



Teachers' Professional Vision: Teachers' Gaze During the Act of Teaching and After the Event

Eva Minarikova^{1*}, Zuzana Smidekova^{2,3}, Miroslav Janik¹ and Kenneth Holmqvist^{4,5,6}

¹Institute for Research in School Education, Faculty of Education, Masaryk University, Brno, Czechia, ²Department of Educational Sciences, Faculty of Arts, Masaryk University, Brno, Czechia, ³HUME Lab—Experimental Humanities Laboratory, Faculty of Arts, Masaryk University, Brno, Czechia, ⁴Department of Psychology, Nicolaus Copernicus University, Torun, Poland, ⁵Department of Computer Science and Informatics, University of the Free State, Bloemfontein, South Africa, ⁶Department of Psychology, Regensburg University, Regensburg, Germany

OPEN ACCESS

Edited by:

Manpreet Kaur Bagga,
Partap College of Education, India

Reviewed by:

Laura Sara Agrati,
University of Bergamo, Italy
Mandeep Bhullar,
Bhutta College of Education, India

*Correspondence:

Eva Minarikova
minarikova@ped.muni.cz

Specialty section:

This article was submitted to
Teacher Education,
a section of the journal
Frontiers in Education

Received: 28 May 2021

Accepted: 05 August 2021

Published: 03 September 2021

Citation:

Minarikova E, Smidekova Z, Janik M
and Holmqvist K (2021) Teachers'
Professional Vision: Teachers' Gaze
During the Act of Teaching and After
the Event.
Front. Educ. 6:716579.
doi: 10.3389/feduc.2021.716579

To date most of our knowledge on professional vision has relied on verbal data or questionnaires that used classroom videos as prompts. This has been used to tell us about a teacher's professional vision. Recently, however, new studies explore professional vision during the act of teaching through the use of mobile eye-tracking. This novel approach poses the question: how do these two "professional visions" differ? Visual attention represented by gaze was used as a proxy to studying professional vision (specifically its noticing component). To achieve this, eye-tracking as a data collection method was used. We worked with three teachers and employed eye-tracking glasses to record teacher eye movements during teaching (4 lessons per teacher; labelled as IN mode). After each lesson, we selected short clips from the lesson recorded by a static camera aimed at pupils and showed them to the same teacher (i.e., providing a similar setting as traditional studies on professional vision) while recording eye movements and gaze behavior data through a screen-based eye-tracker (labelled as ON mode). The two modes differ and due to these differences, comparison is difficult. However, by overlaying them and describing them in detail we want to highlight the exact variance observed. A comparison between IN vs ON condition in terms of dwell time on the same students in either condition was made using both quantitative (correlation) and qualitative (timeline comparison) methods. The findings suggest that the greatest differences in attention given to individual pupils occur when a pupil who was interacted with during the situation is missing from the view in the video recording. Even though individual differences are present in the patterns of gaze in IN and ON modes, the teachers in our sample consistently monitored more pupils more often in the ON mode than in the IN mode. On the other hand, the IN mode was mostly characterized by focused gaze on the pupil that the teacher interacted with in the moment with few side glances. The results aim to open a discussion about our understanding of professional vision in different contexts and about how current research may need to expand its outlook.

Keywords: gaze, eye-tracking, professional vision, teacher, remote, mobile

INTRODUCTION

Professionals perceive events and scenes from their own domain differently from laypeople. This has been the motto behind most research on visual expertise, professional vision, noticing, situational awareness, and other related phenomena. The term professional vision was coined by anthropologist Charles Goodwin in 1994 and since has been studied in various fields including architecture (Styhre, 2010), forensic science (e.g., Mustonen et al., 2015), medicine (e.g., Gegenfurth et al., 2013), and floristry (e.g., Gåfvæls, 2016).

It gained great momentum in research on teachers after 2000 with studies by van Es and Sherin (e.g., Sherin, 2001; Sherin, 2007; Es and Sherin, 2002; Sherin et al., 2009) and has been a household name ever since in the field. The rationale for studying professional vision is clear: teaching is a demanding profession, and teaching situations are characterized by their simultaneity, multidimensionality, and immediacy (Doyle, 1977). It is thus vital to understand how teachers orientate themselves in the plethora of stimuli in the classroom environment.

The term professional vision, however, does not refer solely to visual perception. It is a cognitive concept that encompasses both perceiving events and making sense of them (similarly to situation awareness—Endsley, 2015). In research on teachers, it has been mostly investigated through verbal data, i.e. teachers were asked to comment on classroom situations, usually captured on video (either their own or someone else's teaching). Thus, most of what we know about teachers' professional vision is what teachers think and report outside the act of teaching.

In general, so far it has been difficult to study teaching *in situ* from the perspective of teachers. Observation studies (direct observation, video studies) offer an outsiders' perspective, stimulated recalls provide insights removed from the immediacy of the classroom. Recently, eye-tracking has been introduced as a method to study teacher's perspective in the moment of teaching (Cortina et al., 2015; Dessus et al., 2016; McIntyre et al., 2017; Stürmer et al., 2017; Pouta et al., 2020; Smidekova et al., 2020). The rationale behind using eye-tracking studies is the eye-mind hypothesis, or what is attended to by the eye is processed by the mind (Duchowski, 2007, s. 3). However, as a sole method (without, e.g., stimulated recall) eye-tracking can only offer a partial perspective on what we understand as teachers' professional vision.

To bridge the gap between professional vision studied after the event and teachers' attentional processes when teaching, we use the potential of eye-tracking to investigate the visual attention (as represented by gaze) teachers give to classroom events in these two modes. Visual attention underlies how teachers comment on videos of classroom events (professional vision after the event; Wolff et al., 2020) and also attentional and decision-making processes in the classroom (eye-mind hypothesis).

In this study, we will refer to these two modes as ON action (visual attention that “underlies” professional vision as investigated in previous research; after the event, looking back on the situation) and IN action (visual attention that teachers deploy in the actual act of teaching).

RESEARCH ON PROFESSIONAL VISION

Professional vision in teacher research is usually defined through selective attention and knowledge-based reasoning (Sherin, 2007, p. 384; Seidel et al., 2010, p. 297). Selective attention (sometimes referred to as noticing) can be described as the process of identification of situations and events that are, from the professional point of view, instrumental for the success of pedagogical action (Seidel et al., 2010, p. 297). Knowledge-based reasoning represents the processes of making sense of situations and thinking about them, it presupposes certain knowledge (Seidel et al., 2010) or understanding (Sherin, 2007). These two components of professional vision are, however, interrelated and are applied in a cyclical process. Teachers direct their attention based on their reasoning, and reason about things they give attention to (Sherin et al., 2011, p. 5).

Visual attention as represented by teacher gaze *underlies* the selective attention (noticing) component of professional vision and thus professional vision in itself. Although the word “vision” is used (and some literature equals visual attention and professional vision; e.g., Nückles, 2020), the concept of professional vision as used in teacher research is broader, including also the reasoning processes. Gaze does not guarantee interpretation (although it often happens that we interpret what we look at). Gaze does not guarantee that the likely interpretation is related to one's profession. However, “teacher gaze can serve as an additional operationalization of the noticing component of teacher professional vision” (Seidel et al., 2020, np.).

In the original paper on professional vision by Goodwin (1994), the way professionals view situations is studied as they carry out the tasks of their profession and communicate while performing said tasks (e.g., archeologists working on a dig). However, when adapted for the study of teachers, the setting moved from the actual acting within the profession (i.e., carrying out the tasks) to talking about the acting (talking about carrying out the tasks ex-post; Sherin, 2001). This was given by the nature of the profession—teachers can hardly teach and at the same time comment on their actions. That is why professional vision outside the act of teaching became an established concept in teacher research without being so named.

Professional Vision Outside the Act of Teaching

Previous research on professional vision has yielded important insights into teachers and their thinking, development, and experience. This can be summarized as follows.

Professional vision is closely connected to teachers' professional knowledge (Stürmer et al., 2013; Meschede et al., 2017). This is intrinsic to the study of professional vision. One of the components of professional vision has been conceptualized as knowledge-based reasoning, acknowledging the connection (Sherin, 2007; Seidel et al., 2010). Stürmer et al. (2013) have shown that developing student teachers' professional knowledge influences their professional vision. This is in line with theories of

visual attention that posit it involves both bottom-up processes (i.e., is determined by the stimulus) and top-down processes (i.e., higher cognitive functions; cf. Ashcraft, 2006, p. 97; Duchowski, 2007, p. 12; Holmqvist and Andersson, 2017).

Professional vision is different for teachers of different subjects. Although research on the subject-specific aspects of professional vision is rather limited, it has been shown that the professional vision of student teachers of science and mathematics is different from the professional vision of student teachers preparing to teach school subjects within the domains of social sciences and humanities (Blomberg et al., 2011).

Professional vision can be influenced through various interventions. Research studies have been conducted to gauge the changes in professional vision after different types of intervention; typically this is video-based (Star and Strickland, 2008; Sherin et al., 2009; Minarikova et al., 2015 etc.).

Professional vision develops with experience. Studies on the differences between experienced teachers and novices/student teachers have shown that beginning teachers focus more on the teachers' actions and classroom management when observing video sequences of classroom situations (Sonmez and Hakverdi-Can, 2012) and they tend to describe them (Gonzalez and Carter, 1996). More experienced teachers usually pay attention to the pupils more (especially when watching video sequences of their own teaching; cf. Minarikova et al., 2015) and interpret the events portrayed in the video sequence (Es and Sherin, 2006), see the connections between them and try to suggest alternative courses of action (Copeland et al., 1994).

Of interest to the present study, however, is how professional vision has been investigated. A literature review shows that most studies use a kind of verbalization to explore the concept (c.f. also Jacobs, 2017). Most studies that we analyzed used either a questionnaire (e.g., Observer; Blomberg et al., 2013), an interview (Es and Sherin, 2006), or a written commentary (Vondrova and Zalska, 2012). Only four of the studies used eye-tracking, and only two of those in the act of teaching (Pouta et al., 2020; Seidel et al., 2020).

As an input provided for the verbalization, video sequences of teaching are used in most of the studies we identified. Three types of video sequences can be differentiated connected to the data collection method selected. Many studies use a video of a teacher and a class unknown to the research participant ("other video"), especially in connection with standardized questionnaires such as Observer (Blomberg et al., 2013). Some studies use video of the participant's own teaching ("own video"), usually with interviews (Choppin, 2011) or written commentary (Johannes and Seidel, 2012). The third type of video is mainly used in research within video clubs, i.e., groups of teachers who meet regularly to discuss classroom videos. In this case, the video sequence is a so-called peer video or a video of a person from within the group (e.g., Sherin et al., 2009).

In recent years, eye-tracking has become a novel tool for studying teacher's attentional processes. Although not always referred to as professional vision in the studies, we can assume that where teachers look is one of the prerequisites to processes

underlying professional vision, or that teacher's visual perception can be seen as an important aspect of the noticing component of professional vision (Lachner et al., 2016 as cited by Seidel et al., 2020, np.; and is part of the perception aspect of situation awareness—Endsley, 2015). Screen-based eye-trackers have been used in this respect, attempting to capture the differences between the experienced and the novice teachers when monitoring a video of a teaching situation. Different studies, however, yield different findings. Some studies show different monitoring behavior by more and less experienced teachers (van den Bogert et al., 2014; Wolff et al., 2016; Seidel et al., 2020; Wyss et al., 2020) in that experienced teachers check up on pupils more often and have shorter fixations on students (which indicates faster processing), but also monitor different aspects of the situation (some areas of the visual field tend to be skipped by novices even though they are revisited by experts). On the other hand, Yamamoto and Imai-Matsumura (2013) found no relationship between experience and the ability to notice negative behavior in a class.

