Computers in Human Behavior Reports 16 (2024) 100481

Contents lists available at ScienceDirect

Computers in Human Behavior Reports

journal homepage: www.sciencedirect.com/journal/computers-in-human-behavior-reports

The development and validation of a video tool for capturing teachers' *noticing* in salient and non-salient classroom disruptions

Maxie Kilbury *, Anja Böhnke, Sebastian Haase, Felicitas Thiel

Department of Education and Psychology, Institute of School Pedagogy and School Improvement Research, Freie Universität Berlin, 14195, Berlin, Germany

1. Introduction

Classroom disruptions constitute one of the greatest problems for teachers (Meister & Melnick, 2003; Thiel, 2016). They are events that lead to the disruption of class flow and thus to a loss of learning time for the entire class (Thiel, 2016). Potentially problematic behavior should be perceived and identified early (Ophardt & Thiel, 2013; Thiel, 2016). Teachers can then take preventative action before a classroom disruption arises. The early perception of relevant events constitutes a major challenge for teachers because many complex events occur simultaneously during instruction (Barth, 2017; Doyle, 2006). An additional challenge lies in the perception of non-salient (less easily noticed) events when salient events occur at the same time (van den Bogert, 2016; Wolff et al., 2016). However, the perception of non-salient events is also relevant: On the one hand, critical characteristics of behavior can point to fundamental motives of the pupils (Thiel, 2016). On the other hand, pupils who behave inconspicuously but do not follow the expectation formulated in the program of action endanger their own progress in learning as well as the instruction of the entire class e.g., because they misunderstand contents and assignments, have to ask questions about them, or do not compete them at all.

The perception and identification of events relevant to instruction, often designated as *noticing* in the research literature, constitutes one aspect of competence in professional vision (Barth, 2017; Sherin, 2001). Experts have generally good competence in *noticing* (Bastian et al., 2022; Berliner, 1988; Grub et al., 2022; Kosel et al., 2023), while the attention of novices is often strongly bound to only individual salient events (van den Bogert, 2016). *Noticing* both, salient and non-salient events should therefore be systematically dealt with and developed already in teacher education.

Video instruments have become a common medium to investigate the *noticing*-competence of (pre-service) teachers, (e.g., Hoth et al., 2018; Kleinknecht & Gröschner, 2016; Kramer et al., 2020; Scholten et al., 2020; Seidel & Stürmer, 2014; van Es & Sherin, 2010) and integrated eye-tracking methods brought many new insights into the visual processes of perception (e.g., Cortina et al., 2015; Grub et al., 2022b). Often eye-tracking methods were combined with later group discussions, or questionnaires to determine what had participants identified to be relevant in some particular respect, (e.g., Hoth et al., 2018; Kramer et al., 2020). Up to now, there exists no video instrument which can capture both perception and identification of relevant events simultaneously and "on-the-fly" in such a way that it can be employed practically in the context of teaching and learning. Previously developed instruments are technically very complex (e.g., eye-tracking) and are hardly practical on a larger scale. Furthermore, the evaluation of collected data places intensive requirement on resources.

We have developed a video instrument which is able to capture *noticing* of critical events "on-the-fly", can be employed in a digital context, and in addition, can immediately report back which relevant events were identified (or overlooked). The videos produced for this purpose focus on salient and non-salient critical events that can occur simultaneously. This article describes the development and validation of the video instrument based on an expert-validation, a content-validation via specific criteria and a construct validation, including an expert-novice-comparison.

2. Theoretical background

2.1. Noticing

Noticing as a professional competence of teachers is under extensive research (König et al., 2022). However, very diverse conceptualizations have been utilized, as has been discussed König et al. (2022) in their systematic review. The most well-known concepts are "withitness" (Kounin, 1970), "situation awareness" (Endsley, 1995), and *noticing* as the basis of professional vision (Goodwin, 1994; Seidel et al., 2010; Sherin & van Es, 2009). *Noticing* was long interpreted as a holistic concept involving processes of being attentive/perceiving, interpreting/reasoning, reacting/deciding, and/or establishing connections. Only since 2013 has this competence been increasingly differentiated

* Corresponding author. Habelschwerdter Allee 45, JK 23/240,14195, Berlin, Germany.
E-mail address: maxie.kilbury@fu-berlin.de (M. Kilbury).

https://doi.org/10.1016/j.chbr.2024.100481
Received 3 June 2024; Received in revised form 30 August 2024; Accepted 30 August 2024
Available online 31 August 2024

into various facets and separately examined (König et al., 2022). Barth (2017) made the different approaches explicit, compared them systematically and developed a model of professional vison in which *noticing* as well as other facets like reasoning, generating, deciding, and implementing are built upon professional knowledge (the first facet in the model). In the context of dealing with potentially problematic behavior, for example, professional knowledge includes understanding the establishment of behavioral expectations and a working alliance (Oevermann, 1997; Ophardt & Thiel, 2013), knowledge of management of class flow (e.g., through verbal and nonverbal signals, Doyle, 2006; Ophardt & Thiel, 2013) and knowledge of individual pupil behavior and the modification of this behavior (Ophardt & Thiel, 2013). Barth (2017) distinguishes between two subprocesses of *noticing*: perception and identification:

1. Perception describes the "sensory capture of the environment," the "comparison of the acquired information with existing knowledge," and the "assignment of meaning to the information", that is, the identification of what has been perceived. Attention is decisive for perception and can be controlled variously: salience-driven or knowledge-driven (Chun et al., 2011). Salience-driven attention is controlled by object features and is independent of one's own goals. Salient features have a great stimulus and therefore attract much salience-driven attention. Knowledge-driven attention designates attention directed by one's goals and is controlled by experiences, knowledge, and one's own motives. For the perception of non-salient events, knowledge-driven perception is especially important because these have a weaker stimulus. Good perception means that teachers are continuously oriented about what is happening in the classroom (Kounin, 1970). However, teachers cannot attend to everything that happens in complex classroom situations; they must select events (van Es & Sherin, 2002). In order to do this, they must focus on lesson-relevant events. This localization, the attribution of relevance, occurs in a second process.
2. "Identification" designates the processing of captured information. This process involves identifying features of the perceived events that are relevant to the course of instruction because, for example, they could negatively affect it and should therefore be in the focus of attention. Criteria for processing the events are provided by knowledge-based schemas (Barth, 2017). These schemas describe a structural organization of professional knowledge in memory and serve as filters (Tsui, 2003) for relevant information and thus help to identify structural features (Peterson & Comeaux, 1987)

In order to represent the perception and identification of critical events in a video instrument authentically, typical features and behavioral expressions of critical events should be depicted. The following section therefore addresses the theory of classroom disruptions and potentially problematic pupil behavior.

2.2. Classroom disruptions and potentially problematic behavior

As has already been outlined, dealing with classroom disruptions poses a special challenge for teachers. Disruptions are triggered by potentially problematic behavior that impairs classroom interaction and subsequently impedes the program of action. The term program of action designates the teacher-controlled orientation of classroom interaction. It manifests itself in the expectations of who should do what with whom for how long with which materials and with which goal (Doyle, 2006). The teacher introduces a program of action for certain phases of instruction and is fundamentally unstable. Its implementation can be impaired if the teacher does not introduce it sufficiently, i.e., if the expectations for pupil behavior are not clearly explained, thereby resulting in unrest and extra questions, and/or pupils pursuing other impulses, such as chatting with peers. These critical events equally endanger the program of action (Doyle, 2006; Ophardt & Thiel, 2013). The present article does not focus on dysfunctional teacher behavior, but rather is restricted to the perception and identification of critical pupil behavior. Potentially problematic behavior varies significantly. Daydreaming/lack of focus, motoric restlessness, or so-called passing-time behavior (such as doodling or chatting, alone or with others) do not necessarily require a verbal reaction from the teacher, whereas deliberate provocations or demonstrative refusal to participate usually require intervention (Thiel, 2016, S. 123).

