frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 24 October 2023
DOI 10.3389/feduc.2023.1266826

Assessment of noticing of classroom disruptions: a multi-methods approach

Antje Biermann[1]*, Roland Brünken[1], Doris Lewalter[2] and Ann-Sophie Grub[1]

[1]Department for Empirical Educational Research, Saarland University, Saarbrücken, Germany, [2]TUM School of Social Sciences and Technology, Technical University of Munich, Munich, Germany

Check for updates

OPEN ACCESS

EDITED BY
Kai S. Cortina,
University of Michigan, United States

REVIEWED BY
José Cravino,
University of Trás-os-Montes and Alto Douro, Portugal
Andreas Gegenfurtner,
University of Augsburg, Germany

*CORRESPONDENCE
Antje Biermann
a.biermann@mx.uni-saarland.de

RECEIVED 25 July 2023
ACCEPTED 11 October 2023
PUBLISHED 24 October 2023

CITATION
Biermann A, Brünken R, Lewalter D and Grub A-S (2023) Assessment of noticing of classroom disruptions: a multi-methods approach.
Front. Educ. 8:1266826.
doi: 10.3389/feduc.2023.1266826

Teachers' noticing as a basic precondition for effective teaching is characterized by focusing on relevant events in the classroom and ignoring the irrelevant. In recent years, many researchers have used eye-tracking methodology in classroom observations to gather information about the continuous attentional processes of teachers. Despite the general validity of the eye–mind assumption, methodological triangulation is necessary to draw conclusions about the where and why of the focus of attention. Although in previous studies, different data sources like gaze and verbal data have been used, the analyses were mostly conducted separately, instead of directly combining the data. In our study, we collected verbal data (retrospective think-aloud; RTA) and a reaction-based concurrent measure (keystroke) to assess the noticing process of novice and experienced teachers ($N = 52$) while they watched staged videos of classroom situations. For a direct triangulation, we combined these data with eye-tracking data. The aim of the study was to combine both measures with eye-tracking parameters that indicate attentional processes (fixation count, mean fixation duration, and revisits), and with expertise. We found that participants who were aware of the critical incidents in the videos (they gave a keystroke or mentioned the incident in the RTA), showed—as expected—a higher number of fixations and more revisits to the appropriate area, but a comparable mean fixation duration. However, expertise differences regarding accuracy in both measures could not be shown. We discuss methodological issues regarding the implementation of RTA and keystroke as measurements for the noticing process because—despite only partially significant results—both methods are promising as they allow complementation and possible correction of eye-movement-only data.

KEYWORDS

noticing, classroom management, triangulation, eye-tracking, stimulated retrospective think-aloud, keystroke

1. Introduction

Teaching as an interactive and complex profession needs spontaneous and flexible decisions to handle the requirements of individual situations (Doyle, 1985). In the classroom, features that need attention are, for example, potential disturbances (Grub et al., 2020), student hand-raising (Kosel et al., 2023a), or uninterested or struggling students (Seidel et al., 2021). Only when teachers are able to place a conscious focus on the mentioned aspects, they can react adequately (e.g., Grub et al., in press). Teachers' noticing—also often named as professional vision (PV)—as an important prerequisite for their professional behavior can be understood

10.3389/feduc.2023.1266826

as a situation-specific skill to perceive and interpret the demands of the situation (e.g., Blömeke et al., 2015). In particular, the basic perceptual process of attending to relevant elements in the classroom and ignoring the irrelevant is a widely investigated construct in research on teacher professionalization (König et al., 2022), although there are different conceptualizations and definitions of the terms (Stahnke et al., 2016; König et al., 2022). Despite the various definitions, it has been shown that expert teachers with an elaborated knowledge structure are better able to attend to the specific relevant features of a situation than novices (Lachner et al., 2016; Stahnke et al., 2016; Wolff et al., 2016, 2021; Boshuizen et al., 2020; Grub et al., 2022a). Due to technical development in recent years, a direct assessment of the noticing process is possible with process-based measurements such as eye-tracking and they are increasingly used in research of teachers' noticing (Grub et al., 2020, in press; Stahnke and Blömeke, 2021a).

A basis for eye-tracking research is the eye–mind assumption, which means that conclusions about cognitive processes can be drawn from fixations in a given moment (Just and Carpenter, 1980). This hypothesis is debatable, because it overlooks hidden attention or parafoveal perception, and it is unclear from eye-tracking data alone why a person fixates on a given stimulus (Posner, 1980; Anderson et al., 2004; Holmqvist et al., 2011; Jarodzka et al., 2017; Orquin and Holmqvist, 2018; Schindler and Lilienthal, 2019). This leads to the necessity of triangulation using additional measurement methods (Orquin and Holmqvist, 2018; Gegenfurtner et al., 2023; Grub et al., in press). For this purpose, in research on teachers' noticing, verbalization methods such as think-aloud protocols are often used (Wolff et al., 2016; Stahnke and Blömeke, 2021b; Grub et al., 2022a). However, in most cases, the different data sources are not cross-linked but rather are analyzed separately from each other. A direct link between eye-tracking measurements and verbal data for assessing attentional processes has been implemented only in a few studies (Yamamoto and Imai-Matsumura, 2013; Wyss et al., 2020).

A look toward other research fields, where focused attention and quick reaction are based on the perception of visual stimuli in dynamic scenarios, is worth considering to identify further methods that assess attentional processes. For example, in traffic psychology, a widespread method is the keystroke task to assess attentional processes in hazard perception scenarios (e.g., Malone and Brünken, 2020). Such reaction-based tasks are already used in research on teachers' noticing, but the generated timestamps are used to select material for the analyses (e.g., van den Bogert et al., 2014) or as cues for retrospective think-aloud protocols (RTA; Stahnke and Blömeke, 2021b), and not for analyzing attentional processes itself.

In the present study, we aimed to investigate where the participants had their particular attention during the observation process. For this, we use reaction-based measurement (keystroke) and stimulated RTA to measure the noticing process and validated the data with different parameters of gaze behavior as well as with expertise. More precisely, we aimed to investigate, on one hand, whether the participants who noticed critical incidents (CIs) in the scenarios differed in their attentional gaze behavior (fixation count, fixation duration, and revisits). On the other hand, we aimed to determine whether experts showed higher accuracy in noticing CIs with keystroke measurement and RTA when compared to novices.

1.1. Teachers' noticing

It is indisputable that the perception of essential cues within a profession is a key component of expertise, especially for teachers (Blömeke et al., 2015; Stahnke et al., 2016). Nevertheless, different authors use different conceptualizations and names for the construct of the perceptual processes (Stahnke et al., 2016; König et al., 2022). The concept of Professional Vision was originally introduced as a holistic concept from a sociocultural perspective by Goodwin (1994). It describes a specialized way of seeing and understanding meaningful events in a specific professional context. The adaptation of the concept in the context of teaching entailed a shift to a cognitive psychological perspective and was first used by Sherin (2001) (see also Sherin and van Es, 2009; Seidel and Stürmer, 2014) with the intertwined subprocesses of selective attention (noticing) and knowledge-based reasoning (Weyers et al., 2023b). In later publications, Sherin and van Es subsumed the perception and reasoning process under the term noticing (van Es and Sherin, 2002; Sherin et al., 2011). Other authors added the decision-making process (Jacobs et al., 2010; Kaiser et al., 2015) or rather the concept of shaping (van Es and Sherin, 2021) to noticing; while some concepts only subsumed the perceptual process under this term (Star and Strickland, 2008; Seidel et al., 2021). The common feature of these analytic approaches (König et al., 2022) is the possibility of separate consideration of the subprocesses. An empirical example is the study from Seidel and Stürmer (2014), who referred on the term professional vision and the subprocesses of reasoning. The authors could show in their study that a three-factor model with the components description, interpretation, and prediction fits better than a one-factor model (see also the study of Weyers et al., 2023a). We follow the analytical approach and focus with the term noticing only on the attentional focus of teachers (see the narrower definition of Star and Strickland, 2008 and Seidel et al., 2021).

In all previously mentioned conceptualizations, expertise, and the accompanying professional knowledge, is a crucial basis for noticing (Blömeke et al., 2015; Stahnke et al., 2016; Grub et al., 2020, 2022a; Wolff et al., 2021). Even the perceptual focus on relevant cues presupposes knowledge of the relevance (or irrelevance) of signals for teaching and learning. The direction of attentional processes can be based on professional knowledge (top-down), but also on salient cues in the situation itself (bottom-up; see; for example; Wolff et al., 2021; Gegenfurtner et al., 2023). From expertise research, it is well known that experts with a professional knowledge base, that has well-structured and organized schemata, are better able to direct their attention to relevant cues than novices (Lachner et al., 2016; Stahnke et al., 2016; Wolff et al., 2016, 2021; Boshuizen et al., 2020; Grub et al., 2022a).

