 64%$). The proportion of student attention

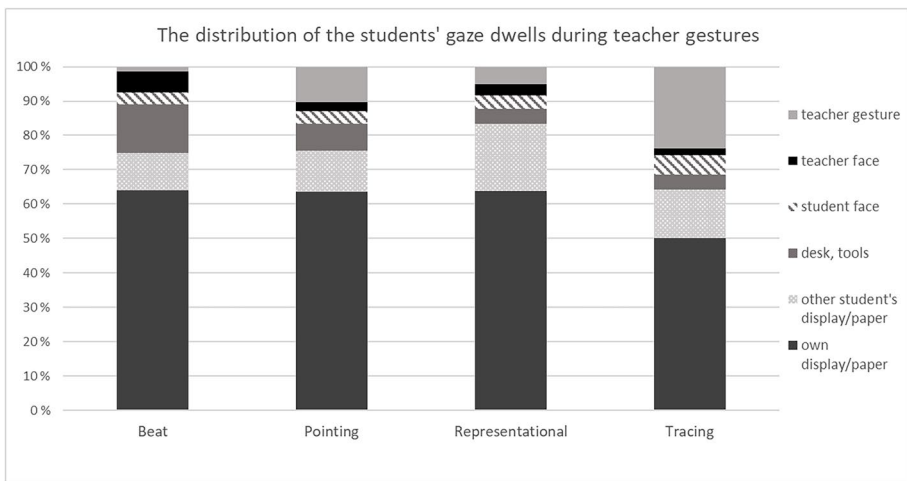


Fig. 3 The distribution of student gaze dwells (% of the total number of dwells) during each category of teacher gestures in two lessons with the same teacher

to teacher gestures was significantly lower than the attention to other targets ($\chi^2(1)=11,29$, $p<0.001$). In fact, the student attention to student desks and tools was significantly higher ($\chi^2(1)=15,92$, $p<0.001$). Therefore, our interpretation was that teacher's beat gestures occurred in moments when the engagement to collaboration was low, and the students were focused on individual thinking or off-task activities.

During the teacher's *pointing gestures* ($n=107$), the students did not pay much attention to the teacher gestures as such ($n=177$, 10%). However, a statistically significant difference was found in the amount of visual attention to students' own papers and displays ($n=1081$, 64%) when compared to other targets ($\chi^2(1)=10,37$, $p=0.001$). The targets at which the teacher pointed were located on the students' papers or displays. However, the students looked at parts of the papers and displays other than those which the pointing gestures referred to. The students' attention was on the task-related targets in these moments. The pointing gestures were often aligned with personal instructions to one of the students, which may have influenced the other students not to be distracted from their work.

Instead, the moments with the teacher's *representational gestures* ($n=32$) often included student attention to each other. There was no statistically significant difference between any gaze categories during representational gestures. The descriptive statistics indicate that 19% ($n=35$) of the visual attention was directed to other students' papers and displays, whereas the gazes at teacher gestures ($n=9$, 5%) were surprisingly few. Once again, the students' own displays and papers ($n=115$, 64%) were the most gazed AOI. The low percentage of gazes at desks ($n=8$, 4%) indicated high attention to task-relevant targets.

The *tracing gestures* ($n=20$) captured the students' attention most effectively, as 24% ($n=91$) of the dwells were directed at teacher gestures during this gesture category ($\chi^2(1)=66,26$, $p<0.001$). When compared to other AOIs within the gesture category, the occurrence of gesture-targeted dwells during the tracing gestures was significantly higher than of the gazes at a student's own ($n=192$, 50%) or a peer's displays and papers ($n=54$, 14%), desks ($n=17$, 4%), or the teacher's face ($n=8$, 2%). In this gesture category, the student's attention to their own papers and displays was significantly lower than it was to other targets ($\chi^2(1)=24,50$, $p<0.001$). The tracing gestures were well aligned with the teacher's speech and thus were easy to follow. They were also clearly related to the geometry contents of the problem task (e.g., tracing the lines in the solution drafts). The low percentage of gazes at the students' belongings on the desks during tracing and representational gestures indicates high task engagement at these moments.

4.2 Visual attention to teacher gestures

The aim of our second research question was to examine how the teacher's gestural cues are associated with the students' geometry problem-solving processes.

4.2.1 The first lesson: students' attention to tracing gestures on imaginary lines

In the first lesson, the students had individually drawn different solution drafts on the papers, and the teacher came to the group. Student 4 asked whether they should find a solution with "even fewer lines," and the teacher corrected the misunderstanding on the task by replying: "Yes, with shorter cables." After this, the students started to draft new solutions. After a few minutes, the students complained to the teacher that they were too stupid to solve the problem, and the teacher encouraged them to continue. The teacher tried verbally

to direct the group's attention to the existing solution options in the notebooks. However, the students were staring at their own papers and hands while listening. The teacher pointed at the X solution on student 2's paper. Student 1 and student 3 mainly looked at this gesture, but student 2 looked mainly at her paper with a short glance on the gesture. Student 4 looked at the gesture and briefly glanced at the research equipment.

T you could consider the solutions you already have.

T if you could combine some of their advantages.