Signs of critical behavior can differ greatly in their salience. One and the same behavioral sign can manifest itself in quite different forms: motoric restlessness can be expressed by both wandering around or by moving legs under the table. Off-task conversation can be loud and conspicuous or quiet and inobtrusive and refusal to participate can manifest itself as a form of silent protest (Thiel, 2016) or very openly and loudly, etc. All the kinds of behavior listed here are equally relevant events. This means that the salience of events does not allow conclusions about their relevance. Therefore, teachers must perceive both salient as well as non-salient potentially problematic behavior early and identify relevant features in order subsequently to be able to draw conclusions about motives of the behavior. Only in this way can they take short- and long-term measures to prevent classroom disruptions and ensure individual learning progress.

Comparisons of experts with novices were able to show that experts generally are equipped with good competence in *noticing* (Bastian et al., 2022; Berliner, 1988; Grub et al., 2022; Kosel et al., 2023), while this requirement constitutes a special challenge for novices (Barth et al., 2018; Berliner, 1994; Endsley, 1995; Huang et al., 2020; Stahnke & Blömeke, 2021). Novices tend to notice only individual salient events (van den Bogert, 2016). Because non-salient events emit a low stimulus, perceiving them is particularly challenging, and novices tend to address them too late or not at all (Schulden et al., 2019). But up to now, studies investigating the perception of critical events do not focus on their salience (e.g., Grub et al., 2022b; Huang et al., 2020; van Driel et al., 2021), do not distinguish between events that are salient or non-salient which occur simultaneously (Telgmann & Müller, 2023; Wolff et al., 2016) or only investigate the perception of very few pupils in a class setting (six pupils in Telgmann & Müller, 2023).

2.3. Noticing of salient and non-salient critical events by experts

In the following it will be shown in detail how experts succeed (a) in perceiving both salient and non-salient critical events and (b) in identifying them as relevant.

(a) How do experts perceive events in the classroom?

Evaluations of process-based data from eye-tracking studies, such as scanpath or fixation duration, revealed that the gaze of experts in the classroom is typically wandering (e.g., Grub et al., 2022; van den Bogert et al., 2014). They glance at individual pupils more frequently but more briefly than novices. Kosel et al. (2023) established that this observational behavior of experienced teachers increases the chances that each pupil in the classroom receives some attention, independently of whether they are active or not and consequently regardless of the salience of their current behavior. Furthermore, experts have the ability to perceive and process parafoveal information, whereas novices still are more limited to foveal information (Gegenfurtner, 2020). Thus, experts generally gather more and different information than novices.

In addition, Kosel et al. (2023) determined that experts repeatedly "backtrack" to selected, relevant areas in order to update their visual information. This was termed "professional monitoring gaze behavior" (Grub et al., 2022), in keeping with the concept of "monitoring" established by Kounin (1970). Monitoring designates the competence of being continuously attentive to deviations from the program of action. The gaze behavior of experts is strongly self-directed and is aimed at perceiving as many pupils as possible, regardless of the salience of their behavior.

Experts thus manage to adequately meet the demands of two areas simultaneously: class focus and individual focus (Ophardt & Thiel, 2013). The class focus is required to maintain the action vector for the entire learning group. The individual focus involves addressing individual pupil behavior. Experts and novices both perceive stimulus-induced events; but experts are less strongly influenced by them and filter them according to their relevance. They are also able to maintain the class focus even during individual salient events, such as in the prevention or intervention of disruptions by individuals (Thiel et al., 2012; Van den Bogert et al., 2014).

The basis of good perception of salient and non-salient critical events is characterized by knowledge-driven perception of as many pupils as possible. Experts' perception is distinguished not only by their stronger distribution of attention to the entire class, but also by their ability to maintain class focus in situations where much attention is required by individual salient events. The dense information perceived involves both relevant and irrelevant events. Unlike novices, experts are not overwhelmed by the flood of information (Sabers et al., 1991), but respond simultaneously with appropriate strategies and signals to critical events in order to maintain the flow of instruction (Thiel et al., 2012).

(b) How do experts recognize relevant events?

Experts cognitively reduce complex classroom situations to essential elements (Carter et al., 1988), that is, they selectively ignore and select perceived events (Gegenfurtner, 2020). They succeed in this all the better because they are familiar with classroom situations (Berliner, 1994), have experienced many similar critical events and were able to store knowledge about them (Van den Bogert et al., 2014). They could cognitively order and condense typical critical events into patterns. These patterns were then stored as various schemas for critical events (Blömeke et al., 2015) and serve as filters for relevant patterns of the features of critical events (Peterson & Comeaux, 1987; Tsui, 2003). This implicates that experts quickly recognize when pupils are not following the program of action. Well-established schemas relieve the demands on the working memory and thereby reduce the cognitive burden of processing visual information (Gegenfurtner, 2020). This results in perceived features quickly being associated with familiar schemas, which leads to the recognition of more relevant events even in complex situations (Huang et al., 2020).

Experts consciously control their perception with the goal of maintaining class focus and perceiving non-salient events as well. Knowledge-guided schemas for features of critical behavioral expressions support them in simultaneously filtering and recognizing relevant events. These insights gained from research on expertise along with the theoretical concepts of *noticing* and critical events in the classroom presented here provide the basis for the content and technical production of the video instrument.

3. Development of the video instrument for capturing *noticing* of salient and non-salient critical events

3.1. Planning of the video instrument

As an aid for investigating the *noticing*-competence of (pre-service) teachers, video-taped lessons are often utilized (e.g., Hoth et al., 2018; Kleinknecht & Gröschner, 2016; Kramer et al., 2020; Scholten et al., 2020; Seidel & Stürmer, 2014; van Es & Sherin, 2002; van Es & Sherin, 2010). In recent years they have been often combined with Eye-Tracking methods, which have brought many new revelations of visual perception processes (e.g., Cortina et al., 2015; Grub et al., 2022b; Kosel et al., 2023; McIntyre & Foulsham, 2018; Schulden et al., 2019; Telgmann & Müller, 2023; Wolff et al., 2016). Here tracking-glasses collect process-based data involving eye movements from a video-taped lesson (or classroom). The output consists partly of information about the position and duration of visual fixation on so-called „Areas of Interest". Conclusions about the visual perception of the test persons can then be drawn from the modelling and analysis of these data (Gegenfurtner, 2020). But the data from eye movements permit no immediate conclusion about what the test persons identify as being relevant. This must therefore be established by triangulation with other methods, e.g., group discussions, questionnaires or interviews (Grub et al., 2022; Kosel et al., 2023; Stahnke & Blömeke, 2021; Telgmann & Müller, 2023; Wolff et al., 2016).

One approach for simultaneously capturing *noticing* as perception and identification has been presented by Grub et al. (2022a). They asked test persons to stop a video as soon as they see events relevant to a particular question and to make a note of this. It can be argued, however, that the authenticity of the perception suffers from the interruption since it is not continuous and therefor is unrealistic. Other methods require the test persons to "think aloud" simultaneously while watching a lesson video, that is, to report what they identify as relevant just at the moment (e.g., Grub et al., 2022b). This method however risks that the verbalization of one's own thoughts draws attention away from the classroom situation. The test persons can thereby focus only on very selective events.

An instrument for investigating the *noticing* of salient and non-salient critical events should capture both perception and recognition simultaneously and "on-the-fly" in such a way that it can be employed practically in the context of teaching and learning. Furthermore, it should address knowledge- and salience-driven perception as well as the identification of typical critical features and validly capture them. The derived content requirements for the videos are as follows:

- To address a "roving gaze" and parafoveal information intake, the events to be recognized should occur locally dispersed.
- Some pupils should display critical behaviors repeatedly to necessitate information updating, requiring a recurrent backward glance.
- To subtract stimulus-induced attention and address goal-directed attention, salient and non-salient events should occur simultaneously.
- Individual and class focus should be simultaneously demanded.
- To address the recognition of features, typical critical behavioral expressions reported by experts from teaching practice should be depicted.

Technical requirements for the development of the instrument:

- To capture *noticing* without distortion, the technical tool for perception should not impair perception.
- The instrument should be easily useable in the teaching-learning context, requiring no elaborate technology for implementation as well as data evaluation.