1.2. Measuring the noticing process

1.2.1. Eye-tracking

In the last nearly 15 years, however, process-based methodologies such as eye-tracking have been increasingly used in research on professional vision of teachers, and thus, have been able to provide further insight into the noticing process (Grub et al., 2020, in press). For example, eye-movement data reflect attention and shifts in attention (van Gog et al., 2009) and can be used to assess and analyze teachers' allocation of attention in the classroom (Bucher and Schumacher, 2012; Beach and McConnel, 2018; Haataja et al., 2019; Grub et al., in press). According to the eye–mind assumption (Just and Carpenter, 1980), the cognitive processing of an object corresponds to the visual focus or rather the fixation on the object. A fixation is defined as "a period of time when the eye is relatively still" (oculomotor definition; Holmqvist et al., 2011, p. 377). Therefore, gaze-based indicators for attention, number of fixations, and fixation duration can be used (Holmqvist et al., 2011). Furthermore, analyzing revisits (areas of interests—AOIs—where people look back) allows conclusions to be made about the relevance of AOIs for participants (Wolff et al., 2016; Grub et al., 2020). Many studies in the context of professional vision in teaching lay their primary focus on contrasting novice and expert teachers (Wolff et al., 2017; Seidel et al., 2021; Shinoda et al., 2021; Stahnke and Blömeke, 2021a,b; Grub et al., 2022b), but not on the attentional focus itself. From these studies, it is known that expert teachers place a stronger focus on relevant aspects of the classroom (e.g., disruptions or hand raising-behavior). Yamamoto and Imai-Matsumura (2013) and Wyss et al. (2020) both directly focused on attentional processes, showing that participants who were aware of a CI fixated on it more often and with a longer duration (see Section 1.2.2 for further description of the studies).

Nevertheless, inferences about the cognitive processes or location of an individual's attention made from eye-tracking measures alone can be imprecise and skewed because of parafoveal processes (Anderson et al., 2004) as well as hidden attention (Posner, 1980). Therefore, it is strongly recommended that gaze data should be complemented with other measurements to identify where attention is actually being directed (methodological triangulation; Denzin, 1989; Holmqvist et al., 2011; Orquin and Holmqvist, 2018; Grub et al., in press).

1.2.2. Post hoc verbalization and standardized tests

For assessment of the noticing process—that is, of *what* was seen—standardized instruments (for an overview see Weyers et al., 2023b), as well as RTA protocols (Wolff et al., 2015; Stahnke and Blömeke, 2021b; Grub et al., 2022a), are worth considering. Standardized test instruments are used in most cases after the actual stimulus materials, in the form of videos or pictures of teaching scenarios in combination with open- or closed-ended questions. Most of these tests differentiate between the dimensions of professional vision or noticing (attending/perception, reasoning, and decision-making; Weyers et al., 2023b), but do not represent the spontaneity of classroom situations (Stahnke and Blömeke, 2021a).

Likewise, the think-aloud method (Ericsson, 2018) in the context of professional vision uses videos or pictures as stimulus materials (Wolff et al., 2015; Stahnke and Blömeke, 2021b). The verbalizations of participants will give insight into their cognitive processes as they handle a task (e.g., the observation of a teaching situation). Although the concurrent collection of verbal data during the task is a more accurate and valid procedure (Ericsson, 2018), it is not suitable for complex tasks such as teaching, or for use while eye-tracking, because the think-aloud process can distort gaze behavior and the cognitive processes themself (van Gog et al., 2005; Prokop et al., 2020). Therefore, the collection of verbal data takes place after the observation of a scenario and is often supported with videos (with or without integrated eye-movements), which are used as cues to increase the validity of the verbal data and avoid oblivion or fabrication problems (van Gog et al., 2005; Prokop et al., 2020). With these stimulated RTA protocols, researchers have investigated attentional processes in teaching (with instructions such as "Could you tell me, what you have seen?"; Wyss et al., 2020; Grub et al., 2022a), as well as reasoning processes, whereby participants are prompted to report their thoughts about what they have seen and the importance of the stimuli (Wolff et al., 2015, 2016, 2017; Gegenfurtner et al., 2020a; McIntyre et al., 2021; Grub et al., 2022a).

Combining verbal methods with eye-tracking as measurement methods of the noticing process is a promising approach (Wolff et al., 2016; Beach and McConnel, 2018), and previous studies of teachers' professional vision have used multiple data sources such as eye-tracking and (stimulated) RTA protocols (Wolff et al., 2016; Pouta et al., 2020; Stahnke and Blömeke, 2021b; Grub et al., 2022b). Nevertheless, in these studies, the data sources were analyzed separately from each other rather than being directly combined. In contrast, a direct combination of eye-tracking data and RTA was used by Muhonen et al. (2021, 2023) to investigate reasoning processes, as well as by Wyss et al. (2020) and Yamamoto and Imai-Matsumura (2013) to investigate attentional processes. Wyss et al. (2020) combined eye-tracking data and RTA for one short video clip to examine the association between the awareness of a CI and the gaze behavior of the participants (student teachers as novices and university teacher educators as experts). From the RTA protocols, they identified who mentioned the CIs and who did not and analyzed the gaze behavior of the two groups. First, they showed that expertise was connected to the identification of the CI (only experts were aware of them), and second, they showed that participants who were aware of the CI evidenced more and longer fixations on the incident. Nevertheless, the study had some drawbacks regarding generalizability: only six of the 56 participants were aware of the CI, and the authors used only one short video clip with one incident. Also using a comparable design with one short video, Yamamoto and Imai-Matsumura (2013) also found differences regarding fixation count (aware participants had more fixations) and fixation duration (a longer duration for aware participants). Since the cultural context of this latter study was situated in Japan, in East Asia, the transferability of the results to the Western context is potentially problematic.

1.2.3. Concurrent measures

In the field of traffic psychology, reaction-based measurement during the observation of traffic scenarios is a common method for investigating the skill of hazard perception (the ability to detect and evaluate road hazards quickly; Horswill and McKenna, 2004; Malone and Brünken, 2020). As the task is for the observers to press a key when they identify a relevant cue within the scenario, an inference to the corresponding cognitive attention process is possible (Chapman et al., 2002). This reaction-based method is a valid measurement in hazard perception research (Malone and Brünken, 2016; Moran et al., 2019): Experienced drivers react faster to CIs, and they identify more potential hazards correctly; this also applies to drivers with less accident liability (Malone and Brünken, 2020). Keystroke tasks have also been applied in research into the PV of teachers, but they are used as stimulus segments for RTA (e.g., Stahnke and Blömeke, 2021b) or for analyzing a selection of video segments in depth (van den Bogert et al., 2014), rather than for analyzing the focus of the teachers' attention. Like eye-tracking, the attentional process can directly be covered with a keystroke during the task but it should be combined with the fixations from eye-tracking to identify the location of the visual attention (Crundall, 2016; Malone and Brünken, 2020).

1.3. Aims and hypotheses

In the present study, first, we aimed to investigate whether keystroke measurements and the RTA are appropriate ways of assessing the noticing process (awareness of a CI). The first step is methodological triangulation with the eye-tracking parameters fixation count, fixation duration, and revisits, which corresponds to the focus of attention and the relevance of the area of interest (AOI) for the participants. Second, we aimed to validate the measures using the expertise of our participants.

For the first research question, we wanted to identify whether participants who were aware of a relevant CI in the videos (assessed with keystroke as well as with RTA) showed differences in their gaze behavior. According to the eye–mind hypothesis, the focus of attention can be inferred from the location of fixations (Just and Carpenter, 1980; Holmqvist et al., 2011). Wyss et al. (2020) and Yamamoto and Imai-Matsumura (2013) both showed that participants who were aware of a CI had more fixations on the CI and a longer fixation duration. Wolff et al. (2016) indicated that participants had more revisits to AOIs of higher relevance, which referred to a conscious noticing of the AOI.

> *Hypothesis 1a:* Participants who identify the CI correctly with a keystroke (they press a key and have a fixation on the CI within the relevant time period) will have more fixations, a longer fixation duration, and more revisits.
> *Hypothesis 1b:* Participants who mention the CI in the stimulated RTA protocols will have more fixations, a longer fixation duration, and more revisits.

For the second research question, we wanted to examine whether expert teachers (experts) were more aware of the CI than prospective teachers (novices). Experts in the field of teaching are better able to direct their attention to relevant cues; that is, they have more fixations on these AOIs than novices (Lachner et al., 2016; Stahnke et al., 2016; Wolff et al., 2016, 2021; Boshuizen et al., 2020).