T for example, what are the advantages of *this cross*? (teacher points with her little finger to solution X on student 2's paper).

S1 it goes to all cities and connects in the middle.

T Yes. Think how you could improve it a bit more.

The teacher wanted the students to reflect on the affordances of the X solution, which she considered the most optimal one at the time, and especially to the middle point where the two lines cross.

Next, the teacher noticed that Student 1 already had the optimal solution on her paper. Pointing to that solution, the teacher said: "You have started to draft this kind of (a solution). How did this develop?" Before this pointing gesture, student 1's visual attention had been wandering on different solutions on her and student 4's papers and the teacher's face. However, student 1 was not convinced by the teacher's hint.

S1 It only makes it longer.

T Compared to what?

S1 That one (X).

T Okay. You could measure as well.

The following teacher gesture (Fig. 4) was crucial for student 1's problem-solving process. The teacher said: "Does it go like that?" and drew an imaginary line on the top of the solution that indicated a more optimal angle of the diagonal line. Student 1 and student 4 attended to this gesture, and student 2 watched and listened to the interaction (Fig. 5). Figure 4 shows the view to the group on the left-hand side, the screen capture of the teacher gesture with the teacher's visual attention presented on it on the right-hand side, and the student solution, which the teacher pointed at, at the bottom left-hand corner. Figure 5 presents the heatmap on the students' visual attention on the left-hand side and the seating and gaze directions of the students and the location of the teacher on the right-hand side.

This representational gesture helped student 1 in attending to the length of the middle line, as she pointed at it with her pencil and then pointed at the H solution saying:

S1 That one is the extreme version of it.

T That's right, this is like the extreme because *here are these* (tracing the straight lines of the H solution with the finger).

The gaze recording showed that both student 1 and student 4 looked at the teacher's tracing gestures on the solutions. After this, without further instructions to student 1 or student 4, the teacher turned toward student 2 and student 3. Student 1 started to draw a new version of the optimal solution, and student 4 copied the optimal solution to her notebook as well.

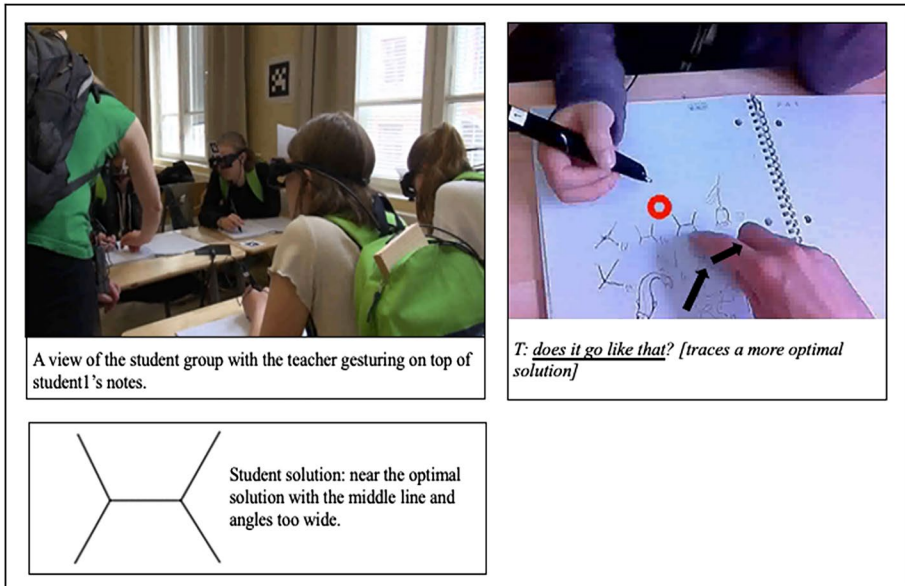


Fig. 4 The student group during the teacher gesture

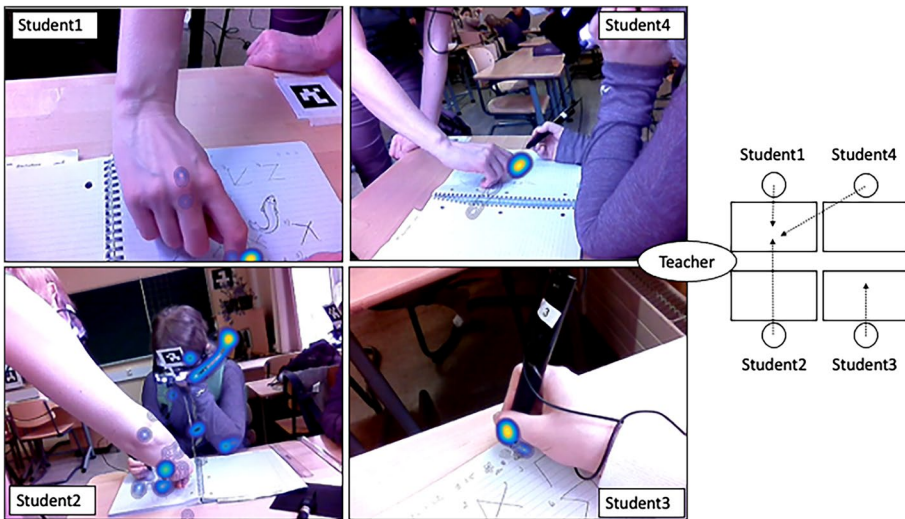


Fig. 5 Heatmaps on student's visual attention to teacher's pointing gesture on student 1's solution paper