3.2. Production of the video clips

To implement the explained theoretical and content-related demands, the use of staged videos is particularly suitable. These can realistically depict specific classroom situations (Deng et al., 2020; Piwowar et al., 2017) for which authentic videos would be very difficult to produce (like seriously disrupted lessons (Thiel et al., 2020). They also allow for didactic condensation according to specific content requirements (Deng et al., 2020; Kilbury et al., 2023). In general, staged videos can be perceived as authentic, cognitively and emotionally engaging by learners (e.g., Codreanu et al., 2020; Dieker et al., 2009; Kramer et al., 2020; Piwowar et al., 2017), but must therefore adhere to certain content quality criteria. The orientation of the video production towards the criteria of *authenticity*, *engaging*, *challenging* and *relevant* has proven itself in this respect (Kim et al., 2006; Piwowar et al., 2017).

Based on the discussed criteria and demands, we developed video clips focusing on typical salient and non-salient critical pupil behaviors.

Initially, we developed lesson scripts based on authentic lesson plans of Mathematics, German, and Political Science. These scripts included lesson content and tasks aligned with various activity structures (Berliner, 1983). The focus was on transition phases from one activity to another, as these are inherently critical (Thiel et al., 2012). In the next step, we designed critical events for each of these sequences, based on theory-driven critical behavioral expressions (Thiel, 2016) and supplemented typical critical events identified in an expert survey of two teachers. When implementing these events in the sequences, care was taken to include in each sequence both non-salient behaviors, such as subtle chatting between two classmates, playing with a phone, or a passive attitude, as well as salient events, such as direct provocations of the teacher or classmates by individual pupils. In some sequences, the (individual) focus is on one pupil in particular. The events always occur at different positions in the classroom and sometimes repeatedly. In addition, the number of non-salient and salient events and their duration varies. (For more details about the content of the video clips, see Kilbury et al., 2023).

The outlined scenarios were translated into scripts, detailing non-verbal aspects such as facial expressions, gestures, and intonation of all actors. The scripts underwent validation in a multi-step process (Kilbury et al., 2023). Subsequently, cooperating theater classes rehearsed the scenes under intensive supervision by the project team. The roles of the teachers were played by professional actors. The final shooting took place in collaboration with a professional film crew under the direction of the project team. Maintaining a total perspective was crucial, ensuring that all pupils are visible in the frame to identify critical behaviors. Due to the nature of the shoot, small changes to the original script had to be decided on location - for example, when events were not visible from the total perspective.

A total of 20 video clips, ranging from seven to 52 s, were produced. The quality criteria (Kilbury et al., 2023; Kim et al., 2006; Piwowar et al., 2017) were addressed as follows:

Relevant: The videos should align with the teaching and learning objectives and occur in an authentic setting. Implementation: The video clips show realistically depicted classes and classrooms. The content was designed in close alignment with a theoretical basis that is taught to pupils within the digital learning environment (this is presented in the following chapter). This criterion can therefore be considered fulfilled and is not taken up in the subsequent study.

Realistic: The videos should appear authentic through complex representations. Implementation: The teaching situations and critical behaviors presented are mainly based on authentic material, such as the collection of critical events by teachers and authentic lesson plans.

Engaging: Learners must be cognitively engaged in the situation in order to have their focus of attention on the task, which is important for positive learning outcomes (Codreanu et al., 2020). Therefore, videos should be multifaceted and stimulating. Implementation: Complex situations are shown and the critical behaviors that occur in them are presented in a multifaceted way. Relevant information must be obtained from both verbal and non-verbal aspects of behavior.

Challenging: The difficulty of the videos should vary. Implementation: The difficulty of the video clips varies due to the number of events that occur simultaneously, their salience and the different durations of the events.

3.3. Development of the video tool and establishment of the digital learning environment

The produced video clips were integrated into the university's proprietary digital learning platform, "tet.folio" (Haase et al., 2016), a programming interface for learning environments. To convey relevant professional knowledge, a multimedia theoretical reader (Thiel, 2016; Thiel & Böhnke, 2023) was developed as the first component of the learning environment. Understanding classroom management and dealing with challenging pupil behaviors are crucial for recognizing critical behaviors. The reader imparts central aspects, core concepts, and action strategies of classroom management and handling disruptions according to Ophardt and Thiel (2013) and Thiel (2016), such as the reduction of inappropriate behavior through negative consequences and the reinforcement of appropriate behavior by using strategies such as prompting (e.g., reminders of agreed rules or use of non-verbal cues) and fading, i.e. the gradual reduction of the reinforcements used.

Following the reader, the participants are processing the video clips. To be quickly oriented in the classroom contexts of the video clips, brief intros are shown beforehand. The intro presents the class in a neutral stance (see Fig. 1) and provides auditory information about the task, the current phase of work, and any special conditions of the situation (e.g., whether work on tablets is permitted). Subsequently, the video clip is shown and the participants are given the following instruction: "Please mark all the events you identify as critical during the lesson. The objective here is to identify student behavior that is or could potentially become critical. For example, if two pupils whisper to each other or pupils seem absent or inattentive. Anything that could disrupt the flow of the lesson. It is not a question of always and immediately reacting to each of these events, but rather of identifying every potentially critical event to begin with." For this purpose, a dynamic marking tool was developed. Participants can place a cross via mouse click (or with a finger on a touch screen) on an event they identify as critical while watching the video clip (see Fig. 2).

To evaluate the placement of the crosses, i.e., to determine which critical events were identified by the participants, we programmed (invisible) boxes, which enclose the areas of each critical event (e.g., pupils who are chatting, see section 4.1) in the video (see Fig. 3). The program displays the placement of a cross within the borders of a box with a "1" (=hit), if no cross is placed within a box it displays "0" (= no hit). Crosses placed outside the boxes are ignored. Additionally, the boxes consider at which time a cross is placed, as some events occur only during specific periods of the video clip. This technical development allows a data output showing which event was identified (and not identified) by each participant via their placement of the dynamic crosses.

To install these boxes, all critical events in each video clip had to be determined locally and temporally, meaning a "maximum solution" of critical events had to be developed. This was created based on an expert validation. The process is described in section 4.1.

In summary, we have developed a video instrument for capturing *noticing* of salient and non-salient critical events, which addresses 1. Knowledge- and salience-driven perception since salient and non-salient events must be perceived simultaneously. These occur scattered in the classroom, requiring a roving gaze and the parafoveal perception of events. 2. The occurring events must be identified based on typical features to be identified as critical. Perception and identification are captured through direct marking in the running video. For use in the teaching-learning context, it is essential that the instrument is flexible and easy to use. We achieved this by integrating the instrument into a learning environment whose platform can be accessed and operated

Fig. 1. Intro of a video clip.

Fig. 2. Marking in a video clip.

from any device (tetfolio.fu-berlin.de). Furthermore, we have developed an automatic evaluation procedure, which is unique in this form and represents a significant gain for the teaching-learning context, as learners can receive immediate feedback.

4. Method

After the development of the video instrument a validation of this instrument took place. To validate the video instrument, three steps were processed:

1. Expert validation of the critical events and their salience.
2. Content validation based on the presented criteria *challenging*, *realistic* and *engaging* (Kilbury et al., 2023; Piwowar et al., 2017).
3. Construct validation on the basis of an expert-novice comparison. The following hypotheses were tested:
   (a) Experts identify more critical events than novices.
   (b) In particular, experts identify more non-salient events than novices.

The marking tool was tested to ensure that it does not impair the *noticing* process.

4.1. Expert validation of the critical events

In order to finalize the video instrument has been validated by five experts with the research focus in classroom management. These experts have validated the video clips and their critical events, resulting in a maximum solution of critical events per each video clip. For this purpose, each of them marked every event in the 20 video clips produced that they considered to be critical. They had an overall interrater reliability of .45, which corresponds to moderate agreement (Altman, 1991). The experts then met with the project team in a workshop where the events were discussed for which the experts' results did not agree. For five video clips these discrepancies were not overcome even in the workshop, and these were therefore discarded. The experts identified a total of 77 critical events.