> *Hypothesis 2a.* Experts will identify more CIs correctly with a keystroke (they press a key and have a fixation on the CI within the relevant time stamp) as novices.
> *Hypothesis 2b.* Experts will identify more CIs in the stimulated RTA protocols than novices.

2. Materials and methods

2.1. Participants

A total of $N = 71$ student and experienced teachers participated in the study. Teachers ($n = 37$) were recruited via email, telephone, and newspaper from schools in South-Western Germany and had at least 5 years of teaching experience. Student teachers ($n = 34$) were recruited via email lists and flyers at Saarland University and had a maximum of 40 h of teaching experience during their regular practice. Data from 12 participants (six experts, six novices) were excluded from the analyses due to insufficient data quality (e.g., an outlier with more than three standard deviations for at least one of the relevant variables). In addition, only those participants for whom both gaze and verbal data were available were used in the analyses. Thus, 26 student teachers and 26 experienced teachers remained ($N = 52$); see Table 1 for the details of the sample.

2.2. Design and procedure

We re-analyzed data from an already published study (Grub et al., 2022b) for the purposes of the present study in support of the intensive use of scientifically generated empirical data (Machado, 2015; see Müller and Gold, 2023; Kosel et al., 2023a, for examples). Overall, the participants observed seven short videos of staged lessons in mathematics and informatics. For the re-analysis presented here, we focused on only two of the videos with, in total, five CIs regarding classroom management (see Section 2.3 for more details).

The study consisted of three sequential parts (see Figure 1). In *Part 1*, participants answered a knowledge test and demographic questions, which were presented online on Unipark. *Part 2* of the experiment, an eye-tracking experiment in the laboratory, took place 10 days later, on average. Here, participants were quasi-randomly assigned to one of six video sequences, in which the videos were balanced regarding the order of presentation by Latin square to exclude sequence effects (the first of the seven videos was always the same and was provided as a tutorial video). The participants' eye movements were recorded as they observed the videos. Meanwhile, the task was to identify CIs in each video via a keystroke. In *Part 3*, a stimulated RTA based on individual eye-tracking data was performed directly after watching all the videos, in which the eye movements presented as scan paths served as a cue for verbalization. During this stage, the noticing process was assessed by asking participants to explicitly specify the previously identified events in more detail. In total, the experiment lasted around two hours (*Part 1*: around 45 min, *Part 2*: around 20 min, *Part 3*: around 55 min).

2.3. Apparatus and videos

Eye movements were recorded under standardized environmental conditions using a stationary, binocular eye tracker (Tobii Pro

TABLE 1 Demographic information.

	Novices (n = 26)	Experts (n = 26)
Age (in years)[a]	23.81 (6.02)	43.15 (9.54)
Gender	♀ 88.5%	♀ 46.2%
Teaching experience[a]	5.00 (9.60) hours	13.85 (7.94) years
Semester[a]	2.35 (2.64)	/

[a]*M(SD).*

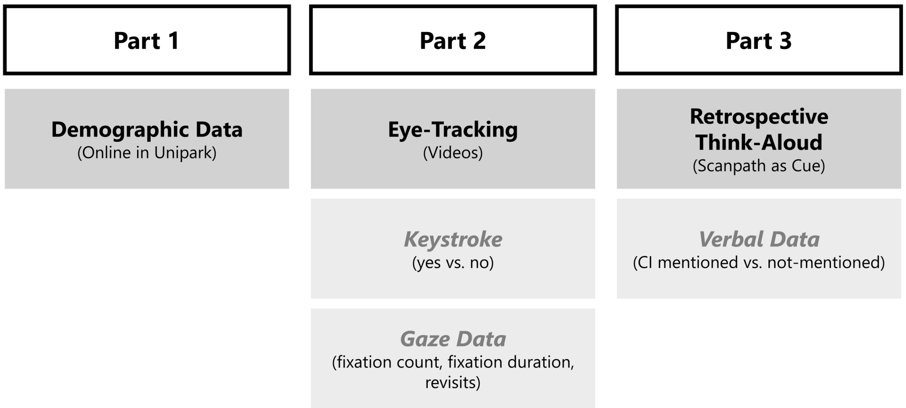

FIGURE 1
Design of the study.

TABLE 2 Specification of the disruptions in video A and video B.

Video	Type of disruption	Location	Accuracy keystroke *M (SD)*	Accuracy RTA *M (SD)*
A (duration: 1:29 min)	1. Striking yawning boy talking with neighbors	Back right	0.58 (0.50)	0.49 (0.50)
	2. Girl throws paper ball to another girl who is answering a teacher's question	Front	0.40 (0.50)	0.47 (0.50)
B (duration: 2:01 min)	3. Chattering pupil group	Back left	0.79 (0.41)	0.41 (0.50)
	4. Boys passing a slip of paper	Left	0.67 (0.47)	0.39 (0.49)
	5. Another chattering pupil group	Back right	0.73 (0.45)	0.33 (0.47)

Range of accuracy: 0–1.

Fusion, 120 Hz).[1] High-quality eye-tracking data were recorded for the participants (calibration accuracy: $M = 0.55$, $SD = 0.18$; calibration precision: $M = 0.35$, $SD = 0.18$).

The videos presented in the eye-tracking experiment were developed by "Toolbox Teacher Education" ("*Toolbox Lehrerbildung*") from the Technical University of Munich (Lewalter et al., 2020). They were based on scripted lessons from the 10th and 11th grades in the advanced track at a German secondary school ("*Gymnasium*"), covered topics in mathematics and informatics, and had already been used in preceding studies (Grub et al., 2022a,b). Each video was presented once during eye-tracking. The videos, in general, were selected based on events related to classroom management, audiovisual quality, and authenticity of the situation by three independent raters. For the aim of the present study, we selected two videos for a deeper analysis. The criteria for the selection of the videos were that they should consist of at least two CIs to increase the complexity of the scenario (Wyss et al., 2020), and approximately half of the participants should be aware of the CIs to give enough power for the analyses. The selected videos showed two or three CIs of classroom disruptions (for details, see Table 2).

2.4. Dependent variables and data analysis

2.4.1. Gaze data

An AOI-based evaluation of the eye-tracking data (fixation count, fixation duration, and revisits)[2] was performed; that is, the parameters were aggregated for the predefined AOIs. For this purpose, polygonal dynamic AOIs were determined deductively, including those that corresponded to the CI, that is, individual students or groups of students. The parameters were calculated for every CI in both video vignettes; therefore, each parameter was present for each CI.

2.4.2. Verbal data (RTA)

Using three general questions, participants were instructed to think aloud while watching their scan path-cued videos again ("What did you see?" "Why did you notice it?" and "How is what you saw relevant to the lesson?"); however, only the answer to the first question

1 We used a 24-inch display monitor (1080×1920), kept the distance between eye tracker and participants as identical as possible (approx. 65cm), and ensured uniform illumination. Before the recording itself, we conducted a 9-point automatic calibration followed by a validation to ensure data quality. The calibration was performed again if the 9-point automatic calibration failed.

2 The data were exported from Tobii with a Tobii I-VT (fixation) filter with a standard setting (I-VT classifier), i.e., a threshold from 30°/s.

10.3389/feduc.2023.1266826

was relevant for the present study. Participants could either verbalize a CI (1) or not mention it (0) in the RTA. Therefore, audio tracks from the RTA were transcribed verbatim. All the data were then read into MAXQDA 2022 and coded with respect to the descriptions of the CIs in the videos based on the master ratings (see Table 2). Coding was performed for each CI by two independent raters (the first author of this paper and an educational psychology student), with a satisfactory interrater reliability (Cohen's kappa) between 0.73 and 0.88.

2.4.3. Concurrent measure (keystroke)

Participants were instructed to watch the videos carefully and received the following instruction regarding the keypress: "If you notice something relevant, press the keyboard." Each CI was checked to identify whether a keystroke was recorded for it during the period in which the CI occurred. This keystroke defined whether the participant was aware of the CI and whether they had at least one fixation on that AOI. Therefore, each participant had either a keystroke score of 0 (not pressed) or 1 (pressed + fixation) for every CI.

3. Results

The analyses were calculated using SPSS IBM (version 29). An alpha level of 0.05 was used for the statistical tests. The analyses were conducted for each CI separately because the size and location of the AOIs could affect the values of the gaze parameters. Therefore, an aggregated analysis can lead to confounded results (Pappa et al., 2020; Holmqvist et al., 2022; Grub et al., in press).

3.1. Preliminary analyses and descriptive results

Because some previous studies have found effects of expertise on gaze behavior (see Grub et al., 2020, in press), we wanted to avoid possible confounding effects in our analyses. Therefore, we conducted MANOVAs for each CI with expertise as the between-subject variable and the eye-tracking parameters (fixation count, mean fixation duration, and revisits) as dependent variables. The means, standard deviations, and detailed results are displayed in Table 3. No statistically significant differences were found, so we removed expertise as a control variable in the subsequent analyses.