An interesting contribution to the field is a study on how teachers observe lessons directly (Egi et al., 2014). These authors asked a group of people with a varying degree of teaching experience to observe a lesson as it was taught. They gave the observers eye-tracking glasses and monitored their eye-movements. This study is an interesting combination of real-life conditions while studying professional vision outside the act of teaching.

Visual Attention in the Act of Teaching

Whereas all the studies mentioned previously provide an invaluable insight into teachers' noticing and thinking, it is clear that the conditions under which professional vision was examined in those studies differ from actual teaching situations. That is why Sherin et al. (2008) attempted to target teacher's professional vision in action (sic) in their exploratory study. Using a wearable camera, they captured the teacher's perspective. To tap into what they found important, the camera only recorded sequences highlighted by the teacher (the teacher had a remote to record anything interesting happening in the lesson and the camera on a loop stored the previous 10 s segment). These sequences were then commented on by the teacher, providing self-selected input for verbal recall.

Recently, a new possibility of capturing a teacher's perspective has emerged, mobile eye-tracking. Eye-tracking glasses allow the capture of teacher's gaze (which often coincides with visual attention) in the act of teaching and, in connection with verbal protocols or classroom observations, can provide deeper insight into actual classroom processes.

Two types of studies can be distinguished, one focuses on real-life classrooms, the other uses altered settings. An example of the latter is a study by Stürmer et al. (2017) that monitored student teachers' eye movements during standardized instructional situations during their university studies. The pupils were played by university students. There were only four of them in the group, each with an assigned pupil profile (strong, weak, uninterested, underestimating pupil). The authors found, in line with expertise research, that the student teachers do not display a

clear pattern of monitoring their classes but differ strongly from one another.

One of the first types of studies, from real-life classrooms, combined data about teachers' visual attention with high-inference data on the quality of instruction and identified differences in teacher feedback. While novice teachers tended to give feedback in a more intimate setting (approaching the pupil, lowering their voice, etc.), experienced teachers shared the feedback with the entire class. This was mirrored in the perception processes—while novice teachers focused on the pupil only, experienced teachers were able to monitor the class while still giving individual attention to the one pupil (Cortina et al., 2015 p. 400). This was corroborated by Dessus et al. (2016), who looking at the relationship of teacher gaze behavior and classroom climate found a small relationship between experience and gaze behavior. McIntyre et al. (2017) showed novice-expert differences in gaze behavior in real-life classrooms in that experts tend to look at students longer while maintaining greater gaze efficiency than novices. Experienced teachers also tend to prioritize students in their gaze over noninstructional regions (McIntyre et al., 2019). However, there are individual differences even among experienced teachers (Smidekova et al., 2020).

Another study focusing on visual attention in action dealt with the specific demands technology-enhanced classrooms place on the teacher and their cognitive load (Prieto et al., 2015). It has been shown that high cognitive load episodes were connected to class-wide interactions, such as explaining or questioning (p. 278), and to the need to make a decision about the subsequent course of action (p. 275). On the other hand, episodes characterized by low cognitive load tended to pertain to individual explanations (p. 275).

AIMS OF THE STUDY

Research on teacher's professional vision is generally justified by referring to the immediacy, simultaneity, and unpredictability of classroom situations (as described by Doyle, 1977). Professional vision is usually defined as the ability to identify and reason about features of classroom situations that are pertinent from an educational perspective (e.g., Sherin and Han, 2004, p. 179; Stürmer et al., 2013, p. 468). The definitions do not usually specify how these situations are perceived—whether they are experienced first-hand or related to the research participant in a different way (verbal account, video). However, this is crucial as these two modes of experiencing the classroom situation differ greatly. "After the event" research (see the previous section) removes both the immediacy of the situation and the need to act. Also, after the event research offers a predefined perspective selected by the researcher (represented by the camera angle). During the event research, on the other hand, poses many technical challenges, as mentioned in the previous section, and is only recently gaining momentum.

These two strands of research offer important insights into teachers' thinking and acting. To bridge the gap between the two and to further our understanding of professional vision in the

broad sense (noticing and reasoning of teachers in their professional capacity, regardless of whether they are teaching or observing), we need to understand the differences of the two modes. One way to do this is to study the visual attention (represented by gaze) that is a prerequisite for reasoning and acting in the two modes.

In our study, we aim to compare, from a qualitative point of view, teacher's gaze (as captured by eye-tracking technology) in the same situations in the act of teaching and after the event. We take this to represent a vital component of professional vision (see Section *Research on Professional Vision*). In our research design, we asked teachers to wear eye-tracking glasses during teaching, while we videotaped the lessons. After each lesson, we selected clips from the video recordings and showed them to the same teachers, recording their eye movements with a remote eye-tracker. In doing so, we captured gaze in the same situation in two different modes (IN vs. ON action).

From a research perspective, the two modes differ by default. They differ in viewing angles, in the freedom to choose where to look in the classroom, the need to act (see **Table 1**). These differences are intrinsic to our research—they copy the approach to studying visual attention and professional vision in previous studies (see above). When researching professional vision using reactions to videotapes of classroom events (specifically video of the participant's own teaching), these reactions are used as a proxy for understanding professional vision (often without acknowledging that this is solely professional vision ON action; e.g., Sherin and Han, 2004; Minarikova et al., 2015). Participants are presented with a different viewing angle to what they experienced, their freedom to choose the area of the classroom to look at is taken away, the situation is not new to them. That is why our research design provides an important insight into the visual attention underlying professional vision IN and ON action that has not been previously addressed.

Eye gaze is task specific (Yarbus, 1967, p. 174). We limited recordings to whole class interaction situations in teaching of English as a foreign language, which comprises the task for our participants. These situations are communicative in nature as they by definition include interaction between the teacher and the class. This generates a different gaze behavior than, for example, lecturing a group of pupils with no mutual communication (cf. Seidel et al., 2020). The task specificity, in our case, then lies at the intersection of the subject taught (English as a foreign language) and the definition of the situation (whole class interaction).

To compare the two modes (IN vs. ON action), we focus on a qualitative investigation of:

What are the characteristics of situations (including ON mode video features) that are associated with greatest/smallest differences in gaze distribution among pupils between IN and ON modes?

What are the characteristics of monitoring behavior (gaze distribution) in the ON mode when presented with a different field of view as/similar field of view than in the IN mode?

TABLE 1 | Research differences of visual attention IN and ON action.

Visual attention IN action (research using eye-tracking glasses)	Visual attention ON action (research input: own video)
<ul style="list-style-type: none"> • High ecological validity • Unlimited visual field • The participant decides where and when to look • Situation experienced first-hand • Need for immediate action • No concurrent verbal protocol to capture teacher thinking is possible • Influence of wearable device on pupils and on teacher 	<ul style="list-style-type: none"> • Lower ecological validity • Scene, angle, duration of the video typically determined by the researcher (some pupils more visible than others due to camera positioning) • The situation had been experienced before (influence of memory) • No need for immediate action • Concurrent verbal protocols available but providing both “recall” (what the teacher was thinking in the time of teaching) and “reflection” (what the teacher had time to consider afterward) • Influence of stationary cameras that were present during classroom video recording

TABLE 2 | Research sample description.

	T1	T2	T3
Education	Teacher education (Technical Education and Civic Education), MA degree	Teacher Education (Civic Education and Geography), MA degree Teacher Education (English as foreign Language), unfinished	Teacher Education (English as foreign language), MA degree
Teaching experience at schools (English as foreign language)	12 years	10 years	6 years
Other teaching experience (English as foreign language)	No	No	Courses in language schools, individual courses for adults and children
Classroom of participation	Grade 4 (pupils aged 9–10)	Grade 6 (pupils aged 11–12)	Grade 7 (pupils aged 12–13)
Number of pupils	16 pupils (5 girls, 11 boys)	24 pupils (13 girls, 11 boys)	16 pupils (8 girls, 8 boys)
Pupil characteristics	3 extra English, 3 special needs, 4 poor performance	1 extra English, 2 special needs, 9 poor performance	4 extra English pupils, 2 special needs, 5 poor performance

What are the general differences in monitoring behavior between the ON and IN modes (irrespective of field of view) and how can they be explained?

METHODOLOGY

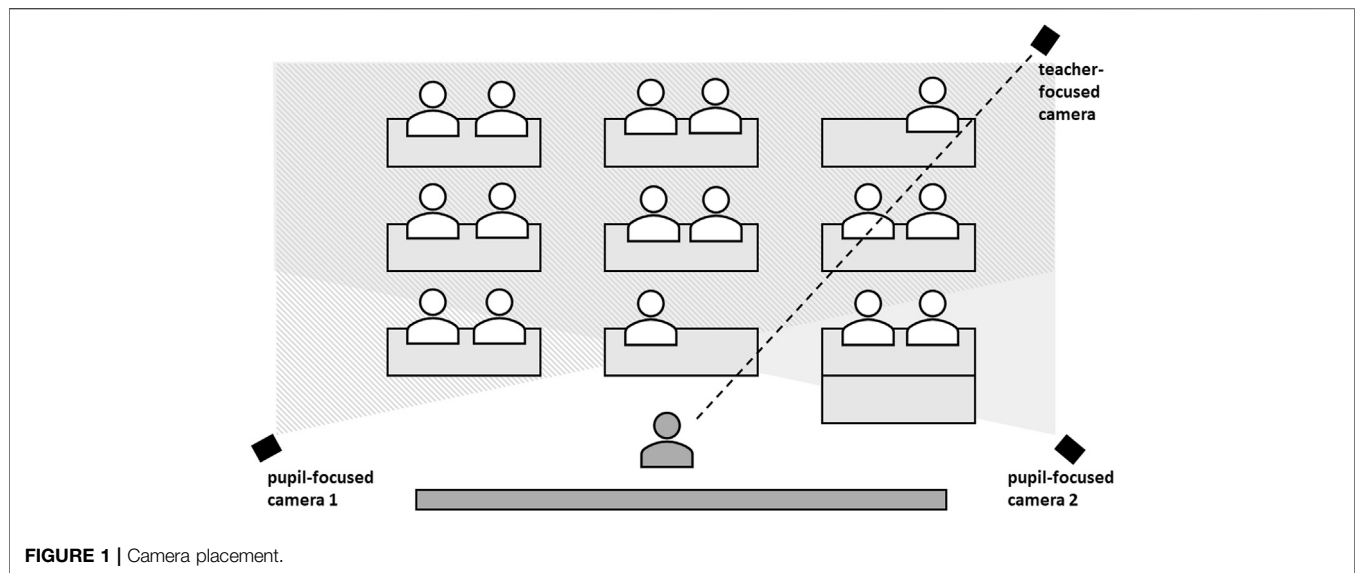
Research Sample

The research sample of this study consists of three experienced practising teachers of English as a foreign language at primary and lower secondary schools. Participating teachers were chosen via recommendation of the school headmaster. All teachers work in the same school located in a small town in the South Moravia Region, Czechia. Each teacher selected one class to participate in the study. The characteristics of the teachers and their classes are provided in **Table 2**. The teachers had no or very limited experience with using classroom videos for professional development and no prior experience with eye-tracking. The research sample is fairly homogeneous, with the teachers having similar education, similar length of teaching experience, and being from the same school (Nückles, 2020, np.).