Interesting processing occurred a while later when the teacher instructed the class to present the solutions on the blackboard. Student 1 wanted to select the solution to be presented and measured the lines of her solutions. The gaze recording showed that on the optimal solution, she did not measure the visible lines she had drawn but the imaginary lines indicated by the tracing gesture of the teacher in the previous scaffolding

interaction. After measuring the “invisible” lines, student 1 calculated the total length of the cables in the imaginary solution and confirmed that it was shorter than the X solution. This was the optimal one, and the group chose to present that to the rest of the class. In sum, in the first lesson, the teacher’s tracing gestures on imaginary lines helped the students to find the optimal solution to the task.

4.2.2 The second lesson: students’ attention associated with multimodal interaction with the teacher

In the second lesson, the students discussed a circular solution and had become quite convinced that it was a promising approach (for details, see Hannula & Toivanen, 2019). The teacher had initiated a discussion with student 6, gesturing at her screen and notebook, which contained two alternative solutions. Student 6 then combined the two solutions into the drawing in her notebook (Fig. 6). The other students mainly followed this dialog. However, due to the seating, other students were not able to see the images the teacher gestured at. By lifting the student’s notebook up and showing it to all group members, the teacher made a simple but crucial move to support the collaboration. The heatmaps in Fig. 7 show how all four students attended to this gesture simultaneously.

- T Now see here. Can I lift your paper a little bit?
 S6 yes.
 S7 so, doesn’t it –.
 T Here we have – here student 6 has drawn *the circle thing* (traces the circle with index finger).
 T *and the straight one* (traces the three straight lines).

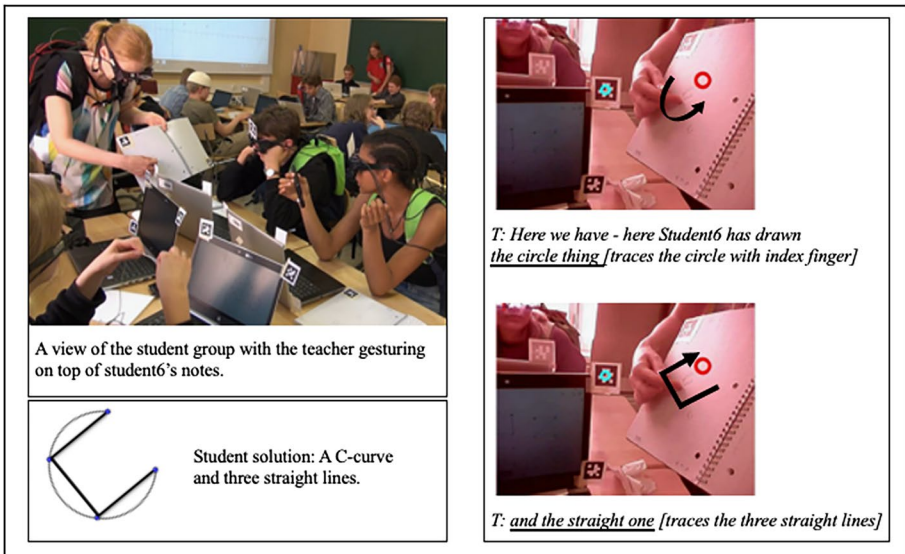


Fig. 6 The teacher gesture is illustrated on the screen capture from the gaze recording of student 8

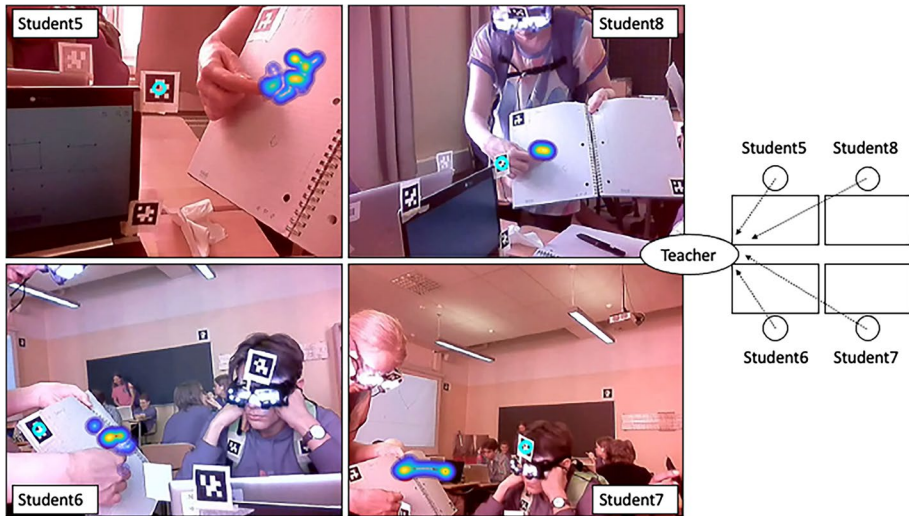


Fig. 7 Heatmaps on student's visual attention to teacher's tracing gestures on student 5's solution paper