The descriptive results and the consistency regarding the accuracy of noticing the CIs using the two measurements (keystroke and RTA) are displayed above in Table 2. In Video A, around half of the participants reacted with a keystroke to the CI and mentioned it in the RTA. The consistency between the accuracy of the two measurements in Video A was moderate with Cohen's $\kappa = 0.46$ for CI 1 and $\kappa = 0.58$ for CI 2. In Video B, we found larger differences between accuracy by keystroke and accuracy by RTA: around 33–75% of the participants reacted with a keystroke, but only 33 to 41% mentioned the CI in the RTA, which led to low consistency values ($\kappa_{CI\,3} = -0.11$, $\kappa_{CI4} = 0.40$, and $\kappa_{CI5} = 0.20$).

3.2. Differences in gaze behavior regarding awareness

3.2.1. Concurrent measurement of noticing (keystroke)

A MANOVA was performed for each CI separately, with the concurrent measurement of noticing (keystroke yes vs. no) as the between-subject variable and the eye-tracking parameters (fixation count, mean fixation duration, and revisits) as dependent variables. The means, standard deviations, and detailed results can be found in Table 4. The omnibus test was significant for four CIs; therefore, the *post hoc* univariate ANOVAs were interpreted. Regarding *fixation count*, generally, participants who responded with the keystroke fixated on them more often. The difference was statistically significant for three CIs (CI 1, CI 2, and CI 4): $F_{CI\,1}$ (1,50) = 5.31, $p = 0.01$, $\eta_p^2 = 0.096$; $F_{CI\,2}$ (1,50) = 32.81, $p < .001$, $\eta_p^2 = 0.401$; $F_{CI\,4}$ (1,50) = 11.82, $p = 0.001$, $\eta_p^2 = 0.188$. Regarding *mean fixation duration*, participants who noticed the CI fixated on it longer only in the case of CI 3: $F_{CI\,3}$ (1,50) = 4.60, $p = 0.02$, $\eta_p^2 = 0.085$. Regarding *revisits*, aware participants looked at the AOIs more often in CI 3 and CI 4: $F_{CI\,3}$ (1,50) = 4.62, $p = 0.02$, $\eta_p^2 = 0.085$; $F_{CI\,4}$ (1,50) = 4.61, $p = 0.02$, $\eta_p^2 = 0.084$.

3.2.2. Verbal measurement of noticing (RTA)

As before, a MANOVA was performed for each CI separately, with the retrospective measurement of noticing (CI mentioned in RTA yes

TABLE 3 Expertise and gaze behavior.

	Descriptive statistics *M* (*SD*)						Values of significance	
	Fixation count		Mean fixation duration		Revisits		*F ratio* (*p*)	η_p^2
Expertise	Expert	Novice	Expert	Novice	Expert	Novice		
Video A								
CI 1	22.58 (9.51)	18.54 (10.09)	0.41 (0.19)	0.42 (0.09)	7.27 (3.13)	5.73 (3.57)	0.90 (0.45)	0.05
CI 2[a]	7.44 (4.57)	9.50 (5.16)	0.45 (0.24)	0.48 (0.25)	1.76 (1.33)	2.11 (1.33)	1.05 (0.38)	0.06
Video B								
CI 3	43.27 (15.76)	39.04 (14.69)	0.33 (0.10)	0.32 (0.13)	8.70 (3.47)	7.96 (3.93)	0.61 (0.62)	0.04
CI 4	40.46 (15.68)	33.50 (18.51)	0.28 (0.08)	0.25 (0.08)	11.84 (3.54)	9.54 (3.83)	2.42 (0.08)	0.13
CI 5	39.54 (12.17)	38.85 (13.27)	0.43 (0.14)	0.41 (0.16)	11.04 (3.26)	10.77 (3.63)	0.13 (0.94)	0.01

[a] $N = 51$. An alpha error correction according to Bonferroni Holm was performed for each CI.

TABLE 4 Keystroke and gaze behavior.

	Descriptive statistics *M* (*SD*)						Values of significance	
	Fixation count		Mean fixation duration		Revisits		*F* ratio (*p*)*	η_p^2
Key-stroke	Yes	No	Yes	No	Yes	No		
Video A								
CI 1	**23.17 (9.77)**	**17.00 (9.19)**	0.43 (0.12)	0.40 (0.18)	7.07 (3.05)	5.73 (3.79)	**2.44 (0.04)**	**0.13**
CI 2[a]	**12.19 (4.01)**	**5.90 (3.75)**	0.42 (0.14)	0.50 (0.29)	2.00 (1.30)	1.90 (1.37)	**11.31 (<0.001)**	**0.42**
Video B								
CI 3	42.59 (13.62)	35.82 (19.97)	**0.34 (0.12)**	**0.26 (0.09)**	**8.88 (3.57)**	**6.27 (3.55)**	**6.17 (<0.001)**	**0.28**
CI 4	**42.17 (15.94)**	**26.29 (15.45)**	0.28 (0.09)	0.25 (0.07)	**11.60 (3.57)**	**8.82 (3.80)**	**4.83 (0.003)**	**0.23**
CI 5	41.50 (12.64)	32.93 (10.54)	0.41 (0.15)	0.43 (0.15)	11.50 (3.17)	9.29 (3.65)	2.07 (0.06)	0.11

[a]$N = 51$; *One-sided significance test. Significant values ($\alpha < 0.05$) are given in bold. An alpha error correction according to Bonferroni Holm was performed for each CI.

TABLE 5 RTA and gaze behavior.

	Descriptive statistics						Values of significance	
	Fixation count		Mean fixation duration		Revisits		*F* ratio (*p*)*	η_p^2
RTA	Yes	No	Yes	No	Yes	No		
Video A								
CI 1	22.96 (10.44)	18.58 (9.14)	0.41 (0.08)	0.41 (0.18)	6.92 (3.24)	6.15 (3.24)	1.16 (0.167)	0.069
CI 2[a]	**11.79 (4.91)**	**5.54 (2.61)**	0.42 (0.17)	0.51 (0.25)	1.96 (1.08)	1.96 (1.56)	**11.62 (<0.001)**	**0.432**
Video B								
CI 3	44.86 (14.82)	39.00 (15.32)	0.28 (0.08)	0.35 (0.13)	9.29 (4.36)	7.73 (3.11)	2.15 (0.05)	0.121
CI 4	**49.05 (15.31)**	**29.94 (14.03)**	0.27 (0.09)	0.27 (0.08)	**12.55 (2.70)**	**9.71 (3.95)**	**6.86 (<0.001)**	**0.304**
CI 5	**44.65 (13.61)**	**36.47 (11.53)**	**0.35 (0.09)**	**0.45 (0.16)**	**12.12 (4.00)**	**10.26 (3.01)**	**2.50 (0.04)**	**0.138**

[a]$N = 51$; *One-sided significance test; RTA: CI mentioned in the stimulated RTA protocol. Significant values ($\alpha < 0.05$) are given in bold. An alpha error correction according to Bonferroni Holm was performed for each CI.

vs. no) as the between-subject variable and the eye-tracking parameters (fixation count, mean fixation duration, and revisits) as dependent variables. The means, standard deviations, and detailed results can be found in Table 5.

The omnibus test was significant for three CIs. Therefore, the *post hoc* univariate ANOVAs were interpreted. Regarding *fixation count*, generally, participants who mentioned the CI in the RTA, fixated on them more often. The difference was statistically significant for three CIs (C1 2, CI 4, and CI 5): $F_{CI\,2}$ (1,48) = 32.34, $p < 0.01$, $\eta_p^2 = 0.403$; $F_{CI\,4}$ (1,49) = 22.01, $p < 0.01$, $\eta_p^2 = 0.300$; $F_{CI\,5}$ (1,49) = 5.05, $p = 0.02$, $\eta_p^2 = 0.093$. Regarding *mean fixation duration*, participants who noticed the CI fixated on it longer only in one case (CI 5): $F_{CI\,5}$ (1,49) = 5.10, $p = 0.02$, $\eta_p^2 = 0.094$. Regarding *revisits*, aware participants looked at the AOIs more often and statistically significant differences were found for two CIs (CI 4 and CI 5): $F_{CI\,4}$ (1,49) = 7.91, $p = 0.003$, $\eta_p^2 = 0.139$; $F_{CI\,5}$ (1,49) = 3.44; $p = 0.02$; $\eta_p^2 = 0.066$.