The participants were asked to provide information about certain pupil characteristics—pupil’s performance, special needs, and extra English. Pupil’s performance is based on their overall grade in last semester; good = 1 or 2; poor = 3 or 4; no 5 (i.e., failed) pupils were among our participants. Pupils with special needs are those officially diagnosed with a learning disability such as dyslexia, or other conditions such as autism. The last category, extra English, includes pupils who display above average English due to non-school factors (extra English classes, specific hobbies, or bilingual upbringing). More information is provided in **Table 2**.

Data Collection

The data were collected in autumn 2018. Eye-tracking was used as a data collection method. The data collection process consisted of four individual recordings of a lesson for each teacher, followed by four interviews. During each lesson (approx. 45 min long), the teacher wore SMI Eye Tracking Glasses 2 Wireless (referred to as ETG; 60 Hz). Before each recording, a 3-pt calibration was used to ensure calibration accuracy. Participants were asked to fixate on three locations in the classroom (within a participant’s field of



view). Alongside this, three standard video cameras were used to capture the lesson overall, one aimed at the teacher (operated by a researcher), two were static aimed at the pupils (see **Figure 1**).

After each lesson, the researchers present in the lesson selected two short situations (1–2 min long) portraying a communicative situation (either teacher talking to the whole class or pupils working in pairs). They edited the video from one of the pupil-focused cameras and the same sequence from the eye-tracking glasses video (showing the gaze point).

This was followed by an interview with the teacher. First, the researcher showed the teacher the video sequence from a static pupil-focused camera. During the viewing, teacher's eye movements were monitored using SMI RED250 mobile System (referred to as REMOTE). After the viewing the teacher was invited to comment on the situation. Subsequently, the sequence from ETG was shown (no eye-movement monitoring) and the teacher was asked to comment.

For this study, the following sources of data were used:

- Data from the ETG in the selected situations;
- Data from the REMOTE pertaining to the same situation;
- Video recordings of the situations;
- Interviews with the teachers.

To ensure data consistency, only situations portraying whole class work were selected. The ETG data are taken to represent the IN action condition, the REMOTE data represent the ON action condition—they pertain to the SAME classroom situations. The data include 11 situations (4 from each teacher; one situation was omitted due to missing data) with two data sets each (IN action x ON action). The situations are 1–2 min long.

Data Analysis—Verbal Data

The interviews with teachers were videotaped and transcribed. To provide further insight into the differences in gaze in the two modes, we identified the statements in the interviews that

comment on the pupils observed, e.g., “I noticed Jane was not paying attention” (T2; pupil pseudonym used) or “everyone seems to have been engaged” (T1). Furthermore, we looked for statements that suggest what drove teachers' attention.

I looked to one side, then the other, so you look at both. And then I noticed I wanted to call on the boy in the first row [pupil Q], but in the meantime the girl sitting behind him raised her hand, so I jumped over and called on her [pupil S]. [. . .] So you actually really pay attention to someone who raises their hand or catches your eye somehow. (T2, when commenting on replay from ETG camera).

Data Analysis—Eye-Tracking Data

The data were analyzed using SMI BeGaze software (see **Table 3**). From the visual field available to the teacher both in IN and ON action conditions, areas of interest (AOI) were selected that pertain to the teaching situation as such—individual pupils, their materials, and board/teacher materials. These AOIs were chosen as they are of value to understanding the situation from an educational point of view. Each pupil is an independent unit who can interact with the teacher or attract their attention. Materials were coded separately as they can provide different kinds of clues to the teacher about how a pupil is coping with their work.

When coding AOIs for each individual pupil, gaze aiming at their face or body above the desk was considered. For pupil materials, the area on each pupil's desk was included. For teacher materials, gaze aiming at the board or teacher's book or other materials was included (**Figure 2**). Each pupil was marked with a letter, materials were marked with M, and pupil's faces and bodies with P (**Figures 2, 3**).

ETG Data Coding

In the first step, the precise timing of each situation was marked in BeGaze so that it corresponds with the situation selected for REMOTE

TABLE 3 | Analysis of eye-tracking data.**Preparation for analysis**

Software: BeGaze

AOIs defined: individual pupils, their materials, and board/teacher materials

Situations defined: Those situations that were selected to be viewed by teachers during interview

ETG data analysis (for each situation)

Time marked to correspond with situations viewed by teachers in the REMOTE viewing
 Preparation of pre-drawn coding image for SMI Semantic Gaze Mapping (see **Figure 3**)
 Manual coding based on defined AOIs using SMI Semantic Gaze Mapping feature
 Inter-coder reliability checked
 Manual coding of data by one researcher
 Dwell time values extracted for AIOs for each situation
 AOI vs. Time graphs extracted for each situation

REMOTE data analysis (for each situation)

Preparation of dynamic AOIs (polygons with a variable number of corners)
 Manually adjusting each polygon for each video frame in BeGaze
 Inter-coder reliability check
 Manual coding of data by one researcher
 Dwell time values extracted for AIOs for each situation
 AOI vs. Time graphs extracted for each situation

Final analysis

AOI vs. Time graphs (ETG, REMOTE) combined for each situation (see **Figure 4**)
 Correlation values (dwell time) between IN and ON modes for each situation calculated

**FIGURE 2** | AOIs marked in BeGaze AOI editor (T2, sequence 1; faces have been obscured for ethical reasons).

viewing. Eye-tracking data were then manually coded using SMI Semantic Gaze Mapping feature, following the guidelines stated previously. To ensure data quality, we checked inter-rater reliability of the coding. As material for inter-coder measurements, we used one trial (one situation) coded by both the coders. For IN action, the same reference view image with predrawn AOIs in SMI Semantic Gaze Mapping in BeGaze software was used by both the coders. The percent agreement between the coders was 88% (number of same code fixations out of all identified fixations). As the inter-coder agreement was satisfactory, the rest of the data were coded by one coder.

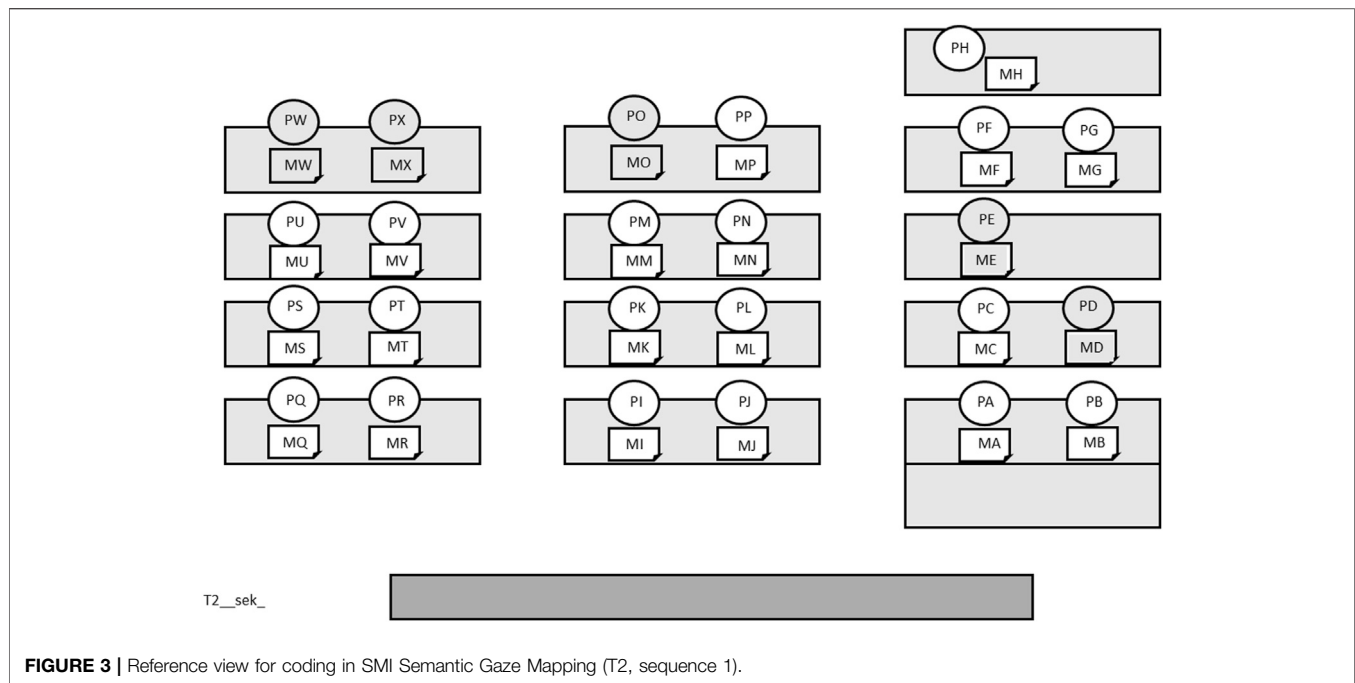
REMOTE Data Coding

For each situation, REMOTE data were analyzed using the dynamic AOI feature in BeGaze. Each AOI was traced individually (polygon option with variable number of corners) and adjusted for each video frame. Coding reliability might be an

issue with this coding procedure as in a classroom video, AOIs are typically placed near each other and must not overlap. Moreover, the variability of the AOIs is considerable as each AOI is adjusted for each frame. One video was thus coded by two independent coders. Each coder was given a starting frame with all AOIs traced out (ensuring the same number of corners for each polygon AOI). Two coders independently adjusted all the AOIs for each video frame in the selected trial. The percent agreement for this part was 92% (number of same code fixations out of all identified fixations). As the inter-coder agreement was satisfactory, the rest of the data were coded by one coder.

Eye-Tracking Data Analysis

A comparison between IN vs ON condition in terms of dwell time on the same AOI was made using both the quantitative and qualitative methods to analyze eye-tracking data. After the data was coded, we retrieved dwell time values for each AOI. As the



proportion of time spent looking at pupils' materials was negligible, we omitted these AOIs from further analyses. Correlations between dwell times for pupil AOIs in IN/ON mode were calculated for each situation separately. Board/teacher materials were not included as they are by default not represented in the ON condition.

Afterward, we retrieved BeGaze graphs AOI vs. time for each situation in both conditions. Each graph was then visually assessed separately in relation to the video recording of the situation. To understand how gaze was distributed timewise in the two modes (e.g., if the same AOIs were looked at at the same time in the situation), we manually combined the AOI vs. time graphs for the two modes—line for AOI for pupil A in the ON condition was placed underneath line for AOI for the same pupil in the IN condition etc. An example of this graph is available in **Figure 4**.

The dwell time values for pupil AOIs for each situation in the two modes, the correlation value, the visual assessment of all the graphs pertaining to each situation (IN, ON, combined), the video recordings, and the teacher's comments from the interview were then considered together and a vignette for each situation was created. We present these in Section Results - Situations.

These vignettes were then considered together and similarities, differences, and patterns were sought that highlight the differences in gaze in the IN and ON modes. Original sources of data were regularly consulted during this process. The results are described in Section Results - Differences in gaze.