It is noteworthy that the teacher placed the lifted student 6's notebook in front of student 5's face. Student 5 had been the least engaged participant in the group, and with this subtle move, the teacher was able to ensure that all the participants took notice of the interaction. During the teacher's tracing gestures, all the students were gazing at the figures on paper and ultimately agreed on curved lines not being optimal. This multimodal interaction helped students to verbalize their thinking and thus compare the features of the different solutions when the teacher left the group. With tracing gestures, the teacher helped the students in focusing on the geometrical features of the solutions (e.g., straight versus curved lines) rather than merely noticing the promising solutions among the drafts. In this kind of geometry problem solving, where the solutions are visual, attention to these features can be crucial for successful learning.

However, the idea of curved line reappeared in many of the students' later solutions, ultimately paving the way for the optimal solution (for details, see Hannula & Toivanen, 2019). It is important to know that student 6 continued to work on the computer and on the other side of the notebook, and only five short fixations on or near this crucial drawing (Fig. 8) were observed during the next 18 min until the teacher pointed at it in the following episode.

The student 6 had produced the H' solution with GeoGebra when the teacher initiated the following dialog with her.

T What did you think about the curves?

S6 What?

T What did you *think about the curves here?* (Pointing at the drawing on S6 notebook).

Fig. 8 The evolving solutions, which we will call X', H', Y', and Y



- S6 That this way is longer than this.
T Hmm, still you have ended up with the curves.
S6 Yes.
T You could think if you could use that *earlier* (emphasizing the word earlier with a beat gesture) observation on the curves compared to straight lines. Could there be *something nicer*? (Pointing at the H' solution on screen).
S6 Ahh, like doing something like this? (S2 draws the Y solution).

In this situation, student 6's attention had been focused on her attempts to construct the H' solution with GeoGebra. However, the teacher's pointing gesture on the paper drew her attention to the critical drawing with straight and curved lines in the same solution (Fig. 5), and we assume the related discussion that had evolved around it. The direction of her attention moved from the paper to the screen even a bit before the teacher actually pointed at the solution on the screen. The teacher's multimodal instruction helped student 6 to combine her earlier conclusion about curved vs. straight lines in the context of her most recent solution idea, so that she found the optimal solution. In sum, the second lesson illustrated how teacher's multimodal interaction with the students directed the visual attention that enabled comparison of two solutions.

5 Discussion

In collaboration, the students' visual attention depends on their social and emotional intentions (Heyd-Metzuyanim et al., 2023) and their teacher's behaviors of friendliness (Haataja et al., 2021). In our data, also the teacher's gestures directed the student's attention to shared targets that were relevant to the students' persistence in the problem-solving process.

5.1 Main findings

Teachers are known to use gestures to make breaks in shared understanding of mathematical reasoning to introduce new ideas that help students in proceeding with the task (Nathan & Walkington, 2017). This study brings insight into how the students direct their visual attention during the teacher's gestures in the context of geometry problem solving. As the teacher restricted the amount of verbal explication of the features that were crucial to finding the optimal solution, the students attended to her gestural cues. Additionally, the teacher seemed to trust the re-establishment of the shared embodied cognition if she perceived that some of the students had attended to her gestures. This underlines the fundamental multimodal nature of mathematical reasoning: as some students attended to the teacher's gestures and some to her verbal scaffolding, the whole group was able to overcome their struggles and continue working.

As expected, the students did not look at the teacher's *beat gestures* much. These gestures may be easy to perceive with peripheral vision (Holmqvist & Andersson, 2017), meaning that people are accustomed to the use of them in normal communication and do not direct their gaze at them to acknowledge their existence. Though not frequent in occurrence, *the tracing gestures* in student solutions were the gestures that directed the student's attention most effectively. During these gestures, the students directed their visual attention according to the teacher's gestural cues more often than during other

gestures. The students often followed the teacher's tracing gestures with their gaze and simultaneously listened to her verbal instruction. With tracing gestures, the teacher was able to provide effective micro-level cognitive scaffolding (cf. Haataja et al., 2019a) and to point out significant visual features of geometry drafts (cf. Levav-Waynberg & Leikin, 2009). Specifically, in the second lesson, the students had drawn curved and straight lines into the same solution draft but failed to notice the clear difference in the length between the versions. The teacher's slow tracing of lines helped the students to notice this mathematically significant aspect and to continue working towards the optimal solution.

The representational gestures were often quite large and thus visible with peripheral vision. These gestures captured students' direct visual attention quite rarely, and they often looked at each other's solution papers and displays, which indicates collaboration and engagement in these moments. However, the previous literature suggests that these dynamic gestures are the most effective for directing students' mathematical reasoning (Nathan & Walkington, 2017). The qualitative analysis indicated that in the first lesson, the teacher's small representational gesture on a student's notebook directed the problem-solving toward a fruitful direction. The imaginary line, drawn by the teacher's finger and noticed by two students, was a crucial feature of the evaluation phase of the process. In practice, this underlines the importance of a teacher's careful interaction that helps the students to direct their attention intentionally, rather than intuitively. Looking directly at the teacher may not always be easy for adolescent students, and the teachers could explicate verbally when they want the students to look at the teacher's hands or body, which they use as visual cues for instruction.