3.3. Differences in noticing regarding expertise

To validate the measurement keystroke and RTA for the assessment of noticing, we investigated whether experts responded to more CIs with a keystroke (Hypothesis 2a) and mentioned it more frequently in the RTA (Hypothesis 2b). A MANOVA was performed for each CI separately, with expertise (expert teachers vs. student teachers) as the between-subject variable and the two CI noticing variables (keystroke, RTA) as dependent variables. The means, standard deviations, and detailed results can be found in Table 6. No statistically significant differences were found for noticing the CIs, for keystroke or RTA. This means that experts and novices noticed the CIs equally.

4. Discussion

In the present study, we used a keystroke task and a stimulated RTA task to assess the noticing process of (novice) teachers and combine the two tasks with eye-tracking measurement (fixation count, fixation duration, and revisits as known parameters for awareness) as well as with expertise through methodological triangulation.

Regarding our first research question, we showed that aware participants who pressed the key (and had a fixation on the corresponding AOI) fixated on the AOI more often and had more revisits than non-aware participants. Comparable results could be found for the verbal task: participants who mentioned the CI in the

TABLE 6 Noticing and expertise.

	Descriptive statistics *M* (*SD*)				Values of significance	
	Keystroke		RTA		*F ratio (p)**	η_p^2
Expertise	Expert	Novice	Expert	Novice		
Video A						
CI 1	0.62 (0.50)	0.52 (0.51)	0.58 (0.50)	0.40 (0.50)	0.78 (0.23)	0.032
CI 2	0.38 (0.50)	0.44 (0.51)	0.50 (0.51)	0.44 (0.51)	0.38 (0.34)	0.016
Video B						
CI 3	0.85 (0.37)	0.72 (0.46)	0.50 (0.51)	0.32 (0.48)	1.52 (0.11)	0.061
CI 4	0.77 (0.43)	0.60 (0.50)	0.46 (0.51)	0.32 (0.47)	0.95 (0.20)	0.038
CI 5	0.81 (0.40)	0.64 (0.49)	0.35 (0.49)	0.32 (0.48)	0.89 (0.21)	0.036

*One-sided significance test. An alpha error correction according to Bonferroni Holm was performed for each CI.

RTA fixated on the AOI more and had also more revisits. Nevertheless, for both gaze parameters, this was statistically ensured only for three of the five CIs. A statistically significant difference in mean fixation duration was found for only one CI. Contrary to our hypothesis, the mean fixation duration for CI 5 was shorter for participants who mentioned the CI in the RTA. Therefore, Hypotheses 1a and 1b were only partially confirmed.

The findings that participants who were aware of the CI (reacted with a keystroke or mentioned it in the RTA) showed gaze behavior that corresponded to attention to a CI (more fixations, more revisits), is aligned with theoretical assumptions (Gegenfurtner et al., 2011; Holmqvist et al., 2011) and empirical evidence on the PV of teachers (Yamamoto and Imai-Matsumura, 2013; Wolff et al., 2016; Wyss et al., 2020). The partial lack of statistical findings can possibly be explained by the fine-grained analysis for each separate CI (Kaakinen, 2020). We decided against an aggregated analysis because the size and location of the AOIs can influence gaze parameters (Pappa et al., 2020; Holmqvist et al., 2022; Grub et al., in press) and therefore confound the results. Further research should take this into account.

Contrary to our hypotheses, there were no mean fixation duration differences between the aware and unaware participants. The studies of Wyss et al. (2020) and Yamamoto and Imai-Matsumura (2013), which showed longer fixation durations for the aware participants, had only one CI in their videos. In contrast, our videos had two or three CIs, so we cannot completely rule out an overlap. Especially in dynamic fields like teaching in a classroom, continuous monitoring is necessary to update what is going on. Therefore, participants may have used more scanning or monitoring gaze behavior, which is characterized by more and shorter fixations (Wolff et al., 2016; Grub et al., 2020; Huang et al., 2021), to detect everything that could be important. That participants show more monitoring gaze behavior, might be also an effect of the presentation mode: Minarikova et al. (2021), for example, in their study, compared the IN-mode (during class instruction with a mobile eye tracker) and the ON-mode (observation of the video of that instruction with a stationary eye tracker) and showed more monitoring gaze-behavior in the ON-mode. During the instruction (IN-mode), the teacher must monitor the classroom, but, for example, they must also interact with the students, whereas in the ON-mode, the task is mostly observation. To interpret the results of our study, the processes of monitoring the whole scene and focusing on single cues could have canceled each other out, what resulted in non-significant differences. This leads also to the question, whether fixation duration is an appropriate parameter for assessing the attentional focus because of its sensitivity for influencing factors like presentation mode or rather the task of the observation (see also 4.2).

Referring to our second research question, we found no differences between experts and novices in the accuracy of noticing, either in the keystroke or the RTA tasks. In many studies, experts have been better able to identify relevant cues of a dynamic scene (Lachner et al., 2016; Stahnke et al., 2016; Wolff et al., 2016, 2021; Boshuizen et al., 2020). Nevertheless, the expertise effect tends to occur in more complex conditions, when elaborated knowledge is needed. Stahnke and Blömeke (2021a), for example, showed expertise differences in perceiving potential disturbances only for partner work scenarios, and not for whole group scenes. Seidel et al. (2021) demonstrated an advantage in detecting the characteristics of students for expert teachers only for inconsistent student profiles; the authors also showed expertise differences only in the seatwork scene. These results suggest that the recognition of potentially disruptive situations may depend on the knowledge or expertise of the observer but also that the context matters. It may be easier to recognize relevant disruptive situations in whole group scenes—as were used here—than in other formats. This could be because salient visual impressions, such as movements, are easier to recognize via bottom-up perception processes, and the advantage of experts (namely, that they already have elaborate, flexible knowledge schemata that enable top-down perception) is not necessary for identifying disruptive situations. In addition, it may be easier for student teachers to recognize situations that they have probably actively experienced themselves as students (disruptive teaching situations) than to identify situations that represent, for example, cognitive activation, since this requires much more elaborated knowledge of educational science and didactics.

The heterogeneous results lead to questions concerning validity and the benefit of using keystroke and RTA to complement the measurement of noticing in the methodological triangulation of data. Despite only partially significant results regarding the eye-tracking parameters, we see the keystroke task as a promising addition to gaze-data in dynamic scenarios. By a keystroke, the researcher can infer that a participant has a conscious focus on a specific AOI at a specific moment, especially when the keystroke is combined with a fixation on that AOI (Crundall, 2016; Malone and Brünken, 2020). Nevertheless, we propose some important conditions for the validity of this inference: (1) The AOIs should be definable and not overlapping. If this is not possible, additional information should be gathered about the thoughts or visual focus of the participant (see also Section 4.2). (2) A given instruction about the task should be well considered due to its influence on the gaze behavior (Yarbus, 1967; DeAngelus and Pelz, 2009; Grub et al., 2022b; Martin et al., 2023) but also on reaction-based measurements such as the keystroke. In our task, the participants had the instruction, "If you *notice* something relevant, press the keyboard," which ensured the focus of attention in any given moment of an AOI, but instructions to press the button when the participant identifies that *a reaction is needed* (Moran et al., 2019) would also be useful. Further research on this is necessary and important.

The verbal data gained through RTA can give important insight into the thought processes of the participants and are, therefore, a useful additional source of information about the noticing process (Yamamoto and Imai-Matsumura, 2013; Wyss et al., 2020). Although, we used the replay of participants' gaze as a stimulus for the RTA to obtain higher validity data, a drawback of our design was the administration of the RTA after the observation of all the videos. We could not prevent oblivion or mixing up the videos by participants. Indeed, in the RTA protocols, we found, for example, several statements such as "Cannot remember why I pressed the button" or about the behavior of students visible in another video. In addition, an active reconstruction process by means of cues instead of reporting the thoughts during observation from memory cannot be avoided (*cf.* van Gog et al., 2005).

Additionally, in some RTA protocols, the verbal allocation to a specific student group was not clear enough to rate it as "mentioned" (for example, with statements such as "they are chattering," where it was unclear which students were being referred to in the use of "they"). This is also a possible explanation for the lack of consistency between RTA and the keystroke in Video B. To prevent these effects, the timespan between observing videos and gaining verbal data should be as short as possible. Furthermore, instructions in the RTA should be very clear (for example, "Name or describe the students you focused on"; see Yamamoto and Imai-Matsumura, 2013). Alternatively, the relevant part of what is being verbalized could be clicked on again during the RTA, either via touchscreen monitor or mouse click (depending on the technical requirements). In follow-up studies, it may also be worth considering a modification of the classic RTA toward a somewhat more guided semi-structured interview with the investigator, so that they could make follow-up inquiries about ambiguous statements to reduce later ambiguities.