The events that occur have varying degrees of salience which were then categorized for each event by three experts. Events that are not obvious, e.g., because they take place far away or are concealed and very inconspicuous, were categorized as "non-salient". Events that are more clearly visible and take place more in the middle of the classroom were taken to be "moderately-salient". Very prominent events, such as direct provocation of the teacher or events that are clearly open and visible, were categorized as "highly-salient". The events that were assessed differently by the experts were then discussed and evaluated jointly.

The final instrument consists of 15 video clips, showing an average of five critical events with a mean duration of 15 s. Overall, there are 25 non-salient events, 29 moderately-salient events and 23 highly-salient events to be identified. An example on how a video clip is composed can be seen in Fig. 4.

4.2. Content validation and construct validation

4.2.1. Samples

Pre-service teachers: The sample consisted of 165 students from the BLINDED. The average age in the sample was 23.08 years ($SD = 6.85$), 72.1% were female, and 91.4% were in the 2nd Bachelor's semester of the teacher training.

Expert teachers: Twelve teachers were selected who, due to their additional service as delegated teachers at the university in pedagogical and didactic departments, have extensive knowledge of classroom management. According to Palmer et al. (2005), they can therefore be regarded as experts for this study.

4.2.2. Instruments

Testing the criterion *realistic*: The video clips should be perceived as authentic as possible. In order to verify this, the pre-service teachers were asked to answer the item "The lesson portrayed in the video clip comes across as believable" on a 5-point Likert scale ("strongly disagree"

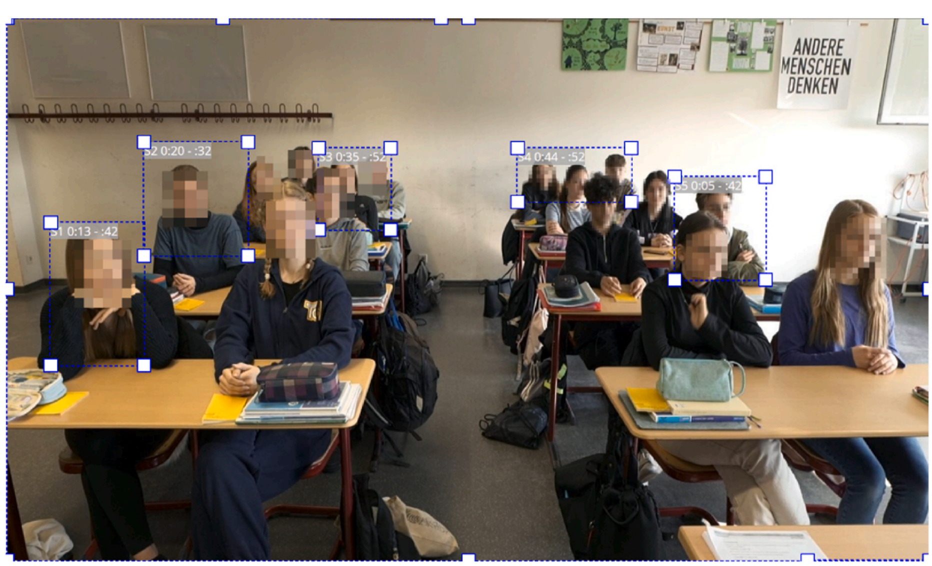

Fig. 3. Predefined boxes in a video clip (programmers' view).

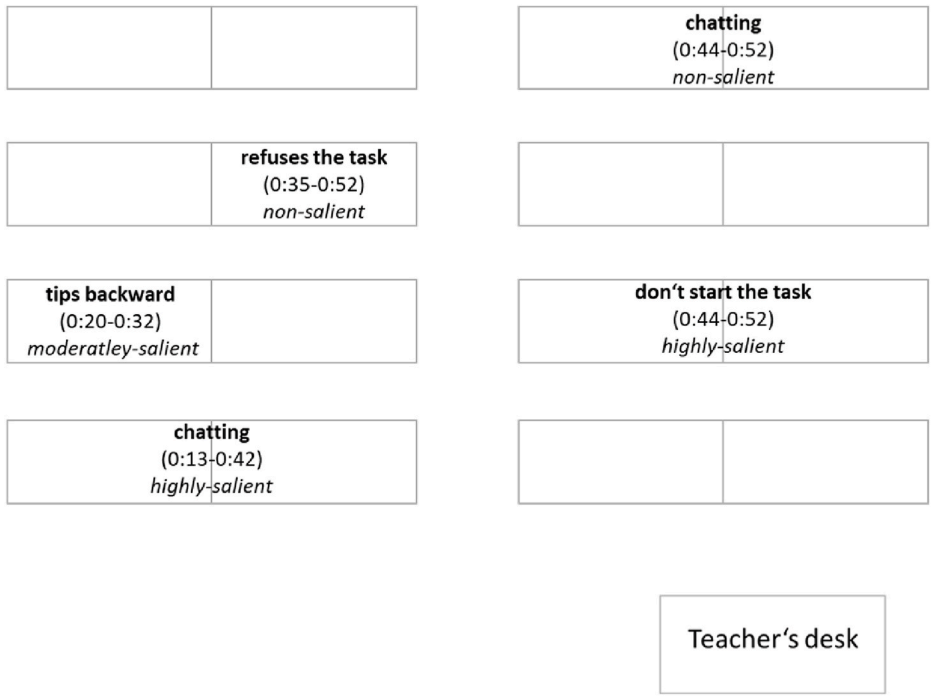

Fig. 4. Sketch of the events in a video clip in top-down perspective.

to "strongly agree") after each video clip they had viewed.

Testing the criterion *engaging*: The videos should be multifaceted and stimulating to engage the pre-service teachers emotionally and cognitively. This criterion was tested on a 5-point Likert scale ("strongly disagree" to "strongly agree") after all the video clips had been processed (see Table 1).

Testing the criterion *challenging*: The participants were asked: "Please mark all the events you identify as critical during the lesson". The item difficulty is defined by their hit rate per video clip. The item difficulty should vary, whereby the hit rate per clip should not fall below 20% or exceed 80%. On the average, the items should have a medium level of difficulty.

As has already been described, the maximum solution developed by experts was used to define the time and location of each event to be recognized in the video clips using a box. Markings within this box were categorized as hits. Before this study was conducted, the self-developed evaluation tool was tested as part of a pilot study. In addition, over 15% of the cases in the study were evaluated manually by the project team. The interrater reliability between the evaluation tool and the project team was very good, with a Cohen's kappa of .80. To further evaluate the item difficulty, the students themselves were asked to rate each clip on a 5-point Likert scale: "It was easy for me to identify critical events in this clip."

Testing the criterion *technical handling*: Marking the events should be technically easy. In order to test this, the expert teachers and pre-service teachers were asked to answer to the item: "Marking the events was technically easy for me." on a 5-point Likert scale ("strongly disagree" to "strongly agree") after all video clips had been processed. The results were analyzed descriptively using correlation analyses

Table 1
Measurement of the Criterion *engaging*.

Item	$\alpha =$ .83
(1) I enjoyed working through the video clips.	
(2) When working through the video clips I felt that I was dealing with challenging classroom situations.	
(3) Working through the video clips was interesting.	
(4) Trying to identify the critical events was an interesting challenge for me.	

4.2.3. *Procedure of the study*[1]

The pre-service teachers first worked through the theoretical content of the digital reader on classroom management and dealing with classroom disruptions (see Fig. 5, grey path). Subsequently they were given instructions on how to process the video clips and then were able to try out the marking tool in an exemplary video at their own pace. Following this, they watched the 15 video clips in a randomized order and were asked to mark the critical events. The video clips could only be played once. The experts processed the learning environment equally, except for the theoretical reader and the items of the content validation.