Our decision to separate tracing gestures from other types of representational gestures appeared to be fruitful. In our data, the tracing gestures seemed to be somewhat similar to the pointing gestures, but more dynamic, with regards to what attentional behavior is related to them. From our gaze recording, we were able to identify details of student attention, such as whether the attention was to teacher gestures per se or to other parts of student papers and displays. Less sophisticated methods of data collection, such as video recordings, would have not provided data to identify targets of attention in such detail. In our qualitative findings, we interpreted how the student's attention shifted across different solutions and how the teacher's timely gestures directed the attention to relevant information.

The tracing gestures often include more information on the dynamics of the geometry representation compared to *the pointing gestures* (e.g., Alibali, 2005). The pointing gestures can be crucial in a multifaceted learning context with several people, notebooks, and screens. In the second lesson, the teacher's single pointing gesture helped the student to combine her ongoing effort with an earlier discussion, which immediately led to inventing the optimal solution. Additionally, the teacher's pointing gesture helped the student to compare between information on her paper and display, even though her glance at the teacher's gesture was very brief. The gesture seemed to remind the student of the contents of the scaffolding interaction that had taken place almost 20 min before this moment. This highlights the importance of situational continuous analyses on multimodal research data.

With respect to the problem-solving process, Salminen-Saari et al. (2021) separated the phases of *Verification* and *Watching and listening* from each other for reliable interpretations on the role of visual attention in collaborative problem solving. We suggest that these phases are intertwined, as a student deliberately watching the teacher gesture and listening to her is actually evaluating the solution at the cognitive level. After seeing a gestural cue from the teacher, the student can continue the evaluation and develop a more sophisticated or precise solution.

When the teacher uses multimodal communication, the modalities complement each other (Alibali et al., 2014; Congdon et al., 2017). With gestures, the teacher may be able to convey deeper meanings of mathematical representations (e.g., Sabena, 2008). However, these gestures are short, delicate, and situational. Hence, it requires pedagogical expertise to convey them to all students in a complex learning situation. From a practical perspective, this underlines the importance of setting up the classroom in a way that supports relevant visual attention. For example, the laptop computers blocked the sight of some group members to others' solutions and some of the gestures. In one moment, the teacher lifted up one student's notebook to show relevant visual information to the group. This indicates that at least unconsciously, the teacher acknowledged the importance of visual attention for successful collaborative problem solving. Theoretically, when the instruction is multimodal, so is the attention, as some students are listening to the teacher, some are looking at her gestures, and some at the mathematical representations on the paper.

5.2 Limitations

From the methodological perspective, the use of mobile eye-tracking glasses was essential for the validity of the data. From the multiple and synchronized gaze recordings, we were able to perceive not only the AOI of the person but also the content of the gestures very accurately. Additionally, the black goggles of our eye-tracking devices prevented some of the peripheral vision of the participants. This validates our interpretations of what they focused on. However, the existence of the glasses may have some influence on students' interactions. Previous research suggests that the reactivity to eye-tracking is minimal among adults, but the equipment may bother children (Magnussen et al., 2017; Praetorius et al., 2017). When asked about this, students in our study did not report equipment to be a distraction to them (see also Hannula et al., 2022).

This study used a mixed-method case study approach, which affects the interpretations drawn from it. Research like this cannot reach the levels of generalizability of studies with larger sample sizes, but it benefits from the depth of the fine-grained analyses on a restricted amount of data and the authenticity of the data collected in real context (cf. Haataja, 2021; Jarodzka et al., 2017). The context of our study, geometry problem solving, framed our study. It highlighted the importance of data triangulation, as the verbal modality was significantly supported by gestures and visual attention as a source of information. However, our geometry task enabled solving it individually and/or with only trying different solutions without collaborative evaluation. Also, some of the student drawings were small, and the teacher's pointing gestures to them may have been difficult to see from the perspectives of others. Future studies could consider that a more dynamic task that requires student gestures to be solved, could provide fruitful new information.

5.3 Conclusion

Purposeful direction of attention is crucial for successful problem solving (Salminen-Saari et al., 2021). The eye-tracking method we used opened new insight to the use of teacher gestures in authentic classroom interaction. We interpret that the teacher gestures simultaneously convey mathematical contents and direct student attention, and therefore, this multimodal interaction intermediates mathematical thinking in problem solving. From the perspective of the problem-solving process, this instructional interaction combines the problem-solving phases of evaluation and watching and listening (cf. Salminen-Saari et al.,

2021). Subtle gestures helped students to align verbal descriptions with visual aspects of the solutions and to overcome misunderstandings about the geometry contents (cf. Svensson & Johansen, 2019). This happened by directing students' situational attention to the pros and cons of the solution drafts and helped them in moving from individual work to collaboration and from drafting to evaluating the solutions.