The finding that awareness based on RTA verbalizations was generally lower than awareness via keystroke (see Table 2) could be due to the method of recording, namely the verbalization itself. Novice teachers, for example, have difficulties applying their knowledge and reasoning about noticed events (Schäfer and Seidel, 2015). Thus, it could be that they had already noticed the events (which can be seen in the keystroke) but they interpreted and verbalized the situation differently compared to our master rating (for more information, see Grub et al., 2022a,b). The experts, on the other hand, might have had difficulties in verbalizing what they actually saw because their knowledge was often rather implicit and so-called tacit, and therefore, was not available for introspection (Sternberg and Horvath, 1999). Tacit knowledge is typically reflected in eye movements "but is not necessarily available for conscious thoughts" (Kaakinen, 2020, p. 172), and therefore is not verbalizable. These are conceivable reasons why the CIs were assigned awareness more often via keystroke than RTA. Further research is therefore necessary to derive, for example, causal relationships.

4.1. Strengths and limitations

Our research has some clear strengths, but also some weaknesses. We used a very well-balanced experimental design with a comparatively large sample size. What is positive in comparison to the studies of Wyss et al. (2020) and Yamamoto and Imai-Matsumura (2013) is that our data collection had a significantly higher sampling rate (120 vs. 60 Hz) and a more accurate calibration procedure (nine-point vs. five-point calibration). Furthermore, we used videos with multiple CIs (two or three CIs vs. one CI in the mentioned studies) and two videos instead of one, which makes the data more accurate and multifaceted and thus slightly increases the external validity.

Another strength is the multi-methods approach, which allowed us to combine the different measurement methods for noticing and examine the validity of the keystroke and verbal data with gaze parameters, a so-called methodological triangulation with three different sources of information. In addition, by systematically analyzing *post hoc* verbalizations (RTA) and concurrent measures (keystroke) in conjunction with eye-tracking data (fixation count, mean fixation duration, revisits), this article represents one of the first methodological attempts in the research area of teachers' PV to combine and triangulate different data sources in order to mitigate the problems associated with the eye–mind assumption and eye-tracking research. Thus, this study lays a foundation for further systematic investigations based on it and represents a starting point for continuing studies.

Nevertheless, on top of the mentioned limitation regarding administrating the RTA protocols, the study has some drawbacks. Although we used standardized video samples, these were very short extracts of very standardized lessons in a frontal teaching setting with low complexity and very "usual" disturbances (Grub et al., 2020, 2022a,b; Grub, 2023). To recognize differences in competence, the situation or the focus to be observed should be more complex, because only then will the top-down-based perception typical for expertise become relevant, and thus, the differences between novice and expert teachers will become visible (Seidel et al., 2021; Stahnke and Blömeke, 2021a). If explicit investigations regarding differences in expertise are to be conducted, care should be taken to ensure that the material used is also suitable for this purpose, i.e., that a knowledge-based top-down perception is a prerequisite for successful PV.

4.2. Implications for further research

With the keystroke and RTA approaches, we were able to investigate two additional measurement methods for noticing, so providing more information about the location of attention than gaze parameters alone. Both approaches have their particular (dis-) advantages. A keystroke is an economic and non-invasive process-based method of assessing the spontaneous noticing process during an observation. Researchers should bear in mind, however, that with a keystroke alone, only information about the timepoint, but not about the location of the noticing process, can be inferred. For this, additional information from eye-tracking (fixation at the moment of the keystroke) or verbalizing is necessary. Another possibility for future research is a mouseover with a click to tag the attentional focus during observation; however, this can lead to distortion of the gaze behavior (Malone and Brünken, 2020). Therefore, the question remains open about what (creative) possibilities there are for matching eye-movement data relative to the two investigated variants, and whether there might be more suitable methods for triangulation.

The RTA is a less economical way of assessing what participants have seen during an observation and can lead to fabrication or oblivion despite a careful research design. Furthermore, it has higher requirements regarding verbalization skills for the participants. Nevertheless, RTA is an appropriate method to gain insight into the knowledge schemata as well as the characteristics of a situation as a basis for making inferences about professionals. In our study, we only rated a defined disturbing behavior if it was mentioned in the RTA. Additionally, it might be interesting to collect information on which concrete characteristics lead to the inference that a specific behavior is relevant. Seidel et al. (2021), for example, investigated whether (student) teachers can assess student profiles (struggling, uninterested, underestimating, etc.) accurately, asking participants which features lead to their decision. In our view, this is promising for gaining insight into the beliefs or intentions of (prospective) teachers. For example, the valuation of a behavior as disturbing depends on one's normative beliefs and expectations, the rules of the situation (e.g., frontal instruction vs. seatwork; Stahnke and Blömeke, 2021a), or the didactical aims of a teaching situation (McIntyre et al., 2017; Muhonen et al., 2021, 2023). Beyond that, however, many other conceivable factors can influence the relevance of an event and a participant's decisions and behavior; for example, what education the (student) teacher has enjoyed, generational effects, more conservative vs. liberal attitudes facing the teaching process (which could influence the estimation process), experience in dealing with disturbances, motivation, or (self-effective) handling of disturbances in one's own teaching practice.

Finally, we want to emphasize a well elaborated selection of eye-tracking parameters carefully derived from a theoretical foundation and the hypotheses of the research (Holmqvist et al., 2011; Carter and Luke, 2020; Grub, 2023). For example, the parameter fixation duration could be influenced by some conditions as the aim of perception (e.g., observing or interacting) or the presentation mode of the stimuli (e.g., ON mode vs. IN mode). As a parameter for attention, it is maybe not sufficient for dynamic tasks, when monitoring or scanning is needed (as observation in the classroom or a traffic scene to detect incidents or hazards). For assessing monitoring gaze behavior, the Gini coefficient as a measure of (un-) equal distribution (Cortina et al., 2015) as well as the gaze relational index (GRI) as a combined measure was used (Gegenfurtner et al., 2020b; Kosel et al., 2023b). The latter is calculated as the relation of fixation count and fixation duration, a lower value indicate a more equally distribution of gaze over the scene. It is an open question, if other combinations of parameters are suitable for assessing the focus of attention or other relevant perceptional skills. For future research, we also see the application of a person-centered approach as promising to combine different eye-tracking parameters as well as other methodological measures within a person to identify different profiles and take the heterogeneity of persons or situations into consideration (Bergman and Andersson, 2012; Hickendorff et al., 2018).

5. Conclusion

With the present paper, we create a first step toward the systematic elaboration of a methodological approach for triangulating eye-tracking data from (prospective) teachers by providing new insight into triangulating eye-movement data with concurrent measurement as well as *post hoc* verbalization. We were able to find heterogeneous results regarding the validity of the two triangulation methods: when considering eye movements and measures of noticing (RTA, keystroke), it is clear that conscious awareness is associated with increased monitoring of video footage (in particular, more fixations and revisits), but not for all CIs. In accordance with these findings, verbal data and concurrent measurements seem to be useful, albeit limited, ways to link eye movements to awareness. Furthermore, we did not find differences in attention between expert and student teachers, either in terms of verbal data or their behavioral responses. All in all, this type of research needs further systematic investigation and purposeful manipulation to ensure a more suitable resource-saving method of data triangulation for eye-tracking data that can be used in the medium term to examine teachers' PV. Only adequate methodological triangulation will allow valid conclusions to be drawn from eye-tracking data, and ultimately, they will be of great importance, especially in the long term, for applying study results to teacher education and training.

With regard to open science research, we would like to emphasize that our study is presented as transparently as possible: the study design was preregistered, our data are accessible for interested researchers, and regarding the eye-tracking procedure, as much information as possible is provided to make replication possible (e.g., information on calibration precision/accuracy, threshold filter, information on the generation of AOIs, etc. see Pappa et al., 2020; Holmqvist et al., 2022; Kosel et al., 2023b).

Data availability statement

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

Ethics statement

The studies involving humans were approved by Data Protection Officer, Saarland University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

 10.3389/feduc.2023.1266826

Author contributions

AB: Conceptualization, Data curation, Formal Analysis, Writing – original draft. RB: Writing – review & editing, Funding acquisition. DL: Resources, Writing – review & editing. A-SG: Writing – review & editing, Conceptualization, Data curation, Investigation, Methodology, Project administration, Resources.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

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.