4.3. *Results*

4.3.1. *Dealing with missing values*

The proportion of missing data per item and video clip was above 5%. We then statistically analyzed the systematics of the cases with missing values with the Little test (Little, 1988) which was conducted in the statistical software SPSS. It was revealed that the data were not missing at random (χ 2 (44) = 167.191, $p < .01$), which meant that no imputation method could be applied (Fielding et al., 2008). Evaluation of the patterns of missing values SPSS showed that the most common reason for missing values was early dropout. We had expected participants to drop out early frequently as they worked through the cases and therefore varied the order of the video clips. In view of these findings, we decided to include all finished items and did not eliminate any cases. As a result, the sample sizes vary for each video clip and item.

In order to present a general overview regarding the results of the aforementioned criterions them being *realistic*, *engaging* and *challenging* they will first undergo a descriptive analysis. Subsequently, to analyze more profoundly how the participants identified the critical events in

[1] In planning and conducting the study, the project team adhered to the standards of the Ethical Principles for Psychologists and the Code of Conduct the American Psychological Association. (2017).

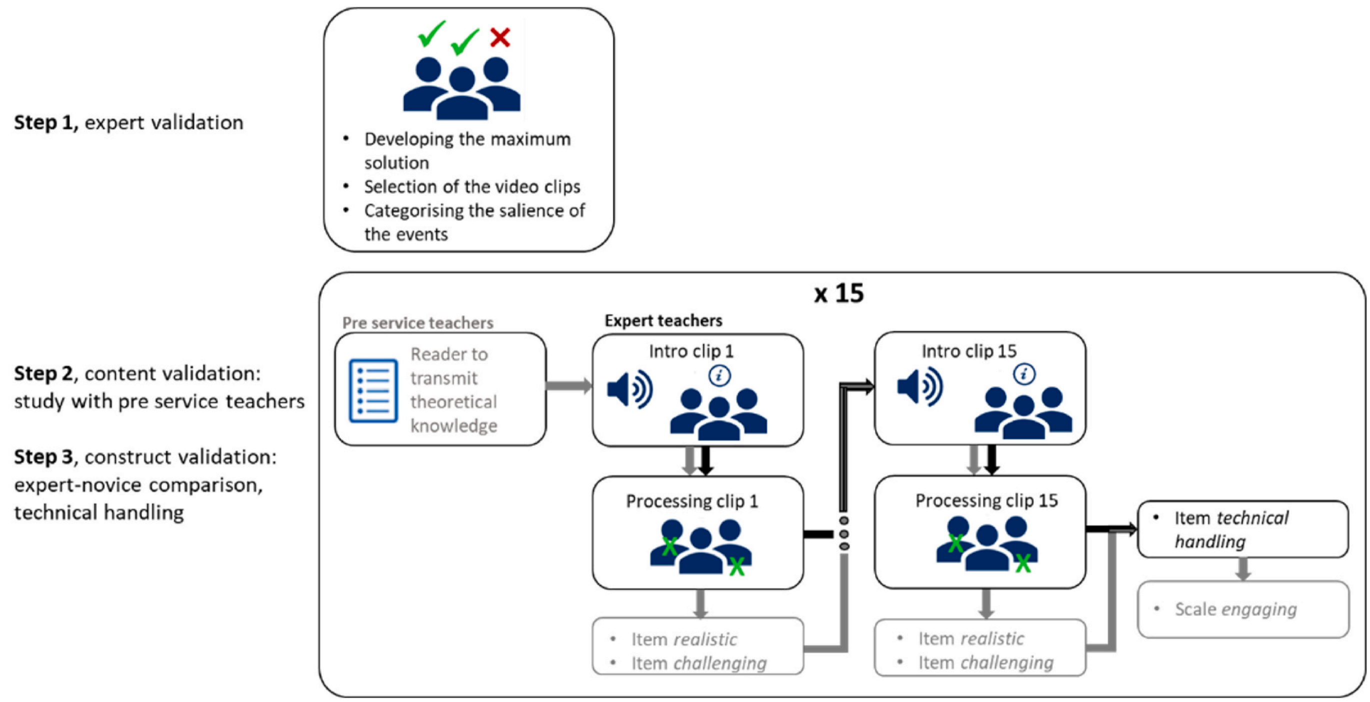

Fig. 5. Procedure of the study.

regard to their salience-levels, paired t-tests are conducted and, to examine the differences between the expert levels independent t-tests are conducted.

4.3.2. Assessment of the video clips for the criterion 'realistic'

The authenticity of all video clips was rated $M = 4$ ($N = 146$; $SD = .63$) on average and were therefore above the theoretical mean of 3. The lowest rating is 3.34 (clip 8) and the highest is 4.17 (clip 10).

4.3.3. Assessment of the video clips for the criterion 'engaging'

The calculated mean value of the scale for recording the criterion engaging was M = 3.78 (SD = .86) (see Table 2) and is thus higher than the theoretical mean value of 3, as is the case for all items of the scale. Especially item (4), i.e., how interesting the challenge was to identify critical events, received very high ratings.

4.3.4. Assessment of the video clips for the criterion 'challenging'

The item difficulty is defined by the hit rate per video clip as well as a self-assessment of the item difficulty. The mean hit rate per video clip was 48.32%. The lowest hit rate was obtained in clip 8 ($M = 28.57\%$; $SD = 15.94\%$) and the highest in clip 2 ($M = 68.40\%$; $SD = 17.99\%$) (see Table 3).

For better understanding, the values of the self-assessment scale are presented in reverse in order to follow the term "difficulty" more logically. This means that a value of '1' on the Likert scale of 1–5 indicates that identifying critical events was perceived as very easy and a value of '5' that it was very difficult. Overall, the processing of the video clips was rated on average as between medium and "tending towards" easy ($M = 2.54$; $SD = 1$). Clips 4 and 11 were rated as the most difficult ($M = 2.89$; $SD = 1.01$ and 1.10) and clip 8 as the easiest ($M = 1.88$; $SD = .77$). The average assessment of clip 8 clearly contradicts the actual hit rate. Possible explanations are addressed in the discussion.

After testing the linearity and excluding statistical outliers, correlation analyses of the hit rate and self-assessment were carried out for each clip. After correction of the Bonferroni alpha error, moderately significant negative correlations were found for seven of the 15 clips (see Table 3). This implicates that in nearly half of the video clips, the self-assessment is related to the actual hit rate.

Table 2
Results for the Criterion *engaging* ($N = 155$).

Item	*M*	*SD*
(1) I enjoyed working through the video clips.	3.71	1.14
(2) When working through the video clips I felt that I was dealing with challenging classroom situations.	3.58	1.08
(3) Working through the video clips was interesting.	3.69	1.07
(4) Trying to identify the critical events was an interesting challenge for me.	4.15	.91

4.3.5. Do experts identify more critical events than novices (=pre-service teachers)?

All together there was a total of 77 events to be identified in the clips. On the average, the experts identified $M = 59.24\%$ ($SD = 7.44\%$) and the novices $M = 48.08\%$ ($SD = 7.94\%$) of the events. After ascertaining a normal distribution, a *t*-test was conducted. The *t*-test showed a significant group difference with a large effect, $t\,(144) = -3.94$, $p \leq .001$, $d = -1.43$. This means that the experts identified significantly more events than the novices.

The group difference was then analyzed depending on the salience of the events. There were 25 events that were non-salient, 29 events that were moderately-salient and 23 events that were highly-salient. Both experts and novices identified non-salient events least often (experts $M = 52.00\%$, $SD = 10.25\%$; novices: $M = 35.01\%$, $SD = 11.78\%$) and highly-salient events most often (experts: $M = 73.91\%$, $SD = 11.30\%$; novices: $M = 67.68\%$, $SD = 12.17\%$) (see Table 4). The experts identified 16.99%pt more of the non-salient events, 11.70%pt more of the moderately-salient events and 6.23%pt more of the highly-salient events than novices.