RESULTS

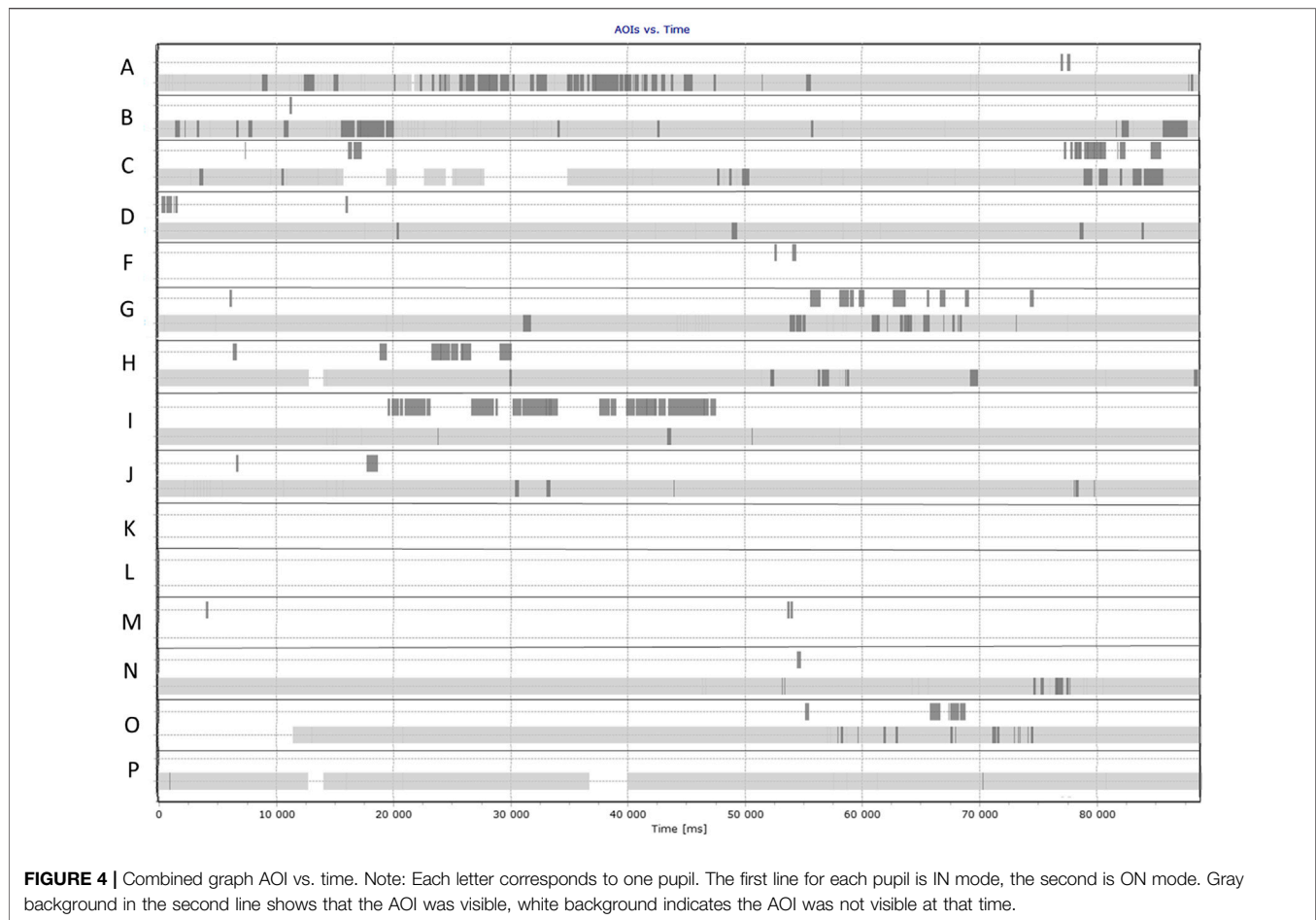
In this section, we are going to provide a detailed analysis of each situation based on the available data and then

look at the differences in gaze in the IN and ON modes in general.

Situations Teacher 1

In the first situation of T1 (see **Figure 5**), the teacher introduces a task from a textbook. First, she wants the pupils to look at a few pictures connected to the task. Then she elicits some information about the task. She calls on individual pupils. At the start, in the IN condition, she looks at her own textbook and briefly monitors the class if they are on the right page, quickly glancing from one pupil to another. Here is the biggest difference between the IN and ON conditions. As the teacher's textbook is not in sight in the ON condition, she focuses mainly on the pupils most visible (A and B). Approx. 20,000 ms into the situation, she starts interacting with pupil I, which is clearly visible in the IN data. However, in the ON condition, pupil I is partly obscured by pupil A; hence, the teacher tends to look at pupil B. After approx. 50,000 ms the direction of the gaze in IN and ON conditions is mostly similar. This is caused by the fact that the teacher stands next to the camera and interacts with pupils who are not obscured by other pupils in the video recording. During the interview, the teacher comments overall on all the pupils ("everyone was engaged"). She also addresses specifically pupil G who is generally weaker, expressing doubts over his engagement. Attention was paid to this pupil in both the IN and ON conditions during the interaction period.

In situation T1_2 (see **Figure 5**), there are cards with items of clothing on the board. After introducing the task briefly and looking mainly at pupils A and B, the teacher asks the pupils e.g., "What can you wear on your head?" and then elicits an answer. The pattern of looking is similar for all interactions in the IN

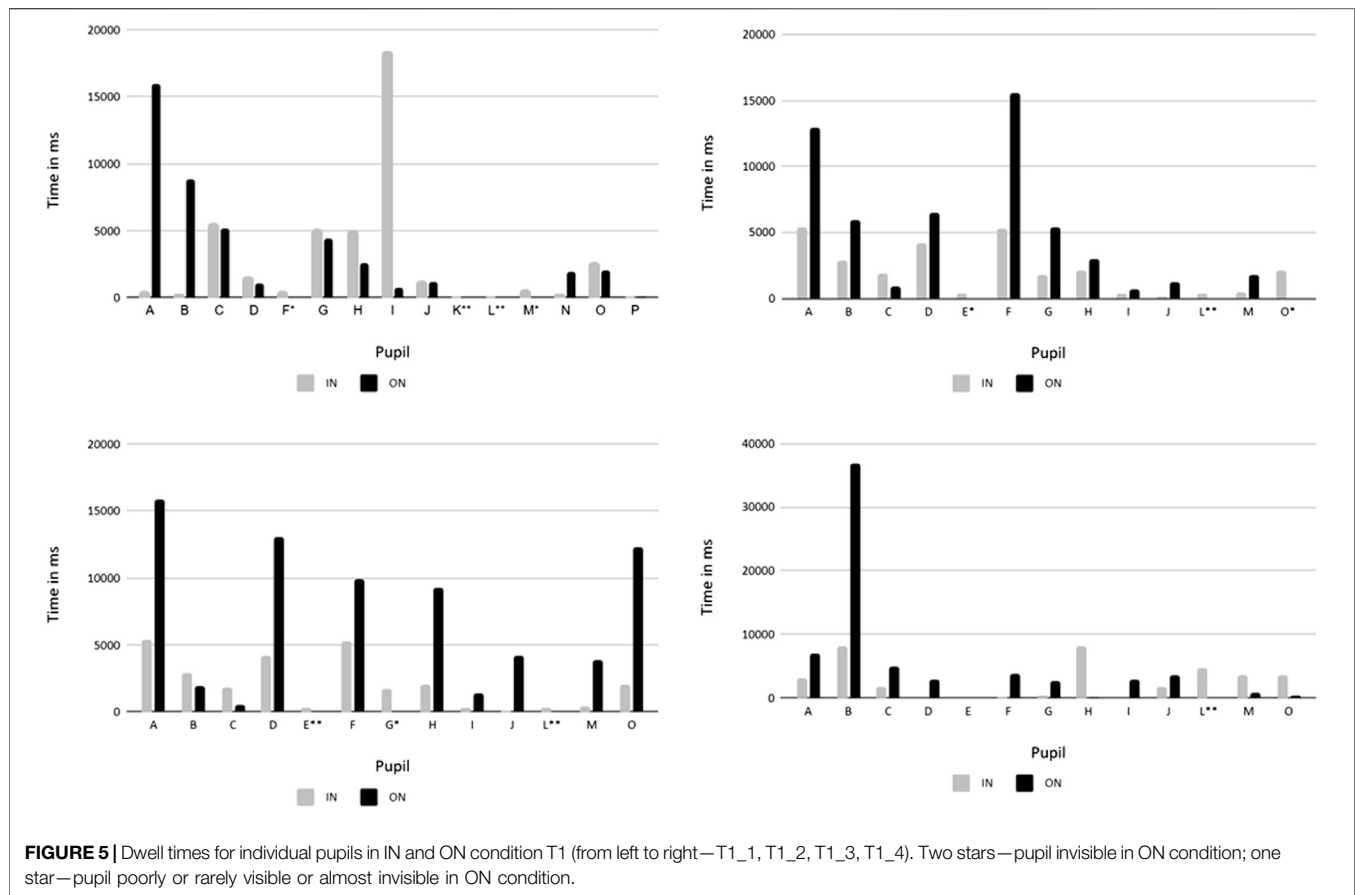


mode: she looks at the board, choosing a word to elicit, then scans the class and selects a pupil to interact with, giving them her full attention. Then she asks the whole class to repeat the word, again briefly scanning the class. This pattern is most obvious between 30,000 and 50,000 ms. The board is not visible in the ON condition so the teacher has more time to look at pupils than in the IN mode. She monitors almost all pupils with quick glances (at least four glances each; except those outside the camera angle or those hidden behind another pupil). This seems more prevalent than in the IN mode in which some pupils receive very limited attention (one or two brief glances—E, I, J, L, M). Quite a considerable amount of attention is given to pupil F who is visible but not right in front of the camera (pupils A and B). This might be explained by the fact that he is very active, raising his hand often, unlike pupils A and B who should in theory attract more attention because of being most visible. Those two, however, are very passive with little facial expression. In the interview she singles out pupil B, whom she paid some but not most attention to in both conditions, commenting he is one of the weaker pupils. Otherwise, she only notes that everyone seems to have been working.

Situation T1_3 (see **Figure 5**) is a continuation of vocabulary practice from T1_2. The teacher now asks more complex

questions, such as “When it’s a really hot day, what are you going to wear?” She first focuses mostly on one pupil (H) who struggles to answer her question, only sometimes looking at pupils around him (I, J). She does not divert her gaze toward pupils whose hands are raised. Sometimes she turns toward the board, presumably attempting to direct the pupil toward the right word card displayed. Then she turns toward another pupil (D) and then yet another (P). The pattern is clear—attention is given mostly to the pupil whom the teacher is interacting with. As the teacher interacts with pupils who are clearly visible in the ON action video, the pattern of attention is similar—attention is given to the pupils the teacher interacts with. An exception is interaction with pupil H at the start—the attention is divided between him and pupil A. This might be caused by the fact that they are in the same line of view (the head of pupil H is directly above the head of pupil A). Interestingly, when asked about pupils’ engagement in the interview, she comments on pupil L and his typical distractedness, even though he is not visible in the ON condition and is paid no attention in the IN condition.

In situation T1_4 (see **Figure 5**), the class work with their textbooks. On the page there is a set of pictures of people. Each pupil is meant to choose one person and describe them to the class who then guess which person the pupil is describing. The



teacher gives most of her attention to pupil B and his textbook. He is a weaker pupil and is struggling to provide the description. For almost 1 minute, the rest of the class (save pupils A and C who are near pupil B) receive no attention. After a period of time the teacher turns to the other pupils to elicit how to say “he has got” to help pupil B and looks at pupils H, L, M, and P. In the ON action video, the teacher mostly pays attention to pupil B and his materials again. Toward the end of the video the teacher turns her attention to pupils I and J—pupil J guesses the person being described. The attention pattern is similar in both IN and ON conditions, mostly focusing on pupil B who the teacher interacts with. However, she often switches to pupil A sitting next to B and scans the class much more in the ON condition. Limited attention is paid to pupil H as he is obscured by pupil A. During the interview, the teacher comments on how many pupils raised their hand in general, and then notes pupil B and his being able to say several sentences (despite being a weaker pupil). This is in line with the interaction pattern in the IN condition and with gaze in both conditions.