References

Anderson, J. R., Bothell, D., and Douglass, S. (2004). Eye movements do not reflect retrieval processes. Limits of the eye-mind hypothesis. *Psychol. Sci.* 15, 225–231. doi: 10.1111/j.0956-7976.2004.00656.x

Beach, P., and McConnel, J. (2018). Eye tracking methodology for studying learning: a review of research. *Int. J. Res. Method Educ.* 42, 485–501. doi: 10.1080/1743727X.2018.1496415

Bergman, L. R., and Andersson, H. (2012). The person and the variable in developmental psychology. *Z. Psychol.* 21, 155–165. doi: 10.1027/0044-3409/a000025

Blömeke, S., Gustafsson, J.-E., and Shavelson, R. J. (2015). Beyond dichotomies. *Z. Psychol.* 223, 3–13. doi: 10.1027/2151-2604/a000194

Boshuizen, H. P. A., Gruber, H., and Strasser, J. (2020). Knowledge restructuring through case processing: the key to generalise expertise development theory across domains? *Educ. Res.* 29:100310. doi: 10.1016/j.edurev.2020.100310

Bucher, H.-J., and Schumacher, P. (2012). "Aufmerksamkeit und Informationsselektion: Blickdaten als Schlüssel zur Aufmerksamkeitssteuerung [Gaze data as a key to attentional control]" in *Interaktionale Rezeptionsforschung: Theorie und Methode der Blickaufzeichnung in der Medienforschung [Interactional Reception Research: Theory and Methods of Eye Tracking in Media Research]*. eds. H.-J. Bucher and P. Schumacher (Wiesbaden: Springer Fachmedien), 83–107.

Carter, B. T., and Luke, S. G. (2020). Best practices in eye tracking research. *Int. J. Psychophysiol.* 155, 49–62. doi: 10.1016/j.ijpsycho.2020.05.010

Chapman, P., Underwood, G., and Roberts, K. (2002). Visual search patterns in trained and untrained novice drivers. *Transport. Res. F: Traffic Psychol. Behav.* 5, 157–167. doi: 10.1016/S1369-8478(02)00014-1

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

Crundall, D. (2016). Hazard prediction discriminates between novice and experienced drivers. *Accid. Anal. Prev.* 86, 47–58. doi: 10.1016/j.aap.2015.10.006

DeAngelus, M., and Pelz, J. B. (2009). Top-down control of eye movements: Yarbus revisited. *Vis. Cogn.* 17, 790–811. doi: 10.1080/13506280902793843

Denzin, N. K. (1989). *The Research Act: A Theoretical Introduction to Sociological Methods*. NJ: Englewood Cliffs

Doyle, W. (1985). Recent research on classroom management: implications for teacher preparation. *J. Teach. Educ.* 36, 31–35. doi: 10.1177/002248718503600307

Ericsson, K. A. (2018). "Protocol analysis and expert thought: concurrent verbalizations of thinking during experts' performance on representative tasks" in *The Cambridge Handbook of Expertise and Expert Performance*. eds. K. A. Ericsson, A. Hoffman, A. Kozbelt and A. M. Williams (Cambridge: University Press), 223–242.

Gegenfurtner, A., Boucheix, J.-M., Gruber, H., Lehtinen, E., and Lowe, R. K. (2020a). The gaze relational index as a measure of visual expertise. *Anim. Proc. Model* 3, 32–40.

Gegenfurtner, A., Gruber, H., Holzberger, D., Keskin, Ö., Lehtinen, E., Seidel, T., et al. (2023). "Towards a cognitive theory of visual expertise: methods of inquiry" in *Re-Theorising Learning and Research Methods in Learning Research*. eds. C. Damşa, A. Rajala, G. Ritella and J. Brouwer (London: Routlegde), 146–163.

Gegenfurtner, A., Lehtinen, E., and Säljö, R. (2011). Expertise differences in the comprehension of visualizations: a meta-analysis of eye-tracking research in professional domains. *Educ. Psychol. Rev.* 23, 523–552. doi: 10.1007/s10648-011-9174-7

Gegenfurtner, A., Lewalter, D., Lehtinen, E., Schmidt, M., and Gruber, H. (2020b). Teacher expertise and professional vision: examining knowledge-based reasoning of pre-service teachers, in-service teachers, and school principals. *Front. Educ.* 5:59. doi: 10.3389/feduc.2020.00059

Goodwin, C. (1994). Professional vision. *Am. Anthropol.* 96, 606–633. doi: 10.1525/aa.1994.96.3.02a00100

Grub, A.-S. (2023). PRONOEA—professional vision of novice and expert teachers. Doctoral Dissertation. Saarland University, Saarbrücken

Grub, A.-S., Biermann, A., and Brünken, R. (2020). Process-based measurement of professional vision of (prospective) teachers in the field of classroom management: a systematic review. *J. Educ. Res. Online* 12, 75–102. doi: 10.25656/01:21187

Grub, A.-S., Biermann, A., and Brünken, R. (in press). "Eye tracking as a process-based methodology to examine teachers' professional vision" in *Teacher Professional Vision: Theoretical and Methodological Advances (Part 1). EARLI-book series by Routledge "New Perspectives on Learning and Instruction"*. eds. A. Gegenfurtner and R. Stahnke (London: Routlegde).

Grub, A.-S., Biermann, A., Lewalter, D., and Brünken, R. (2022a). Professional knowledge and task instruction specificity as influencing factors of prospective teachers' professional vision. *Teach. Teach. Educ.* 109:103517. doi: 10.1016/j.tate.2021.103517

Grub, A.-S., Biermann, A., Lewalter, D., and Brünken, R. (2022b). Professional vision and the compensatory effect of a minimal instructional intervention: a quasi-experimental eye-tracking study with novice and expert teachers. *Front. Educ.* 7:890690. doi: 10.3389/feduc.2022.890690

Haataja, E., Garcia Moreno-Esteva, E., Salonen, V., Laine, A., Toivanen, M., and Hannula, M. S. (2019). Teacher's visual attention when scaffolding collaborative mathematical problem solving. *Teach. Teach. Educ.* 86:102877. doi: 10.1016/j.tate.2019.102877

Hickendorff, M., Edelsbrunner, P. A., McMullen, J., Schneider, M., and Trezise, K. (2018). Informative tools for characterizing individual differences in learning: latent class, latent profile, and latent transition analysis. *Learn. Individ. Differ.* 66, 4–15. doi: 10.1016/j.lindif.2017.11.001

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R. J. H., and Van de Weijer, J. (2011). *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford: OUP.

Holmqvist, K., Orbom, S. L., Hooge, I. T. C., Niehorster, D. C., Alexander, R. G., Andersson, R., et al. (2022). Eye tracking: empirical foundations for a minimal reporting guideline. *Behav. Res. Methods* 55, 364–416. doi: 10.3758/s13428-021-01762-8

Horswill, M. S., and McKenna, F. P. (2004). "Drivers' hazard perception ability: situation awareness on the road" in *A Cognitive Approach to Situation Awareness*. eds. S. Banbury and S. Tremblay (Aldershot: Ashgate), 155–175.

Huang, Y., Miller, K. F., Cortina, K. S., and Richter, D. (2021). Teachers' professional vision in action. *Z. für Pädagog. Psychol.* 37, 122–139. doi: 10.1024/1010-0652/a000313

Jacobs, V. R., Lamb, L. C., and Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *J. Res. Math. Educ.* 41, 169–202. doi: 10.5951/jresematheduc.41.2.0169

Jarodzka, H., Holmqvist, K., and Gruber, H. (2017). Eye tracking in educational science: theoretical frameworks and research agendas. *J. Eye Mov. Res.* 10, 1–18. doi: 10.16910/jemr.10.1.3

Just, M. A., and Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychol. Rev.* 87, 329–354. doi: 10.1037/0033-295X.87.4.329

Kaakinen, J. K. (2020). What can eye movements tell us about visual perception processes in classroom contexts? Commentary on a special issue. *Educ. Psychol. Rev.* 33, 169–179. doi: 10.1007/s10648-020-09573-7

Kaiser, G., Busse, A., Hoth, J., König, J., and Blömeke, S. (2015). About the complexities of video-based assessments: theoretical and methodological approaches to overcoming shortcomings of research on teachers' competence. *Int. J. Sci. Math. Educ.* 13, 369–387. doi: 10.1007/s10763-015-9616-7

König, J., Santagata, R., Scheiner, T., Adleff, A.-K., Yang, X., and Kaiser, G. (2022). Teacher noticing: a systematic literature review of conceptualizations, research designs, and findings on learning to notice. *Educ. Res. Rev.* 36:100453. doi: 10.1016/j.edurev.2022.100453

Kosel, C., Böheim, R., Schnitzler, K., Holzberger, D., Pfeffer, J., Bannert, M., et al. (2023a). Keeping track in classroom discourse: comparing in-service and pre-service teachers' visual attention to students' hand-raising behavior. *Teach. Teach. Educ.* 128:104142. doi: 10.1016/j.tate.2023.104142

Kosel, C., Mooseder, A., Seidl, T., and Pfeffer, J. (2023b). Measuring Teachers' visual expertise using the gaze relational index based on real-world eye-tracking data and varying velocity thresholds. arXiv [Preprint] doi: 10.48550/arXiv.2304.05143

Lachner, A., Jarodzka, H., and Nückles, M. (2016). What makes an expert teacher? Investigating teachers' professional vision and discourse abilities. *Instr. Sci.* 44, 197–203. doi: 10.1007/s11251-016-9376-y

Lewalter, D., Titze, S., Bannert, M., and Richter-Gebert, J. (2020). Lehrer*innenbildung digital und disziplinverbindend. *J. Für LehrerInnenbildung jlb* 20, 76–84. doi: 10.35468/jlb-02-2020_06

Machado, J. (2015). "Open data and open science" in *Open Science, Open Issues*. eds. S. Albagli, M. L. Maciel and A. H. Abdo (Unirio: IBICT), 189–214.