In order to analyze whether the mean values between the two groups differ statistically depending on their salience, further t-tests were conducted. After applying the Bonferroni correction of the alpha error, a significant group difference for the non-salient events with a large effect, $t\,(148) = -3.99$, $p \leq .001$, $d = -1.45$ and also a significant group difference with a large effect for the moderately-salient events, $t\,(145) = -3.42$, $p \leq .001$, d $= -1.24$ was found. The groups do not differ significantly in identifying the highly-salient events, $t\,(152) = 1.5$, $p =$

Table 3
Hit rate of the pre-service teachers, their self-assessment of the item difficulty and the correlation of these two variables.

	total events	Clip duration	Hit rate in %			Self-assessment item difficulty		
			N	*M* in %	*SD*	*M*	*SD*	*r*[a]
M all clips	5.07	00:21		48.32	19.74	2.54	1.00	
Clip 1	5	00:39	152	64.74	21.80	2.28	.89	−.024
Clip 2	5	00:18	150	68.40	17.99	2.29	.88	−.215
Clip 3	5	00:52	147	48.44	17.19	2,66	1.01	.026
Clip 4	5	00:28	150	49.73	22.31	2.89	1.01	−.268**
Clip 5	6	00:28	153	46.62	16.15	2.47	.93	−.124
Clip 6	5	00:17	149	41.34	18.70	2.66	.99	−.204
Clip 7	4	00:22	150	55.00	29.11	2.71	1.07	−.287**
Clip 8	5	00:19	140	28.57	15.94	1.88	.77	−.043
Clip 9	8	00:35	152	45.72	16.53	2.40	.95	−.436**
Clip 10	2	00:07	115	40.87	25.29	2.98	1.23	−.205
Clip 11	3	00:07	135	30.12	18.61	2.89	1.10	.072
Clip 12	8	00:17	150	58.75	17.36	2.29	.91	−.326**
Clip 13	6	00:14	151	51.37	16.09	2.31	1.02	−.314**
Clip 14	4	00:15	145	57.24	23.00	2.53	1.07	−.329**
Clip 15	5	00:07	142	37.89	20.03	2.80	1.16	−.272**

**p < .01.
[a] Pearson product moment correlation.

Table 4
Hit rate of the events depending on their salience.

	M non-salient	*SD* non-salient	*M* moderately-salient	*SD* moderately-salient	*M* highly-salient	*SD* highly-salient
Novices	35.01%	11.78%	43.04%	9.50%	67.68%	12.17%
Experts	52.00%	10.25%	54.74%	7.70%	73.91%	11.30%

.204.

While novices identified significantly more of the moderately-salient events than non-salient events, t (137) = −7.16, $p \leq .001$, $d = -.61$, there was no significant difference between the experts' identification of these two categories t (7) = -. 654, $p = .267$. Both novices and experts recognize significantly more highly-salient events than moderately-salient events (novices: t (139) = −25.47, $p \leq .001$, $d = -2.16$; experts t (7) = −5.8, p ≤ .001, $d = -2.05$).

Experts obtained a hit rate between 50% and 73% in most of the video clips. However, clips 8, 10, 11 and 15 diverge from this range with an average hit rate between 30% and 40%. In these clips, the experts have a similar hit rate to the novices (see Table 3). Possible causes of this anomaly are addressed in the discussion.

4.3.6. *Assessment of the technical handling of the marking tool*

In order to ensure that marking was easy to manage technically and did not interfere with the process of *noticing*, both groups were asked to answer one item after processing all video clips. Both the experts (N = 11; M = 3.91; SD = 1.22) and novices (N = 155; M = 3.73; SD = 1.22) found the technical handling rather easy. No significant group difference was found, U = 922.5, Z = .47, p = .636.

5. Discussion

Noticing non-salient and salient disruptive events is an important prerequisite for dealing effectively with classroom disruptions and should therefore be addressed and trained during the teaching education program. A corresponding instrument should be easy to handle technically and be resource-efficient. In this article, the development of a video instrument was described and validated on the basis of findings from *noticing* expertise research as well as on the basis of well-founded theories on classroom management and on dealing with classroom disruptions (Doyle, 2006; Oevermann, 1997; Ophardt & Thiel, 2013; Thiel, 2016; Thiel et al., 2012).

In the following, the reported results will first be discussed under the aspect of the content validity of the produced video clips. This was examined with the quality criteria *realistic*, *engaging* and *challenging*. Then it will be discussed whether the developed instrument validly captures the competence of *noticing* (construct validity). For this purpose, the hypotheses that experts identify more of the critical events in the video clips than novices and that the group difference is particularly evident for non-salient events were tested. In addition, the perceived use of the marking tool was analyzed.

5.1. *Content validation of the video clips with quality criteria*

The video clips developed were meant to be as *realistic* as possible, i. e. authentically described and at the same time *engaging*, in order not to negatively affect the learning process and sufficiently involve the students cognitively (Hölzer, 2019; Pawek, 2009; Rank et al., 2012). Overall, the video clips were rated above the theoretical average of 3 in terms of both *realistic* and *engaging* (see Table 2). No video clip was rated lower than 3.14 on average in terms of authenticity, so this criterion can be considered fulfilled for all video clips (c.f. Codreanu et al., 2020; Gold et al., 2013). Each item on the *engaging* scale was also rated better than the theoretical mean value of 3. The results of the study suggest that the pre-service teacher found identifying the events interesting on the one hand, and also that they enjoyed working on them (see Table 2).

The criterion *challenging* designates the requirement for difficulty of the item that is relevant in the teaching-learning context. This should vary and still allow an increase in competence. The range of the average hit rate is 28.57%–68.40% and the mean value is 48.32%, indicating good item difficulty. The self-assessed difficulty is on average slightly below the mean value of 3 (see Table 3), which indicates that the pre-service teachers did not find the task to be too difficult and therefore potentially frustrating. The correlation analyses between the hit rate and the students' self-assessment of their performance showed that there were moderate correlations for half of the clips, while no significant correlations were found for the other half (see Table 3). The relatively very optimistic self-assessment of clip 8 (M = 1.88; SD = .77) is particularly striking in view of the very low hit rate (28.57%; SD = 15.94%). To explain this, the events to be identified in the clip are examined in more detail: Two of the five events to be identified refer to pupils who do not have any materials ready on the table (contrary to the work instruction). The critical behavior is thus expressed here by the fact that something is not done and therefore does not meet the expectations of the program of action. Another event that occurs in the clip relates to a pupil who is eating a chocolate bar in the last row, half concealed and behaving extremely unobtrusive. These three events, which were barely identified by the pre-service teachers (and experts), are extremely difficult to identify. As a result, the teaching situation does not appear obviously problematic, which could have led to the students' misjudgment of their own performance. The average, somewhat optimistic self-assessment could also be explained (beyond clip 8) by the fact that pre-service teachers assessed situations with many non-salient events as less critical than they actually were. Another possible explanation is the Dunning-Kruger effect (Kruger & Dunning, 2000), i.e. the tendency to overestimate oneself in the case of one's own lack of competence.

5.2. Construct validation for capturing noticing using an expert-novice comparison and assessment of the technical implementation

Experts have better *noticing* competences than novices (Gegenfurtner, 2020; Grub et al., 2022b; Van den Bogert et al., 2014). To determine whether the developed instrument captures *noticing* validly, the hypothesis that experts identify more events than novices was tested. The inferential statistical analysis showed that experts identified significantly more events than novices, thus confirming this hypothesis. In addition, differential analyses showed that experts were significantly better at identifying non-salient and moderately-salient events than novices, which also confirms the research findings described above. No significant group difference was found in identifying of highly-salient events (see also van Driel et al., 2021). Furthermore, it was established that novices recognize more moderately-salient events than non-salient events and, in turn, more highly-salient events than moderately-salient events. This justifies the assumption that the instrument successfully differentiates *noticing* with regard to salience. One assumption suggested by the results with regard to *noticing* expertise is that it does not make any difference to experts whether an event is non-salient or moderately-salient. Further research on this should be conducted.