Teacher 2

In situation T2_1 (see **Figure 6**) the teacher elicits answers for a task in the textbook which she then writes up on the board. We can see a clear pattern of writing on the board and then calling on individual pupils. It is very interesting that in this situation,

the teacher does not “scan” the class when she selects a pupil to answer but rather selects immediately (focuses on one other pupil at most before turning her attention to the selected pupil). Toward the end of the sequence, the teacher turns from the board, looks at pupil Q for about 3 s, and then says his name. At the same time, pupil S who sits behind pupil Q raises her hand. The teacher immediately calls on her, without waiting for the answer from pupil Q. The teacher pays attention to the same pupils in the IN and ON conditions at the same time (except for pupils Q, S, and V who are not visible in the ON condition). A lot of attention is given to pupils A and B, mostly in times when she was looking at the board in the act of teaching. She also seems to monitor the whole class more in the ON condition (short glances to more pupils). Pupils G, H, J, and K are monitored, in contrast to no attention during the actual teaching. After viewing the video replay focused on pupils, the teacher singles out pupil B who did not pay attention. She also comments that during teaching, if she stands by the teacher’s desk (placed right against the desk of pupils A and B), her field of view encompasses the whole class but pupil B “is in the corner of the field and sort of blurred.” During the interview, an interesting contrast emerges. After viewing the pupil camera replay, the teacher comments that she noticed she calls on pupils with raised hands, even though usually calls on pupils who did not raise their hands in the particular interaction exchange

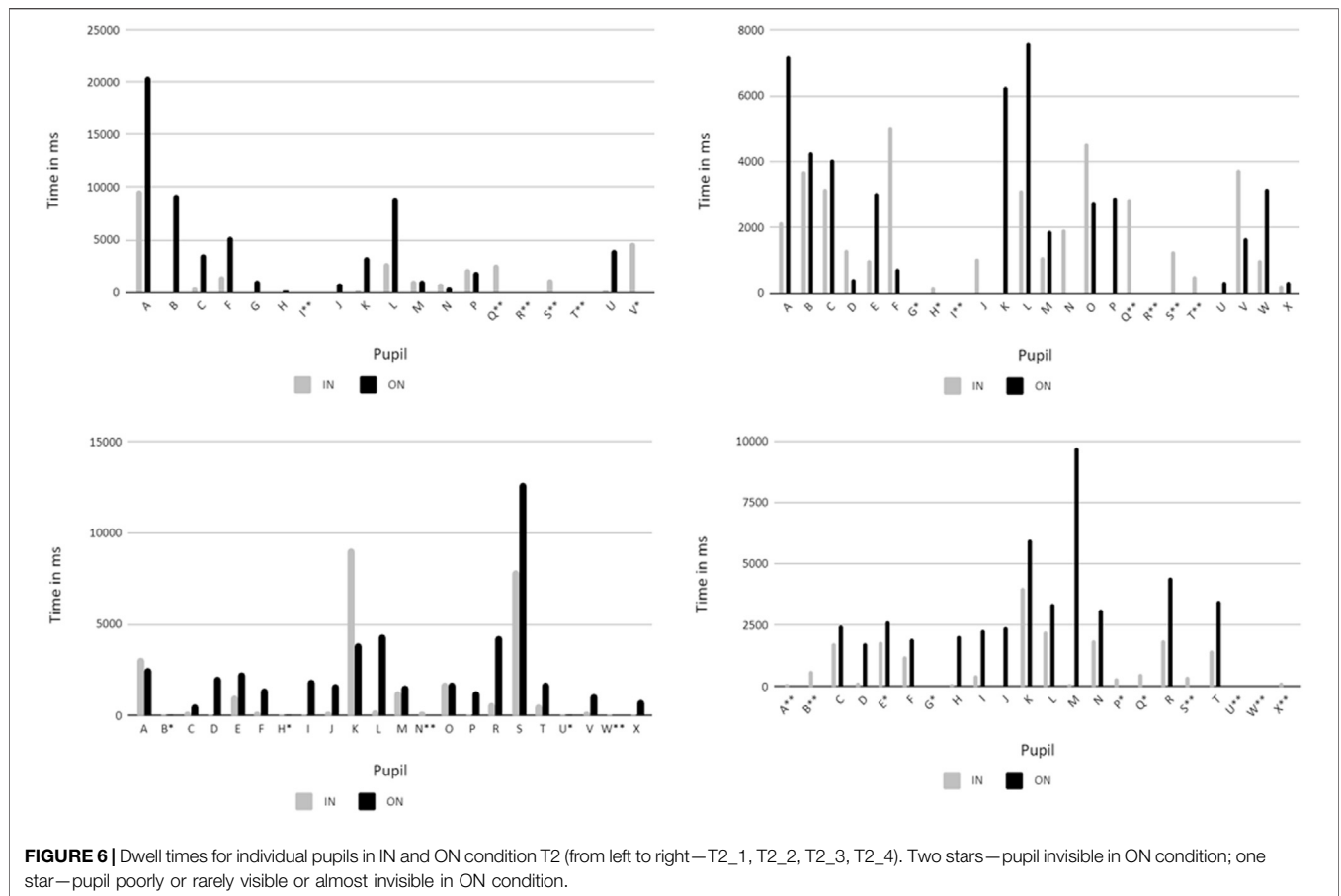


FIGURE 6 | Dwell times for individual pupils in IN and ON condition T2 (from left to right—T2_1, T2_2, T2_3, T2_4). Two stars—pupil invisible in ON condition; one star—pupil poorly or rarely visible or almost invisible in ON condition.

(even when they protest). However, when confronted with the ETG camera replay of the same situation, she observes:

“I looked to one side, then the other, so you look at both. And then I noticed I wanted to call on the boy in the first row, [pupil Q], but in the meantime the girl sitting behind him raised her hand, so I jumped over [from looking at him to looking at her], and called on her [pupil S]. [...] So you actually really pay attention to someone who raises their hand or catches your eye somehow.”

Interestingly, out of seven pupils she calls on in this sequence, three did not raise their hands. Only one is given a chance to answer; one is provided with moderate wait time (pupil P, 4 s), and one with no wait time before the teacher calls on another pupil. Pupil L was called on twice, both after she raised her hand. No pattern of difference between IN and ON mode gaze was discernible in this particular situation when gender, performance, and classroom position were taken into account.

In situation T2_2 (see Figure 6), the teacher elicits parts of the body from different pupils. Again, she gives attention to selected pupils. This time, however, she scans the class a little more between each turn (fixates up to five pupils before giving her attention to the one pupil). In the ON condition, the teacher’s

attention is more scattered (shorter attention spans distributed over more pupils). The pupils who had been given attention in the IN mode are given attention in the ON mode at approximately the same time (but for a shorter time), provided they are visible in the video (except for pupils D, N, and V). A lot of attention is given to pupils A, K, and L who are in a central position in the ON video. Pupils A and K are commented on in the interview specifically, highlighting their behavior (pupil A in connection with her overall characteristics). The teacher also comments on how everyone was engaged, pointing to the parts of the body as requested. While watching the replay of the ETG camera, she again discusses her overall motivation for giving pupils visual attention. This time she mentions both raised hands (as an overt reason) and her consideration of who might know the answer given their competence level (focus on weaker pupils given the simple nature of the task). She also notices that she distributes her visual attention over the whole class but singles out a pupil who had his hand raised for a long time but was not called on (in IN action, she only fleetingly looked at him once). Her reflection on visual attention also includes being easily distracted from monitoring the class, such as when an outside person comes into the class.

In situation T2_3 (see Figure 6), the teacher shows pupils a paper clock and elicits the time. In the IN condition, the teacher looks at the clock itself. Pupils being interacted with are looked at,

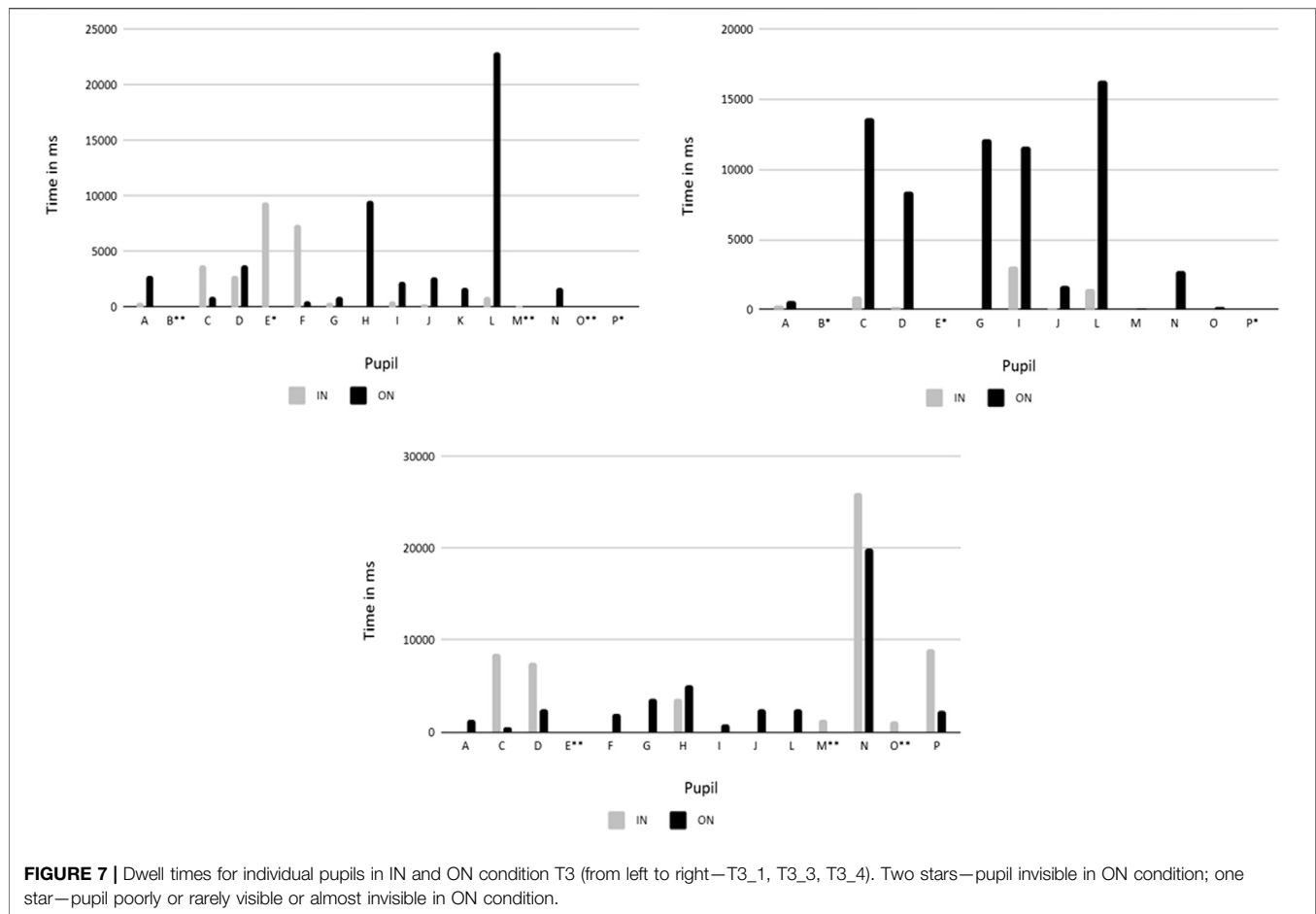


FIGURE 7 | Dwell times for individual pupils in IN and ON condition T3 (from left to right—T3_1, T3_3, T3_4). Two stars—pupil invisible in ON condition; one star—pupil poorly or rarely visible or almost invisible in ON condition.

with brief whole class monitoring at the start of the activity and some quick glances around a specific area of the classroom before calling on a pupil. In the ON mode, attention is scattered as the teacher monitors the whole class. Exception is a spell of attention given to pupil S who looks bored and plays with her pencil case and then struggles to answer the teacher's question (the teacher comments on her in the interview). This attention period takes place at the same time as in the video. Again, pupils being spoken with draw attention also in the ON mode (pupils S, K, A, E), but it is less pronounced than in the previous case. During the interview, the teacher notes the body posture of many of the pupils (slumped, resting their chins on their hands, lying on their desks) and how bored they look. She also says she focuses now more on pupil A who seems to be out of her field of vision and (after watching the ETG camera replay) says she will pay extra attention to her in the future. Again, she makes overall comments about what drives her attention in lessons (raised hands, looking at pupils she wants to call on, general monitoring at the start of the activity).