However, when a teacher helps students with gestures, students' timely visual attention seems essential for the efficacy of the instruction. The students failed to notice many teacher gestures even though the learning content was as visual as in this case. This underlines the importance of investigating the situational interaction processes in classrooms in detail but also the need for professional development. Through professional learning, teachers could develop awareness of what the most important features of the geometry problems are and how to convey them to the students in a way that does not funnel their learning process but improves students' learning of mathematical problem solving.

Acknowledgements We wish to express our thanks to our participant teacher and students for making this research possible. We also thank Dr. Enrique Garcia Moreno-Esteva for his contribution in the data collection and Dr. Anu Laine for her advice during the process of writing up this manuscript.

Funding Open Access funding provided by University of Helsinki (including Helsinki University Central Hospital). We thank the Academy of Finland (grant 297856) for funding this research. The research is part of the MathTrack project (2016–2020), which was funded by the Academy of Finland (grant 297856).

Declarations

Conflict of interest The authors declare no competing interests.

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

References

- Alibali, M. W. (2005). Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation*, 5(4), 307–331. https://doi.org/10.1207/s15427633sc0504_2
- Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences*, 21(2), 247–286. <https://doi.org/10.1080/10508406.2011.611446>
- Alibali, M. W., Nathan, M. J., Wolfram, M. S., Church, R. B., Jacobs, S. A., Martinez, J., & Knuth, E. J. (2014). How teachers link ideas in mathematics instruction using speech and gesture: A corpus analysis. *Cognition and Instruction*, 32(1), 65–100. <https://doi.org/10.1080/07370008.2013.858161>
- Artzt, A. F., & Armour-Thomas, E. (1992). Development of a cognitive metacognitive framework for protocol analysis of mathematical problem solving in small groups. *Cognition and Instruction*, 9(2), 137–175. https://doi.org/10.1207/s1532690xcio902_3
- Arzarello, F., Robutti, O., & Thomas, M. (2015). Growth point and gestures: Looking inside mathematical meanings. *Educational Studies in Mathematics*, 90(1), 19–37. <https://doi.org/10.1007/s10649-015-9611-5>
- Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, 12(3), 307–359. https://doi.org/10.1207/S15327809JLS1203_1

- Beach, P., & McConnel, J. (2019). Eye tracking methodology for studying teacher learning: A review of the research. *International Journal of Research & Method in Education*, 42(5), 485–501. <https://doi.org/10.1080/1743727X.2018.1496415>
- Böckler, A., Eskenazi, T., Sebanz, N., & Rueschemeyer, S. A. (2016). (How) observed eye-contact modulates gaze following. An fMRI study. *Cognitive Neuroscience: Social Attention in Mind and Brain*, 7(1–4), 55–66. <https://doi.org/10.1080/17588928.2015.1053442>
- Buswell, G. T. (1935). *How people look at pictures*. University of Chicago Press.
- Carlson, M. P., & Bloom, I. (2005). The cyclic nature of problem solving: An emergent multidimensional problem-solving framework. *Educational Studies in Mathematics*, 58(45), 45–75. <https://doi.org/10.1007/s10649-005-0808-x>
- Chen, C. L., & Herbst, P. (2013). The interplay among gestures, discourse, and diagrams in students' geometrical reasoning. *Educational Studies in Mathematics*, 83, 285–307. <https://doi.org/10.1007/s10649-012-9454-2>
- Congdon, E. L., Novack, M. A., Brooks, N., Hemani-Lopez, N., O'Keefe, L., & Goldin-Meadow, S. (2017). Better together: Simultaneous presentation of speech and gesture in math instruction supports generalization and retention. *Learning and Instruction*, 50, 65–74. <https://doi.org/10.1016/j.learninstruc.2017.03.005>
- Cook, S., Duffy, R., & Fenn, K. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development*, 84(6), 1863–1871. <https://doi.org/10.1111/cdev.12097>
- Cook, S. W., Friedman, H. S., Duggan, K. A., Cui, J., & Popescu, V. (2017). Hand gesture and mathematics learning: Lessons from an avatar. *Cognitive Science*, 41(2), 518–553. <https://doi.org/10.1111/cogs.12344>
- Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research*, 6(2), 80–88. <https://doi.org/10.1177/1558689812437186>
- Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., Hall, R., Koschmann, T., Lemke, J. L., Sherin, M. G., & Sherin, B. L. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *Journal of the Learning Sciences*, 19(1), 3–53. <https://doi.org/10.1080/10508400903452884>
- Edwards, L. (2009). Gestures and conceptual integration in mathematical talk. *Educational Studies in Mathematics*, 70, 127–141. <https://doi.org/10.1007/s10649-008-9124-6>
- ELAN (Version 5.3) [Computer software]. (2019). Nijmegen: max planck institute for psycholinguistics. <https://tla.mpi.nl/tools/tla-tools/elan/>
- Garcia-Perez, M. A., & Nunez-Anton, V. (2003). Cellwise residual analysis in two-way contingency tables. *Educational and Psychological Measurement*, 63(5), 825–839. <https://doi.org/10.1177/0013164403251280>
- Glöckner, A., & Herbold, A. K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*, 24(1), 71–98. <https://doi.org/10.1002/bdm.684>
- Haataja, E. S. H. (2021). *Understanding the role of gaze in momentary teacher-student scaffolding interaction during collaborative problem solving* [Doctoral dissertation, University of Helsinki]. <http://urn.fi/URN:ISBN:978-951-51-6947-1>
- Haataja, E. S. H., Moreno-Esteva, G., Salonen, E., Laine, V., Toivanen, A., M., & Hannula, M. S. (2019a). Teacher's visual attention when scaffolding collaborative mathematical problem solving. *Teaching and Teacher Education*, 86. <https://doi.org/10.1016/j.tate.2019.102877>
- Haataja, E. S. H., Toivanen, M., Laine, A., & Hannula, M. S. (2019b). Teacher-student eye contact during scaffolding collaborative mathematical problem-solving. *LUMAT: International Journal on Math Science and Technology Education*, 7(2), 9–26. <https://doi.org/10.31129/LUMAT.7.2.350>
- Haataja, E. S. H., Salonen, V., Laine, A., Toivanen, M., & Hannula, M. S. (2021). The relation between teacher-student eye contact and teachers' interpersonal behavior during group work: A multiple-person gaze-tracking case study in secondary mathematics education. *Educational Psychology Review*, 33(1), 51–67. <https://doi.org/10.1007/s10648-020-09538-w>
- Hannula, M. S., & Toivanen, M. (2019). Making and observing visual representations during problem solving: An eye tracking study. In M. Graven, H. Venkat, A. A. Essien, & P. Vale (Eds.), *PME43: Proceedings of the 43rd Conference of the International Group for the Psychology of Mathematics Education* (vol. 2, pp. 328–335). <http://www.igpme.org/publications/current-proceedings/>
- Hannula, M. S., Toivanen, M., & Garcia Moreno-Esteva, E. (2019). Eye movements during collaborative geometry problem solving lesson. In A. Shvarts (Ed.), *Proceedings of the PME and Yandex Russian Conference: Technology and Psychology for Mathematics Education* (pp. 138–145). Moscow, Russia: HSE Publishing House.