When considering the experts' results with regard to the video clips, the comparatively very low hit rates of clips 8, 10, 11 and 15 were remarkable. It has already been noted in the previous section that the events in clip 8 were also much more difficult for the experts to identify than those in others. Clips 10, 11 and 15 are characterized especially by the fact that they are shorter than the other clips, lasting less than 10 s (see Table 3). This suggests that even experts need a few seconds to orient themselves in unfamiliar teaching situations and thus raises the question of whether video clips shorter than 10 s are at all suitable for validly capturing *noticing*. Further research into this would also be advisable.

The second requirement for valid assessment of *noticing* competence is uncomplicated use of the marking tool, which should influence the perception process as little as possible. On average, both experts and novices rated the technical application better than the theoretical mean of 3, with no significant difference between the two groups. The marking tool thus does not lead to interference with the perception of events and is independent of *noticing* expertise.

The sum of these results can be interpreted as an indication of the construct validity of the instrument produced.

Never the less, some limitations of the study should also be noted. The study was conducted online. Unfortunately, the conditions under which the participants worked on the clips, e.g., how attentive or distracted they were by external circumstances, could not be included in the study.

The sample size of the two groups was very different and the group of experts was very small, which resulted in limitations in the analysis. It would be advisable to repeat this study with a higher sample size of experienced teachers.

6. Further directions

This article presents the development and validation of a video instrument for capturing *noticing* for used in a teaching-learning context. In principle, the video instrument can also be used to train *noticing* skills. However, the evaluation of the hits in the chronological order of the clips shows no increase in the hit rate over the duration of the processing (see Fig. 6).

It can be assumed that an increase in *noticing* skills cannot be achieved through identifying alone, but that the video-clips should be enriched with attention-focusing elements for training purposes. Findings from teaching contexts in medicine, sports and the humanities point to the effectiveness of cueing techniques in improving novices' *noticing* competences (De Koning et al., 2009; Merkt & Sochatzy, 2015; Ryu et al., 2013). Cues are visual signals that direct learners' attention to essential elements and relevant information (Boucheix & Guignard, 2005; Brunvand, 2010; Jarodzka et al., 2013). This is achieved, for example, by increasing the luminance or visually highlighting relevant areas. Using cues in the video clips to help students direct their attention in a more goal-driven manner and to less salient events is a promising approach for the developed *noticing* tool that is currently being developed by the project team.

For further technical development it would also be interesting to investigate whether the video instrument could be linked further to other technologies. For example, Theelen et al. (2019) are researching the use of 360-degree instructional videos. The advantage of this approach is that it allows participants to choose their own perspective in the classroom.

In their content the video cases developed also offer potential for further processing options in the teaching-learning context. In the cases developed (in addition to typical critical behavior of pupils), specific characteristics of inadequate classroom management by the teacher were also implemented, such as a lack of signal control, lack of establishment of routines, etc. (Ophardt & Thiel, 2013; Thiel, 2016). In this way, further possible questions may be directed at critical teacher behavior.

7. Conclusion

In summary, it can be said that the criteria-based production of the video clips, the strong reference to a theoretical framework and the development of the instrument on the basis of findings from expert research into *noticing* have led to the development of a valid video instrument for capturing *noticing*.

The resulting instrument offers great potential for teacher training. It can be used flexibly and in an economical manner and can support the promotion of professional vision in a variety of ways.

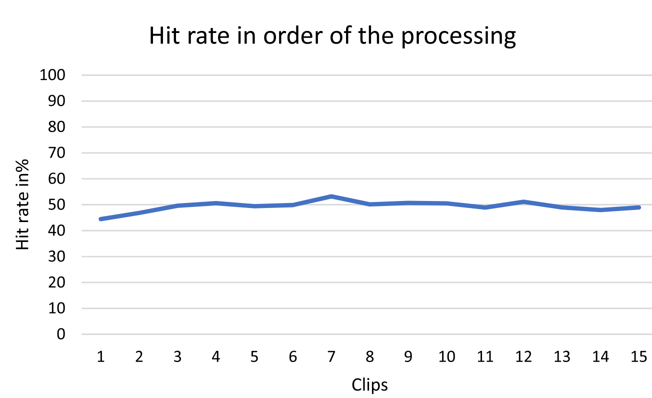

Fig. 6. Hit rate of the pre-service teachers in chronological order.

Disclosure statement

No potential conflict of interest was reported by the authors.

CRediT authorship contribution statement

Maxie Kilbury: Writing – original draft, Formal analysis, Data curation, Conceptualization. **Anja Böhnke:** Writing – review & editing, Project administration, Conceptualization. **Sebastian Haase:** Software. **Felicitas Thiel:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Maxie Kilbury reports financial support was provided by Federal Ministry of Education and Research of Germany. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgements

The research project is funded by the BMBF (Bundesministerium für Bildung und Forschung; Federal Ministry of Education and Research of Germany, funding code 16DHB3031). We would like to thank the students, teachers, and the administrative staff involved for this research.

Appendix A. Overview of the lessons

Clip	Subject and Topic	Activity Structure
1	German – Analyzing a novel	The teacher is starting the lesson by recapitulating the previous lesson.
2		The teacher is explaining today's task.
3		The pupils are starting with the task (individual work).
4	Mathematics – Integral calculus	The teacher is starting the lesson by explaining today's content.
5		The teacher is giving instruction for the partner work.
6		The handouts for the task are being distributed.
7		The results are discussed in plenum.
8	Politic Sciences – Analyzing a speech	The teacher welcomes the class and discusses the last lesson.
9		The teacher is introducing today's task and political speeches.
10		The pupils form groups of for and start the task.
11		One group presents their results in front of the class.
12	German – film genres	The teacher is welcoming the class.
13		The teacher is explaining the task, which is to read a film review.
14		The pupils are reading und underlining a film review.
15		The teacher brings the working phase to a close and starts a discussion.

References

Altman, D. G. (1991). *Practical statistics for medical research.* Chapman and Hall. https://doi.org/10.1201/9780429258589

American Psychological Association. (2017). *Ethical principles of Psychologists and code of Conduct.* American Psychological Association.

Barth, V. L. (2017). *Professionelle Wahrnehmung von Störungen im Unterricht (1. Aufl. 2017).* Fachmedien Wiesbaden: Springer.

Barth, V.L., Thiel, F. & Ophardt, D. (2018). Professionelle Wahrnehmung von Unterrichtsstörungen: Konzeption einer videobasierten Fallanalyse mit offenem Antwortformat. In A. Krüger, F. Radisch, A.S. Willems, T.H. Häcker & M. Walm (Hrsg.), *Empirische Bildungsforschung im Kontext von Schule und Lehrer*innenbildung* (S. 141–153). Klinkhardt.

Bastian, A., Kaiser, G., Meyer, D., Schwarz, B., & König, J. (2022). Teacher noticing and its growth toward expertise: An expert–novice comparison with pre-service and in-service secondary mathematics teachers. *Educational Studies in Mathematics, 110*. https://doi.org/10.1007/s10649-021-10128-y

Berliner, D. (1983). Developing conceptions of classroom environments: Some light on the T in classroom studies of ATI. *Educational Psychologist, 18*(1), 1–13. https://doi.org/10.1080/00461528309529256

Berliner, D. (1988). *The development of expertise in pedagogy*.

Berliner, D. (1994). *Expertise: The wonder of exemplary performances*.

Boucheix, J.-M., & Guignard, H. (2005). What animated illustrations conditions can improve technical document comprehension in young students? Format, signaling and control of the presentation. *European Journal of Psychology of Education, 20*(4), 369–388. https://doi.org/10.1007/BF03173563

Brunvand, S. (2010). Best practices for producing video content for teacher education. *Contemporary Issues in Technology and Teacher Education, 10*(2), 247–256 (10 Seiten) http://www.citejournal.org/articles/v10i2currentpractice1.pdf.