Malone, S., and Brünken, R. (2016). The role of ecological validity in hazard perception assessment. *Transport. Res. F: Traffic Psychol. Behav.* 40, 91–103. doi: 10.1016/j.trf.2016.04.008

Malone, S., and Brünken, R. (2020). Studying gaze behavior to compare three different hazard perception tasks. *Hum. Factors* 62, 1286–1303. doi: 10.1177/0018720819873462

Martin, M., Farrell, M., Seidel, T., Rieß, W., Könings, K. D., van Merriënboer, J. J., et al. (2023). Knowing what matters: short introductory texts support pre-service teachers 'professional vision of tutoring interactions. *Teach. Teach. Educ.* 124:104014. doi: 10.1016/j.tate.2023.104014

McIntyre, N. A., Draycott, B., and Wolff, C. E. (2021). Keeping track of expert teachers: comparing the affordances of think-aloud elicited by two different video perspectives. *Learn. Instr.* 80:101563. doi: 10.1016/j.learninstruc.2021.101563

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. Instr.* 49, 41–53. doi: 10.1016/j.learninstruc.2016.12.005

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

Moran, C., Bennett, J. M., and Prabhakharan, P. (2019). Road user hazard perception tests: a systematic review of current methodologies. *Accid. Anal. Prev.* 129, 309–333. doi: 10.1016/j.aap.2019.05.021

Muhonen, H., Pakarinen, E., and Lerkkanen, M.-K. (2021). Do teachers' professional vision and teaching experience always go hand in hand? Examining knowledge-based reasoning of Finnish grade 1 teachers. *Teach. Teach. Educ.* 106:103458. doi: 10.1016/j.tate.2021.103458

Muhonen, H., Pakarinen, E., and Lerkkanen, M.-K. (2023). Professional vision in the classroom: teachers' knowledge-based reasoning explaining their visual focus of attention to students. *Teach. Teach. Educ.* 121:103907. doi: 10.1016/j.tate.2022.103907

Müller, M. M., and Gold, B. (2023). Videobasierte Erfassung wissensbasierten Verarbeitens als Teilprozess der professionellen Unterrichtswahrnehmung—Analyse eines geschlossenen und offenen Verfahrens. *Z. Erzieh.* 26, 7–29. doi: 10.1007/s11618-022-01128-6

Orquin, J. L., and Holmqvist, K. (2018). Threats to the validity of eye-movement research in psychology. *Behav. Res. Methods* 50, 1645–1656. doi: 10.3758/s13428-017-0998-z

Pappa, C. I., Kosel, C., Schnitzler, K., and Seidel, T. (2020). Using eye-tracking to investigate teacher gaze: data accuracy and drawing of meaningful dynamic areas of interest in video stimuli. *Pedagog. Orient.* 29, 401–423. doi: 10.5817/PedOr2019-4-401

Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25. doi: 10.1080/00335558008248231

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

Prokop, M., Pilař, L., and Tichá, I. (2020). Impact of think-aloud on eye-tracking: a comparison of concurrent and retrospective think-aloud for research on decision-making in the game environment. *Sensors* 20, 739–771. doi: 10.3390/s20102750

Schäfer, S., and Seidel, T. (2015). Noticing and reasoning of teaching and learning components by pre-service teachers. Journal for educational research. *Online* 7:11489. doi: 10.25656/01

Schindler, M., and 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. *Educ. Stud. Math.* 101, 123–139. doi: 10.1007/s10649-019-9878-z

Seidel, T., Schnitzler, K., Kosel, C., Stürmer, K., and Holzberger, D. (2021). 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

Seidel, T., and Stürmer, K. (2014). Modeling and measuring the structure of professional vision in preservice teachers. *Am. Educ. Res. J.* 51, 739–771. doi: 10.3102/0002831214531321

Sherin, M. G. (2001). "Developing a professional vision of classroom events" in *Beyond Classical Pedagogy: Teaching Elementary School Mathematics*. eds. T. L. Wood, B. S. Nelson and J. Warfield (Mahwah, NJ: Lawrence Erlbaum Associates), 75–93.

Sherin, M. G., Jacobs, V. R., and Philipp, R. A. (2011). "Situating the study of teacher noticing" in *Mathematics Teacher Noticing: Seeing Through Teachers' Eyes*. eds. M. G. Sherin, V. R. Jacobs and R. A. Philipp (London: Routledge), 3–13.

Sherin, M. G., and van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *J. Teach. Educ.* 60, 20–37. doi: 10.1177/0022487108328155

Shinoda, H., Yamamoto, T., and Imai-Matsumura, K. (2021). Teachers' visual processing of children's off-task behaviors in class: a comparison between teachers and student teachers. *PLoS One* 16:e0259410. doi: 10.1371/journal.pone.0259410

Stahnke, R., and Blömeke, S. (2021a). Novice and expert teachers' noticing of classroom management in whole-group and partner work activities: evidence from teachers' gaze and identification of events. *Learn. Instr.* 74:101464. doi: 10.1016/j.learninstruc.2021.101464

Stahnke, R., and Blömeke, S. (2021b). Novice and expert teachers' situation-specific skills regarding classroom management: what do they perceive, interpret and suggest? *Teach. Teach. Educ.* 98:103243. doi: 10.1016/j.tate.2020.103243

Stahnke, R., Schueler, S., and Roesken-Winter, B. (2016). Teachers' perception, interpretation, and decision-making: a systematic review of empirical mathematics education research. *ZDM* 48, 1–27. doi: 10.1007/s11858-016-0775-y

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

Sternberg, R. J., and Horvath, J. A. (Eds.) (1999). *Tacit Knowledge in Professional Practice: Researcher and Practitioner Perspectives*. Hove, UK: Psychology Press.

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. Educ.* 37, 208–216. doi: 10.1016/j.tate.2013.09.001

van Es, E. A., and Sherin, M. G. (2002). Learning to notice: scaffolding new teachers' interpretations of classroom interactions. *J. Technol. Teach. Educ.* 10, 571–596.

van Es, E. A., and Sherin, M. G. (2021). Expanding on prior conceptualizations of teacher noticing. ZDM. *Math. Educ.* 53, 17–27. doi: 10.1007/s11858-020-01211-4

van Gog, T., Jarodzka, H., Scheiter, K., and Paas, F. (2009). Attention guidance during example study via the model's eye movements. *Comput. Hum. Behav.* 25, 785–791. doi: 10.1016/j.chb.2009.02.007

van Gog, T., Paas, F., van Merriënboer, J. J. G., and Witte, P. (2005). Uncovering the problem-solving process: cued retrospective reporting versus concurrent and retrospective reporting. *J. Exp. Psychol. Appl.* 11, 237–244. doi: 10.1037/1076-898X.11.4.237

Weyers, J., König, J., Rott, B., Greefrath, G., Vorhölter, K., and Kaiser, G. (2023a). Mathematics teachers' professional noticing: transfer of a video-based competence assessment instrument into teacher education for evaluation purposes. *Z. Erzieh.* 26, 627–652. doi: 10.1007/s11618-023-01159-7

Weyers, J., König, J., Santagata, R., Scheiner, T., and Kaiser, G. (2023b). Measuring teacher noticing: a scoping review of standardized instruments. *Teach. Teach. Educ.* 122:103970. doi: 10.1016/j.tate.2022.103970

Wolff, C. E., Jarodzka, H., and Boshuizen, H. P. A. (2017). See and tell: differences between expert and novice teachers' interpretations of problematic classroom management events. *Teach. Teach. Educ.* 66, 295–308. doi: 10.1016/j.tate.2017.04.015

Wolff, C. E., Jarodzka, H., and Boshuizen, H. P. A. (2021). 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

Wolff, C. E., van den Bogert, N., Jarodzka, H., and Boshuizen, H. P. A. (2015). Keeping an eye on learning. *J. Teach. Educ.* 66, 68–85. doi: 10.1177/0022487114549810

Wyss, C., Rosenberger, K., and Bührer, 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. *Compreh. Psychol.* 2, 1–7. doi: 10.2466/01.It.2.6

Yarbus, A. L. (1967). *Eye Movements and Vision (B Haigh, trans.)*. New York: Plenum Press.