- Hannula, M. S., Haataja, E. S. H., Löfström, E., Moreno-Esteve, G., Salminen-Saari, E., & Laine, A. (2022). Advancing video research methodology to capture the processes of social interaction and multimodality. *ZDM – Mathematics Education*. <https://doi.org/10.1007/s11858-021-01323-5>
- Heyd-Metzuyanım, E., Haataja, E. S. H., Hannula, M. S., & Moreno-Esteve, G. (2023). What can eye-tracking, combined with discourse analysis, teach us about the ineffectiveness of a group of students solving a geometric problem? *Instructional Science*, 51(3), 363–396. <https://doi.org/10.1007/s11251-023-09617-9>
- Holmqvist, K., & Andersson, R. (2017). *Eye tracking: A comprehensive guide to methods, paradigms and measures*. Lund Eye-Tracking Research Institute.
- Jarodzka, H., Holmqvist, K., & Gruber, H. (2017). Eye tracking in Educational Science: Theoretical frameworks and research agendas. *Journal of Eye Movement Research*, 10(1), 1–18. <https://doi.org/10.16910/jemr.10.1.3>
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 17, 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>
- Kendon, A. (2000). Language and gestures: Unity or duality? In D. McNeill (Ed.), *Language and gestures* (pp. 47–63). Cambridge University Press.
- Koskinen, A., Seitamaa-Hakkarainen, P., & Hakkarainen, K. (2015). Interaction and embodiment in craft teaching. *Techne Series (Oslo)*, 22(1). <https://journals.hioa.no/index.php/techneA/article/view/1253>
- Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research*, 25(3), 296–324. <https://doi.org/10.1016/j.preteyeres.2006.01.002>
- Levav-Waynberg, A., & Leikin, R. (2009). Multiple solutions to a problem: A tool for assessment of mathematical thinking in geometry. In V. Durand-Guerrier, S. Soury-Lavergne, & F. Arzarello (Eds.), *Proceedings of the Sixth Conference of the European Society for Research in Mathematics Education - CERME-6* (pp. 776–785). <http://ife.ens-lyon.fr/publications/edition-electronique/cerme6/wg5-11-levav-leikin.pdf>
- Lukander, K., Jagadeesan, S., Chi, H., & Müller, K. (2013). OMG! A new robust, wearable and affordable open source mobile gaze tracker. In *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services (MobileHCI '13)* Association for Computing Machinery (pp. 408–411) New York. <https://doi.org/10.1145/2493190.2493214>
- Määttä, O., McIntyre, N., Palomäki, J., Hannula, M. S., Scheinin, P., & Ihantola, P. (2021). Students in sight: Using mobile eye-tracking to investigate mathematics teachers' gaze behaviour during task instruction-giving. *Frontline Learning Research*, 9(4), 92–115. <https://doi.org/10.14786/flr.v9i4.965>
- Magnussen, R., Zachariassen, M., Kharlamov, N., & Larsen, B. (2017). Mobile eye tracking methodology in informal E-learning in social groups in technology-enhanced science centres. *Electronic Journal of E-Learning*, 15(1), 46–58.
- McIntyre, N. A., & Foulsham, T. (2018). Scanpath analysis of expertise and culture in teacher gaze in real-world classrooms. *Instructional Science*, 46(3), 435–455. <https://doi.org/10.1007/s11251-017-9445-x>
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.
- Nathan, M. J., & Walkington, C. (2017). Grounded and embodied mathematical cognition: Promoting mathematical insight and proof using action and language. *Cognitive Research*, 2(9). <https://doi.org/10.1186/s41235-016-0040-5>
- Nathan, M. J., Schenck, K. E., Vinsonhaler, R., Michaelis, J. E., Swart, M. I., & Walkington, C. (2021). Embodied geometric reasoning. *Journal of Educational Psychology*, 113(5), 929–948. <https://doi.org/10.1037/edu0000638>
- Ng, O.-L. (2016). The interplay between language, gestures, dragging and diagrams in bilingual learners' mathematical communications. *Educational Studies in Mathematics*, 91, 307–326. <https://doi.org/10.1007/s10649-015-9652-9>
- Pier, E. L., Walkington, C., Clinton, V., Boncoddò, R., Williams-Pierce, C., Alibali, M. W., & Nathan, M. J. (2019). Embodied truths: How dynamic gestures and speech contribute to mathematical proof practices. *Contemporary Educational Psychology*, 58, 44–57. <https://doi.org/10.1016/j.cedpsych.2019.01.012>
- Pólya, G. (1945). *How to solve it: A new aspect of mathematical method*. Princeton University Press.
- Praetorius, A.-K., McIntyre, N. A., & Klassen, R. M. (2017). Reactivity effects in video-based classroom research: An investigation using teacher and student questionnaires as well as teacher eye-tracking. *Zeitschrift für Erziehungswissenschaft*, 20(1), 49–74. <https://doi.org/10.1007/s11618-017-0729-3>
- Radford, L. (2009). Why do gestures matter? Sensuous cognition and the palpability of mathematical meanings. *Educational Studies in Mathematics*, 70(2), 111–126. <https://doi.org/10.1007/s10649-008-9127-3>
- Sabena, C. (2008). On the semiotics of gestures. In L. Radford, G. Schubring, & F. Seeger (Eds.), *Semiotics in Mathematics Education: Epistemology, History, Classroom, and Culture* (Vol. 1, pp. 19–38). Sense. https://doi.org/10.1163/9789087905972_003