In situation T2_4 (see **Figure 6**) the teacher elicits days of the week (e.g., “What day was it yesterday?”) and asks pupils (N, E, Q) to write them up on the board. In the IN mode, a lot of visual attention is given to the board. Even though the teacher interacts with individual pupils, the interactions are rather short and

involve the teacher checking what the pupil is writing on the board. Attention is given to many pupils in short spells of time—the teacher monitors the class and selects a person to answer. These pupils are mostly positioned on the right-hand side or center of the classroom. Especially pupils K (in this and previous interviews commented on as distracted, not paying attention) and L (always raising her hand, even before a question is posed) are revisited many times despite not being called on. The answering pupils do not get prolonged visual attention time. In the ON mode, attention is also distributed among many pupils, no clear pattern can be discerned. A lot of attention and revisiting can be observed for pupils in the center of the class (e.g., pupil M). In the interview, the teacher comments specifically on pupils H, K, and T. None of these have been called on in this situation, and except for pupil H, they are well visible in the video.

Teacher 3

In situation T3_1 (see **Figure 7**) the pupils describe people on a worksheet. The teacher calls on one pupil and during their interaction two other pupils join in (F, E, C). Other pupils receive very little attention. The interacting pupils, however, are not visible in the ON mode. The teacher gives most attention to two pupils in the front row who are inattentive,

TABLE 4 | Correlations between dwell time per pupil in IN and ON conditions.

Situation	Highlights of the situation	Correlation
T3_1	Very different conditions; pupils interacted with in IN condition are not visible in ON condition (E, F—most interaction throughout the situation)	-0.24
T1_1	Different conditions; teacher is looking at her textbook, pupil who was interacted with in IN condition is obscured in ON condition	-0.16
T2_2	Different conditions because of pupils not being visible in ON mode (F, Q, W), a lot of attention to prominent pupils in ON condition	0.24
T2_4	Different conditions; board work, one pupil who was interacted with now out of view (E, for some time)	0.33
T1_4	Different conditions; teacher looking at materials (her own or pupil B's), two pupils who were paid attention to in IN condition out of view in ON condition (M, O)	0.53
T3_3	Very different conditions; lots of boardwork, some attention to the same pupils (C, I) but much more monitoring in ON condition	0.63
T1_3	Similar conditions; some attention taken from pupil A to pupil H, board gaze only limited	0.69
T2_1	Different conditions; a lot of board work in IN condition, pupil V who was interacted with not visible in ON condition	0.71
T2_3	Different conditions; teacher looking at a paper clock in IN condition; one pupil (A) who was interacted with not visible in ON condition, more monitoring	0.71
T3_4	Slightly different conditions; pupil C not visible in IN condition, more monitoring	0.85
T1_2	Different conditions to a certain degree; some board glances when selecting a word to elicit but most attention on pupils	0.89

display off-task behavior, and were outside her scope of vision during the IN mode. She comments on their behavior in the interview as being typical for them. In the interview, a slight difference is visible between commenting on the ON video and on the ETG replay video. In the former, the teacher says that “how the kids work here, so they work in lessons. I know who to expect to work more and who to expect to turn around all the time.” After ETG she says that all the pupils were behaving well and paying attention. The video data, however, suggest withdrawn and possibly off-task behavior. This might be caused by the “field of view” bias—the teacher did not see the off-task pupils in the ETG replay as much as when watching the ON video where they were clearly visible.

In situation T3_3 (situation T3_2 was omitted due to missing data; see **Figure 7**), the teacher checks answers to a textbook exercise. She writes up the sentences on the board herself. In the IN mode, her gaze is directed mostly to the board and the pupils she is interacting with. In the ON condition, she focuses on pupils who she interacted with (D and I) and on pupils who stand out—one who raises his hand often (C) and two who are displaying disruptive behavior (H and L). She comments on their disruptive behavior in the interview and notes that it ceased after they remembered they were being videotaped. It is interesting that pupil D is lying on her desk, after her interaction with the teacher is over, but the teacher no longer pays her attention in either of the conditions. During the interview, the teacher comments on most of the pupils. She comments on their behavior in the sequence in relation to their general characteristics, sometimes seen from a long-term and whole person perspective. With pupil D, she comments on the change she sees in her over the past more than 2 years,

including her situation at home in her considerations. In her comments, she revisits most of the pupils several times, giving very detailed accounts.

In situation T3_4 (see **Figure 7**), we see the beginning of the lesson. The teacher opens the lesson by asking her pupils how they feel that morning, addressing individuals directly (C, D, N, P, H). In the IN mode she follows the pupils she is speaking to with her gaze. In the ON mode, she either follows the same pupil as in the IN mode (e.g., N) or, when the pupil is not visible, scans the class, paying more attention to pupils who she usually reports to be disruptive (G, H, A) and who are inattentive and appear to be exchanging silent jokes. All pupils visible in the ON condition, however, receive short, repeated visits throughout the situation. During the interview, she first comments on pupils she had no interaction with, commenting on their engagement (A, I, L, C, D) or lack thereof (G, H—disruptive; E, F—off task but not disruptive). After watching the ETG replay, she also comments on pupils N and P, who she interacted with, and their neighbors (M, O), giving their overall characteristics and pair dynamics. The only person not mentioned in the interview is J.

Differences in Gaze

Based on the IN and ON action default differences discussed above, two types of relationship between the IN and ON modes can occur. There can be the same conditions (i.e., the teacher was in a similar position during teaching as the camera that recorded the pupils; visual field is similar) or different conditions (the ON mode does not depict the same visual field). The latter is typically present when the teacher looks at the board or at their own materials during teaching or when the pupil that the teacher focuses on in the IN mode is not visible or partially hidden by

another pupil in the IN mode. In each of our situations, both conditions are usually present, albeit in a different ratio.

Both of these sets of conditions are interesting in themselves. What characteristics of each condition cause greatest differences between gaze behavior in IN and ON modes? When given the opportunity to view the same visual field in the same situation, do teachers focus on the same? And when given the opportunity to see what was happening “behind their back,” what do they choose to zoom in on?

Correlations Between Dwell Times in IN and ON Conditions

The analysis of correlations between dwell time per pupil in IN and ON conditions (see **Table 4**) shows that the conditions causing the biggest difference (i.e., the weakest correlation) between gaze in IN versus ON mode are those where the pupils that the teacher interacted a lot with in the IN condition are missing from view in the ON condition (situations T3_1, T1_1, T2_2). Medium range of correlations can be observed for situations that were different because of the teacher looking at the board or at materials during IN condition, with some pupils with minor interactions in IN condition missing from view in ON condition (T2_4, T1_4, T3_3). Different conditions (e.g., teacher looking at the board) in certain parts of the situation do not seem to result in low correlations where teachers tend to consistently monitor the same pupils at the same time in both conditions when possible (T2_1, T2_3).

Instances With Different Conditions

With all three teachers there were instances where the conditions in the IN and ON modes differed—either the teacher looked at the board/into their textbook/at some other material in the IN mode or at a pupil who was not visible in the ON mode.

In instances where different conditions occur, three different sets of behaviors can be observed. In such circumstances, T1 tends to focus on pupils who are clearly visible in the ON mode video. After reviewing the video sequences and the corresponding interviews we did not find another indication as to why the teacher focused on the pupils she focused on other than their position relative to the camera. On the other hand, T3 seems to focus on pupils who are central to the video but not most prominent who display signs of inattentiveness and off-task behavior. She comments on this in each interview, labeling their behavior as typical of them. T2 combines the two types of behavior—she monitors the class in general with frequent short glances (often the pattern is not discernible) or she focuses on pupils of interest. Her comments in the interviews highlight this as she points out these pupils for raising their hand, looking bored or disrupting others, for example.

Instances With Similar Conditions

For all teachers there were times when the conditions in the IN and ON modes were similar. Situations with mostly similar conditions yield high correlation values (see **Table 3**). A look at the combined timelines (for an example see **Figure 4**) reveals

that often the teachers looked at the same pupil in both the conditions. The pupils looked at in the IN condition for longer periods of time are typically those the teacher interacts with (all situations pertain to whole class work where the teacher interacts with the class—see above). It is thus understandable that this interaction draws their attention again in the ON mode—it represents the core of the situation. However, not all teachers' visual attention was the same. T1 in many cases looked at the same pupils in both conditions and gave them the same or greater amount of attention (see **Figure 4** from 50,000 ms onward). On the other hand, T2 and sometimes T3 looked at the same pupils at the same (or similar) time, too, but often gave them less attention, just looking at them briefly and then continuing with monitoring the whole class or focusing on pupils of interest in the ON mode.

General Differences

Irrespective of the set of conditions (same or different in the IN and ON modes), several observations can be made. Overall monitoring behavior differs in the IN and ON modes. In the IN mode, most attention is usually focused on the pupil who the teacher interacts with (apart from the board, if applicable). Teachers T1 and T2 usually scan the class briefly before calling on a pupil. This is not so prevalent with T3. During the interaction with a pupil, their attention is focused mostly on them, with some glances to the pupils around (in the same area). This concentrated attention is prominent especially with pupils who struggle to answer (e.g., B in situation T1_4). T2 provides insight into the monitoring when teaching during her interviews. She says that her attention seems to be driven by pupils standing out (hands raised) in situation T2_1, where scanning is minimal. However, in situations with more scanning, she says she calls on pupils who either raise their hands or who she thinks will be able to answer, considering their competence. Here, we could surmise that the scanning of the class before calling on a pupil does not only help with decision-making (assessing who is attentive, on-task etc.), but also gives time to consider who can answer and who should be given an opportunity to answer correctly (e.g., easier questions for weaker pupils to provide them with opportunity to interact, experience success).

In the ON mode, irrespective of the conditions (same or different), monitoring behavior is different. All teachers in most situations monitor all (or most) visible pupils. This is evident in the amount of dwell time as represented by the graphs in Figures 5–7. Despite a clear focus noted above, the teachers manage to glance at all visible pupils and revisit them at least once (usually more times) during the situation, typically with some time in between glances. This suggests that despite the interaction often leading the attention in the ON mode, the teachers are free of the need to give the interacting pupil their undivided attention (as is usual in communication) and are available to monitor other pupils.