- Salminen-Saari, J. F. A., Moreno-Esteva, G., Haataja, E., Toivanen, E. S. H., Hannula, M., & Laine, A. (2021). Phases of collaborative mathematical problem solving and joint attention: A case study utilizing mobile gaze tracking. *ZDM – Mathematics Education*, 53, 771–784. <https://doi.org/10.1007/s11858-021-01280-z>
- Schindler, M., & Lilienthal, A. J. (2019). Domain-specific interpretation of eye tracking data: Towards a refined use of the eye-mind hypothesis for the field of geometry. *Educational Studies in Mathematics*, 101(1), 123–139. <https://doi.org/10.1007/s10649-019-9878-z>
- Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (2018). Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. *International Journal of Computer-Supported Collaborative Learning*, 13, 241–261. <https://doi.org/10.1007/s11412-018-9281-2>
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. Academic Press.
- Shvarts, A. (2018). Joint attention in resolving the ambiguity of different presentations: A dual eye-tracking study of the teaching-learning process. In N. Presmeg, L. Radford, W.-M. Roth, & G. Kadunz (Eds.), *Signs of signification: Semiotics in mathematics education research* (pp. 73–102). Springer. https://doi.org/10.1007/978-3-319-70287-2_5
- Strohmaier, A. R., MacKay, K. J., Obersteiner, A., & Reiss, K. M. (2020). Eye-tracking methodology in mathematics education research: A systematic literature review. *Educational Studies in Mathematics*, 104, 147–200. <https://doi.org/10.1007/s10649-020-09948-1>
- Svensson, M., & Johansen, G. (2019). Teacher's didactical moves in the technology classroom. *International Journal of Technology and Design Education*, 29, 161–176. <https://doi.org/10.1007/s10798-017-9432-1>
- Toivanen, M., Lukander, K., & Puolamäki, K. (2017). Probabilistic approach to robust wearable gaze tracking. *Journal of Eye Movement Research*, 10(4). <https://doi.org/10.16910/jemr.10.4.2>
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore, & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Lawrence Erlbaum Associates.
- Wakefield, E., Novack, M. A., Congdon, E. L., Franconeri, S., & Goldin-Meadow, S. (2018). Gesture helps learners learn, but not merely by guiding their visual attention. *Developmental Science*, 21(6), e12664. <https://doi.org/10.1111/desc.12664>
- Wathne, U., & Carlsen, M. (2022). Third grade students' multimodal mathematical reasoning when collaboratively solving combinatorial problems in small groups. *Mathematical Thinking and Learning*, 1–20. <https://doi.org/10.1080/10986065.2022.2099611>

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

We confirm that this manuscript has not been published or submitted for publication elsewhere.