Individual Differences

The analysis showed that there are discernible patterns of monitoring behavior that set apart the IN and ON modes. It also became apparent that these patterns are not only mode

specific but also person specific. Our three participants each displayed a slightly different monitoring behavior under the same conditions, despite our sample being fairly homogenous.

T1 tended to follow the same pupils with her gaze in the IN and ON modes. If this was not possible, she tended to focus on most visible pupils—no other reason for focusing on them was found. T2 also focused on the same pupils in the ON mode that she interacted with in the IN mode but she tended to give them less attention and instead tended to monitor the whole class. Bigger chunks of attention, if present, were given to pupils of interest. Reasons for focusing on them were diverse—from active pupils, to struggling to disruptive. The monitoring behavior of T3 in the ON mode was characterized by focusing on pupils who displayed disruptive behavior. Here, the interviews provide an interesting insight. In some instances, after watching the ON mode video, the teacher would point out misbehaving or off-task behavior. For instance, after watching situation T3_1 she comments that “how the kids work here, so they work in lessons. I know who to expect to work more and who to expect to keep turning around [i.e. communicating with other pupils around them] all the time.” indicating different engagement from different pupils. However, when confronted with the ETG replay, she comments that “everyone was engaged” and the pupils “were being good” (T3_1). A similar thing happens after viewing situation T3_4. First, she highlights pupils who “did their own things,” but after ETG replay she concludes that everyone was paying attention. This suggests a “field of view” bias—in the ON condition, the teacher monitors more pupils, whereas in the IN condition mostly focuses on pupils who she interacts with. Thus during ETG replay, her field of view is again focused on these pupils. Even though several pupils are in the view (given the angle), attention is probably affected by the dot (gaze point) showing her gaze during teaching.

DISCUSSION AND CONCLUSION

The study set out to explore visual attention represented by gaze that underlies decision processes in the act of teaching (IN mode) and that is also the base for professional vision as a cognitive concept investigated after the event (ON mode). We focused on classroom situations where the whole class works together and the teacher interacts with individual pupils or groups of pupils.

The greatest differences between IN and ON mode gaze behavior (weakest correlations) are caused by pupils missing from view in the ON mode who were interacted with a lot during the IN mode. In general, two types of relationship between IN and ON mode field of view can be described—one with similar conditions (field of view is similar in both) and one with different conditions (field of view differs, because of the teacher looking at the board in the IN mode or some pupils not being visible in the ON mode). Clear patterns are discernible in both conditions. When similar conditions are present, the teachers tend to look at the same pupils, albeit to different degrees. It stands to reason as the pupils paid attention to are

mostly those who the teacher interacts with. This means they draw attention during the ON mode too. With different conditions, teachers differ in what seems to drive their attention—for some it is mainly the position of the pupil relative to the camera, for others it is the behavior of the pupil, be it them raising hands or displaying off-task behavior.

It is most noteworthy that monitoring behavior is different in the ON mode for the whole sample. More pupils are focused on—for most situations all pupils visible in the ON mode were focused on at least once, usually even more times, during each episode. Despite research suggesting that experienced teachers maintain higher awareness of class in general even when providing feedback to an individual pupil (Cortina et al., 2015), our sample in the selected episodes did not confirm this. Teachers were mostly focused on the pupil they interacted with, giving limited attention to pupils in the same area. This field of view bias was confirmed in interviews with T3 who seemed to be aware of disruptive pupils only when confronted with ON mode video and disregarded them when ETG replay was provided (mimicking IN action conditions). It is thus of great interest that this general monitoring behavior is introduced when teachers are provided with a different field of view and freed from the immediacy of the classroom and the interaction itself. We can hypothesize that the attention processes are less guided by bottom-up factors (such as motion or event focus; cf. comments of T2 on her selection of pupils to interact with) and more by top-down factors (such as conscious monitoring of all pupils; cf. Seidel et al., 2020) as more cognitive resources are made available by removing the pressure of the hot action (Eraut, 1994) and the high cognitive load of whole-class situations (Prieto et al., 2015), and by the removal of the interaction demands (e.g., maintaining eye contact). That is not to say that during the act of teaching top-down factors are not in play; our data might suggest that the ratio of top-down and bottom-up factors changes between IN and ON action.

From our data it also transpires that there are individual differences, despite general trends. Teacher 1 (T1) followed the same pupils at the same time, giving them a similar amount of attention. If this was not possible, her gaze seems to have been directed by the visibility of the pupils. On the other hand, Teacher 2 (T2) and Teacher 3 (T3) gave the same pupils less attention and focused more on pupils of interest. Individual differences can be expected as each teacher has a unique set of experiences and knowledge, despite the sample being fairly homogenous in terms of education, years of teaching, and even school culture (Stürmer et al., 2017).

The results presented contribute to our understanding of the concept of professional vision of teachers. Previous research on professional vision has mostly focused on professional vision after the act of teaching and was investigated through verbal reports (see overview in Professional vision outside the act of teaching). This is logical as studying teacher thinking, reasoning, and noticing during the act of teaching is extremely challenging. The fact, however, is often not reflected in the papers—many of them justify research on professional vision by the immediacy and simultaneity of classroom events. However, common sense tells us, and our results confirm this, that the monitoring behavior of teachers is different in IN and ON action conditions—the

angles are different, the choice of where to look is taken away to a certain extent, and, most importantly, the need to act, the need to actively communicate with the pupils and show interest are not present. This is not to say that ON action condition research on professional vision is not valuable and meaningful. Our results simply highlight the fact that the two conditions are different and result in different types of monitoring behavior. Or, in other words, we need to be cautious when generalizing our knowledge of monitoring behavior as observed in ON action studies to IN action conditions.

Further research into the issue, even though challenging, is much needed. As mentioned above, visual attention represented by gaze is only one facet of professional vision as defined in literature. Data describing visual attention (gaze) and teacher's reasoning can be collected and integrated for ON action condition (through interviews etc.) relatively easily, but further ways of investigating not only visual attention, but also reasoning during the IN action condition need to be sought, such as focusing on teacher's actions (e.g. Cortina et al., 2015) or rigorous video-stimulated recall.

This study also broadens our understanding of the use of video in teacher education. It is commonplace to include a kind of video work into both pre- and in-service teacher education programmes (Gaudin and Chaliès, 2015). Our results can inform teacher educators who use participants' own videos in their courses. We show that video recordings of classroom situations provide a different view of the classroom with more time to monitor the pupils. However, the example of T3 also shows that even this might not (at least in the short term) lead to desired effects. T3 was able to see the classroom situations more "objectively" when presented with the pupil camera replay but returned to her "field of view" bias and commented that pupils were on task when presented with the ETG camera replay, even though these two replays happened shortly after each other. This suggests that during the act of teaching she might be "seeing" a more favorable image of her class than when removed from the situation, which hinders her from activating appropriate scripts (Wolff et al., 2020).

It is important to remember, however, that there are limits to our study. The sample was rather limited due to the character of data collection and the demands on all participants stemming from the European data protection policy. With the advent of GDPR, the European Data Protection Regulation (EU) 2016/679, which is applicable as of May 25th, 2018 in all member states to harmonize data privacy laws across Europe, effective data collection in real classrooms has become hugely complicated. Before data gathering, a large number of consent and agreement forms signed by the teachers and all pupils' parents need to be obtained to meet the requirements of the GDPR policy. The aim of this note is not to criticize the established rules, which are highly needed and appropriate in many ways, but to express a

concern about the collection of eye-tracking data in schools, which is currently almost impossible due to the above-mentioned regulations. For homogeneity reasons the sample only comprised teachers from the same school and with similar education and length of experience. Participation of teachers and pupils from different schools would broaden the perspective but provide a less homogenous sample. Only one type of situation was included, namely such where the teacher works with the whole class. Future studies in different contexts are needed. Lastly, it is important to remember that any disruption to classrooms can have an effect on the data collected—be it the use of eye-tracking glasses, video cameras, or just the presence of researchers during lessons. We tried to mitigate these by getting both teachers and pupils familiar with the process and the equipment beforehand and by regularly checking with the teachers how much the lessons were influenced. Despite these limitations, the article provides an insight into a new area of research and can be an inspiration for further studies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Masaryk University Ethics Committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

EM—research design and supervision, data analysis; ZS—research design, data collection and analysis, MJ—data collection, KH—data analysis.

FUNDING

This work was supported by the project English teachers' professional vision in/on action in communicative activities from the perspective of eye tracking (GA17-15467S) funded by the Czech Science Foundation and by the research infrastructure HUME Lab Experimental Humanities Laboratory, Faculty of Arts, Masaryk University.

REFERENCES

- Ashcraft, M. H. (2006). *Cognition*. Upper Saddle River: Pearson Education.
- Blomberg, G., Renkl, A., Sherin, M. G., Borko, H., and Seidel, T. (2013). Five Research-Based Heuristics for Using Video in Pre-service Teacher Education. *J. Educ. Res. Online* 5, 90–114.
- Blomberg, G., Stürmer, K., and Seidel, T. (2011). How Pre-service Teachers Observe Teaching on Video: Effects of Viewers' Teaching Subjects and the Subject of the Video. *Teach. Teach. Edu.* 27, 1131–1140. doi:10.1016/j.tate.2011.04.008
- Choppin, J. (2011). Learned Adaptations: Teachers' Understanding and Use of Curriculum Resources. *J. Math. Teach. Educ* 14, 331–353. doi:10.1007/s10857-011-9170-3

- Copeland, W. D., Birmingham, C., DeMeulle, L., D'Emidio-Caston, M., and Natal, D. (1994). Making Meaning in Classrooms: an Investigation of Cognitive Processes in Aspiring Teachers, Experienced Teachers, and Their Peers. *Am. Educ. Res. J.* 31, 166–196. doi:10.3102/00028312031001166
- Cortina, K. S., Miller, K. F., McKenzie, R., and Epstein, A. (2015). Where Low and High Inference Data Converge: Validation of CLASS Assessment of Mathematics Instruction Using mobile Eye Tracking with Expert and Novice Teachers. *Int. J. Sci. Math. Educ.* 13, 389–403. doi:10.1007/s10763-014-9610-5
- Dessus, P., Cosnefroy, O., and Luengo, V. (2016). “Keep Your Eyes On'em All!": A Mobile Eye-Tracking Analysis of Teachers' Sensitivity to Students,” in European Conference on Technology Enhanced Learning, Lyon, France, September 13-16, 2016 (Springer), 72–84. doi:10.1007/978-3-319-45153-4
- Doyle, W. (1977). “Learning the Classroom Environment: An Ecological Analysis of Induction into Teaching,” in *Annual Meeting of American Educational Research Association* (New York. Available at <http://files.eric.ed.gov/fulltext/ED135782.pdf>).
- Duchowski, A. (2007). *Eye Tracking Methodology: Theory and Practice*. London: Springer Science & Business Media.
- Egi, H., Ozawa, S., and Mori, Y. (2014). “Analyses of Comparative Gaze with Eye-Tracking Technique for Peer-Reviewing Classrooms,” in Advanced Learning Technologies (ICALT), 2014 IEEE 14th International Conference, Athens, Greece, 7-10 July 2014 (IEEE), 622–623. doi:10.1109/ICALT.2014.181
- Endsley, M. R. (2015). Situation Awareness Misconceptions and Misunderstandings. *J. Cogn. Eng. Decis. Making* 9, 4–32. doi:10.1177/1555343415572631
- Eraut, M. (1994). *Developing Professional Knowledge and Competence*. London: The Falmer Press.
- Es, E. A., and Sherin, M. G. (2006). How Different Video Club Designs Support Teachers in “Learning to notice.” *J. Comput. Teach. Edu.* 22, 125–135.
- Es, E. A., and Sherin, M. G. (2002). Learning to Notice: Scaffolding New Teachers' Interpretations of Classroom Interactions. *J. Tech. Teach. Edu.* 10, 571–596.
- Gäfväls, C. (2016). Vision and Embodied Knowing: The Making of Floral Design. *Vocations Learn.* 9, 133–149. doi:10.1007/s12186-015-9143-2
- Gamoran Sherin, M., and van Es, E. A. (2009). Effects of Video Club Participation on Teachers' Professional Vision. *J. Teach. Edu.* 60, 20–37. doi:10.1177/0022487108328155
- Gaudin, C., and Chaliès, S. (2015). Video Viewing in Teacher Education and Professional Development: A Literature Review. *Educ. Res. Rev.* 16, 41–67. doi:10.1016/j.edurev.2015.06.001
- Gegenfurtner, A., Siewiorek, A., Lehtinen, E., and Säljö, R. (2013). Assessing the Quality of Expertise Differences in the Comprehension of Medical Visualizations. *Vocations Learn.* 6, 37–54. doi:10.1007/s12186-012-9088-7
- Gonzalez, L. E., and Carter, K. (1996). Correspondence in Cooperating Teachers' and Student Teachers' Interpretations of Classroom Events. *Teach. Teach. Edu.* 12, 39–47. doi:10.1016/0742-051X(95)00024-E
- Goodwin, C. (1994). Professional Vision. *Am. Anthropologist* 96, 606–633. doi:10.1525/aa.1994.96.3.02a00100
- Hayes, A. F., and Krippendorff, K. (2007). Answering the Call for a Standard Reliability Measure for Coding Data. *Commun. Methods Measures* 1, 77–89. doi:10.1080/19312450709336664
- Holmqvist, K., and Andersson, R. (2017). *Eye Tracking: A Comprehensive Guide to Methods, Paradigms and Measures*. Lund, Sweden: Lund Eye-Tracking Research Institute.
- Jacobs, V. R. (2017). “Complexities in Measuring Teacher Noticing: Commentary,” in *Teacher Noticing: Bridging and Broadening Perspectives, Contexts, and Frameworks*. Editors E. O. Schack, M. H. Fisher, and J. A. Wilhelm (Cham: Springer), 273–279. doi:10.1007/978-3-319-46753-5_16
- Janík, T., Minaříková, E., Pišová, M., Kostková, K., Janík, M., and Hublová, G. (2014). Profesní Vidění Učitelů: Pokus O Zmapování Výzkumného Pole. *Pedagogika* 64, 151–176.
- Johannes, C., and Seidel, T. (2012). Professionalisierung von Hochschullehrenden. *Z. Erziehungswiss* 15, 233–251. doi:10.1007/s11618-012-0273-0
- Kim, W.-J., Byeon, J.-H., Lee, I.-S., and Kwon, Y.-J. (2012). Repeat after Me. *Nat. Med.* 18, 1443. doi:10.1038/nm.2978
- McIntyre, N. A., and Foulsham, T. (2018). Scanpath Analysis of Expertise and Culture in Teacher Gaze in Real-World Classrooms. *Instr. Sci.* 46, 435–455. doi:10.1007/s11251-017-9445-x
- McIntyre, N. A., Jarodzka, H., and Klassen, R. M. (2019). Capturing Teacher Priorities: Using Real-World Eye-Tracking to Investigate Expert Teacher Priorities across Two Cultures. *Learn. Instruction* 60, 215–224. doi:10.1016/j.learninstruc.2017.12.003
- McIntyre, N. A., Mainhard, M. T., and Klassen, R. M. (2017). Are You Looking to Teach? Cultural, Temporal and Dynamic Insights into Expert Teacher Gaze. *Learn. Instruction* 49, 41–53. doi:10.1016/j.learninstruc.2016.12.005
- Meschede, N., Fiebranz, A., Möller, K., and Steffensky, M. (2017). Teachers' Professional Vision, Pedagogical Content Knowledge and Beliefs: On its Relation and Differences between Pre-service and In-Service Teachers. *Teach. Teach. Edu.* 66, 158–170. doi:10.1016/j.tate.2017.04.010
- M. G. Sherin, V. R. Jacobs, and R. A. Philipp (2011). in *Mathematics Teacher Noticing. Seeing through Teachers' Eyes* (New York: Routledge). doi:10.4324/9780203832714
- Minaříková, E., Pišová, M., Janík, T., and Uličná, K. (2015). Video Clubs: EFL Teachers' Selective Attention before and after. *Orbis scholae* 9, 55–75. doi:10.14712/23363177.2015.80
- Mustonen, V., Hakkarainen, K., Tuunainen, J., and Pohjola, P. (2015). Discrepancies in Expert Decision-Making in Forensic Fingerprint Examination. *Forensic Sci. Int.* 254, 215–226. doi:10.1016/j.forsciint.2015.07.031
- Nückles, M. (2020). Investigating Visual Perception in Teaching and Learning with Advanced Eye-Tracking Methodologies: Rewards and Challenges of an Innovative Research Paradigm. *Educ. Psychol. Rev.* 33, 149–167. doi:10.1007/s10648-020-09567-5
- Pouta, M., Lehtinen, E., and Palonen, T. (2020). Student Teachers' and Experienced Teachers' Professional Vision of Students' Understanding of the Rational Number Concept. *Educ. Psychol. Rev.* 33, 109–128. doi:10.1007/s10648-020-09536-y
- Prieto, L. P., Sharma, K., and Dillenbourg, P. (2015). “Studying Teacher Orchestration Load in Technology-Enhanced Classrooms,” in *Design for Teaching and Learning in a Networked World*. Editors G. Conole, T. Klobučar, J. K. C. Rensing, and E. Lavoué (Heidelberg: Springer International Publishing), 268–281. doi:10.1007/978-3-319-24258-3_20
- Prieto, L. P., Wen, Y., Caballero, D., Sharma, K., and Dillenbourg, P. (2014). “Studying Teacher Cognitive Load in Multi-Tabletop Classrooms Using Mobile Eye-Tracking,” in ITS '14: Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces, November 2014 (Heidelberg, Germany: ACM), 339–344. doi:10.1145/2669485.2669543
- Schön, D. A. (1983). *The Reflective Practitioner*. London: Basic Books.
- Seidel, T., Blomberg, G., and Stürmer, K. (2010). OBSERVE - Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. *Z. für Pädagogik* 56, 296–306.
- Seidel, T., Schnitzler, K., Kosel, C., Stürmer, K., and Holzberger, D. (2020). Student Characteristics in the Eyes of Teachers: Differences between Novice and Expert Teachers in Judgment Accuracy, Observed Behavioral Cues, and Gaze. *Educ. Psychol. Rev.* 33, 69–89. doi:10.1007/s10648-020-09532-2
- Sherin, M. G. (2001). “Developing a Professional Vision of Classroom Events,” in *Beyond Classical Pedagogy: Teaching Elementary School Mathematics*. Editors T. Wood, B. S. Nelson, and J. Warfield (Hillsdale: Erlbaum), 75–93.
- Sherin, M. G. (2007). “The Development of Teachers' Professional Vision in Video Clubs,” in *Video Research in the Learning Sciences*. Editors R. Goldman, R. Pea, B. Barron, and S. J. Derry (London: Lawrence Erlbaum Associates Publishers), 383–396.
- Sherin, M. G., and Han, S. Y. (2004). Teacher Learning in the Context of a Video Club. *Teach. Teach. Edu.* 20, 163–183. doi:10.1016/j.tate.2003.08.001
- Sherin, M. G., Russ, R. S., Sherin, B. L., and Colestock, A. (2008). Professional Vision in Action: An Exploratory Study. *Issues Teach. Edu.* 17, 27–46.
- Sherin, M. G. (2017). “Exploring the Boundaries of Teacher Noticing. Springer, 401–408. doi:10.1007/978-3-319-46753-5_23

- Smidekova, Z., Janik, M., Minarikova, E., and Holmqvist, K. (2020). Teachers' Gaze over Space and Time in a Real-World Classroom. *Jemr* 13. doi:10.16910/jemr.13.4.1
- Sonmez, D., and Hakverdi-Can, M. (2012). Videos as an Instructional Tool in Pre-service Science Teacher Education. *Egitim Arastirmalari-Eurasian J. Educ. Res.* 12, 141–158.
- Star, J. R., and Strickland, S. K. (2008). Learning to Observe: Using Video to Improve Preservice Mathematics Teachers' Ability to Notice. *J. Math. Teach. Educ.* 11, 107–125. doi:10.1007/s10857-007-9063-7
- Stürmer, K., Könings, K. D., and Seidel, T. (2013). Declarative Knowledge and Professional Vision in Teacher Education: Effect of Courses in Teaching and Learning. *Br. J. Educ. Psychol.* 83, 467–483. doi:10.1111/j.2044-8279.2012.02075.x
- Stürmer, K., Seidel, T., Müller, K., Häusler, J., and Cortina, K. S. (2017). What Is in the Eye of Preservice Teachers while Instructing? an Eye-Tracking Study about Attention Processes in Different Teaching Situations. *Z. für Erziehungswissenschaft* 20, 75–92. doi:10.1007/s11618-017-0731-910.1007/978-3-658-15739-5_4
- Styhre, A. (2010). Disciplining Professional Vision in Architectural Work. *Learn. Organ.* 17, 437–454. doi:10.1108/09696471011059822
- van den Bogert, N., van Bruggen, J., Kostons, D., and Jochems, W. (2014). First Steps into Understanding Teachers' Visual Perception of Classroom Events. *Teach. Teach. Edu.* 37, 208–216. doi:10.1016/j.tate.2013.09.001
- Vondrová, N., and Žalská, J. (2012). Do student Teachers Attend to Mathematics Specific Phenomena when Observing Mathematics Teaching on Video? *Orbis scholae* 6, 85–101. doi:10.14712/23363177.2015.42
- Wolff, C. E., Jarodzka, H., and Boshuizen, H. P. A. (2020). Classroom Management Scripts: a Theoretical Model Contrasting Expert and Novice Teachers' Knowledge and Awareness of Classroom Events. *Educ. Psychol. Rev.* 33, 131–148. doi:10.1007/s10648-020-09542-0
- Wolff, C. E., Jarodzka, H., van den Bogert, N., and Boshuizen, H. P. A. (2016). Teacher Vision: Expert and Novice Teachers' Perception of Problematic Classroom Management Scenes. *Instr. Sci.* 44, 243–265. doi:10.1007/s11251-016-9367-z
- Wyss, C., Rosenberger, K., and Bühner, W. (2020). Student Teachers' and Teacher Educators' Professional Vision: Findings from an Eye Tracking Study. *Educ. Psychol. Rev.* 33, 91–107. doi:10.1007/s10648-020-09535-z
- Yamamoto, T., and Imai-Matsumura, K. (2013). Teachers' Gaze and Awareness of Students' Behavior: Using an Eye Tracker. *Compr. Psychol.* 2, 01.IT.2.6. doi:10.2466/01.IT.2.6
- Yarbus, A. L. (1967). *Eye Movements and Vision*. New York: Plenum Press.
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
- Copyright © 2021 Minarikova, Smidekova, Janik and Holmqvist. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.