Carter, K., Cushing, K., Sabers, D., Stein, P., & Berliner, D. (1988). Expert-novice differences in perceiving and processing visual classroom information. *Journal of Teacher Education, 39*(3), 25–31. https://doi.org/10.1177/002248718803900306

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology, 62*, 73–101. https://doi.org/10.1146/annurev.psych.093008.100427

Codreanu, E., Sommerhoff, D., Huber, S. A., Ufer, S., & Seidel, T. T. (2020). Between authenticity and cognitive demand: Finding a balance in designing a video-based simulation in the context of mathematics teacher education. *Teaching and Teacher Education, 95*, Article 103146. https://doi.org/10.1016/j.tate.2020.103146

Cortina, K., Miller, K., McKenzie, R., & 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. *International Journal of Science and Mathematics Education, 13*. https://doi.org/10.1007/s10763-014-9610-5

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review, 21*(2), 113–140. https://doi.org/10.1007/s10648-009-9098-7 (28 Seiten).

Deng, M., Aich, G., Bakaç, C., & Gartmeier, M. (2020). Fictional video cases on parent-teacher conversations: Authenticity in the eyes of teachers and teacher education students. *Education Sciences, 10*(3), 63. https://doi.org/10.3390/educsci10030063

Dieker, L. A., Lane, H. B., Allsopp, D. H., O'Brien, C., Butler, T. W., Kyger, M., Lovin, L., & Fenty, N. S. (2009). Evaluating video models of evidence-based instructional practices to enhance teacher learning. *Teacher Education and Special Education, 32*(2), 180–196. https://doi.org/10.1177/0888406409334202

Doyle, W. (2006). Ecological approaches to classroom management. In C. M. Evertson, & C. S. Weinstein (Eds.), *Handbook of classroom management* (pp. 97–126). Routledge (Hrsg.).

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors The Journal of the Human Factors and Ergonomics Society, 37*(1), 32–64. https://doi.org/10.1518/001872095779049543

Gegenfurtner, A. (2020). *Professional vision and visual expertise*.

Gold, B., Förster, S., & Holodynski, M. (2013). Evaluation eines videobasierten Trainingsseminars zur Förderung der professionellen Wahrnehmung von Klassenführung im Grundschulunterricht. *Zeitschrift für Padagogische Psychologie, 27*(3), 141–155. https://doi.org/10.1024/1010-0652/a000100

Goodwin, C. (1994). Professional vision. *American Anthropologist, 96*(3), 606–633. https://doi.org/10.1525/aa.1994.96.3.02a00100

Grub, A.-S., Biermann, A., Lewalter, D., & Brünken, R. (2022). Professional knowledge and task instruction specificity as influencing factors of prospective teachers' professional vision. *Teaching and Teacher Education, 109*, Article 103517. https://doi.org/10.1016/j.tate.2021.103517

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

Grub, A.-S., Biermann, A., Lewalter, D., & 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. *Frontiers in Education, 7*. https://doi.org/10.3389/feduc.2022.890690

Haase, S., Kirstein, J., & Nordmeier, V. (2016). The technology enhanced textbook: an HTML5-based online system for authors, teachers and learners. In *Ludwig-Maximilians-Universität München (Vorsitz), Selected papers from the 20th international conference on multimedia in physics teaching and learning. Symposium im Rahmen der Tagung*. München: von Ludwig-Maximilians-Universität.

Hölzer, H. (2019). Authentizität spielen lernen. Simulation in der medizinischen Ausbildung. In M. Ankele, C. Kaiser, & S. LedeburHrsg (Eds.), *Aufführen - Aufzeichnen - Anordnen: Wissenspraktiken in Psychiatrie und Psychotherapie* (pp. 91–111). Springer. https://doi.org/10.1007/978-3-658-20151-7_6. S.

Hoth, J., Kaiser, G., Döhrmann, M., König, J., & Blömeke, S. (2018). A situated approach to assess teachers' professional competencies using classroom videos. In O. Buchbinder, & S. KuntzeHrsg (Eds.), *ICME13Mo. Mathematics teachers engaging with representations of practice* (pp. 23–45). Springer International Publishing. https://doi.org/10.1007/978-3-319-70594-1_3. S.

Huang, Y., Richter, E., Kleickmann, T., Wiepke, A., & Richter, D. (2020). Classroom complexity affects student teachers' behavior in a VR classroom. *Computers & Education*. https://doi.org/10.1016/j.compedu.2020.104100. Vorab-Onlinepublikation.

Jarodzka, H., van Gog, T., Dorr, M., Scheiter, K., & Gerjets, P. (2013). Learning to see: Guiding students' attention via a Model's eye movements fosters learning. *Learning and Instruction, 25*, 62–70. https://doi.org/10.1016/j.learninstruc.2012.11.004

Kilbury, M., Böhnke, A., & Thiel, F. (2023). Producing staged videos for teacher education: development and content validation of video scripts on the topic of handling classroom disruptions. *Education Sciences, 13*(1). https://doi.org/10.3390/educsci13010056.

Kim, S., Phillips, W. R., Pinsky, L., Brock, D., Phillips, K., & Keary, J. (2006). A conceptual framework for developing teaching cases: A review and synthesis of the literature across disciplines. *Medical Education, 40*(9), 867–876. https://doi.org/10.1111/j.1365-2929.2006.02544.x

Kleinknecht, M., & Gröschner, A. (2016). Fostering preservice teachers' noticing with structured video feedback: Results of an online- and video-based intervention study. *Teaching and Teacher Education, 59*, 45–56. https://doi.org/10.1016/j.tate.2016.05.020

König, J., Santagata, R., Scheiner, T., Adleff, A.-K., Yang, X., & Kaiser, G. (2022). Teacher noticing: A systematic literature review of conceptualizations, research designs, and findings on learning to notice. *Educational Research Review, 36*. https://doi.org/10.1016/j.edurev.2022.100453

Kosel, C., Böheim, R., Schnitzler, K., Holzberger, D., Pfeffer, J., Bannert, M., & Seidel, T. [T. (2023). Keeping track in classroom discourse: Comparing in-service and pre-service teachers' visual attention to students' hand-raising behavior. *Teaching and Teacher Education, 128*, Article 104142. https://doi.org/10.1016/j.tate.2023.104142

Kounin, J. S. (1970). *Discipline and group management in classrooms*. R.E. Krieger Pub. Co.

Kramer, M., Förtsch, C., Stürmer, J., Förtsch, S., Seidel, T. T., & Neuhaus, B. J. (2020). Measuring biology teachers' professional vision: Development and validation of a video-based assessment tool. *Cogent Education, 7*(1). https://doi.org/10.1080/2331186X.2020.1823155

Kruger, J., & Dunning, D. (2000). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*, 1121–1134. https://doi.org/10.1037//0022-3514.77.6.1121

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*(404), 1198. https://doi.org/10.2307/2290157

McIntyre, N. A., & Foulsham, T. (2018). Scanpath analysis of expertise and culture in teacher gaze in real-world classrooms. *Instructional Science, 46*(3), 435–455. https://doi.org/10.1007/s11251-017-9445-x

Meister, S., & Melnick, D. (2003). National new teacher study: Beginning teachers' concerns. *Action in Teacher Education, 24*(4), 87–94.

Merkt, M., & Sochatzy, F. (2015). Becoming aware of cinematic techniques in propaganda: Instructional support by cueing and training. *Learning and Instruction, 39*, 55–71. https://doi.org/10.1016/j.learninstruc.2015.05.006

Oevermann, U. (1997). Theoretische Skizze einer revidierten Theorie professionalisierten Handelns. In A. Combe, & W. HelsperHrsg (Eds.), *Suhrkamp-taschenbuch wissenschaft: Bd. 1230. Pädagogische professionalität: Untersuchungen zum typus pädagogischen handelns* (pp. 70–182). Suhrkamp. S.

Ophardt, D., & Thiel, F. (2013). *Klassenmanagement: Ein Handbuch für Studium und Praxis*. W. Kohlhammer.

Palmer, D. J., Stough, L. M., Burdenski, J., Thomas, K., & Gonzales, M. (2005). Identifying teacher expertise: An examination of researchers' decision making. *Educational Psychologist, 40*(1), 13–25. https://doi.org/10.1207/s15326985ep4001_2

Pawek, C. (2009). *Schülerlabore als interessefördernde außerschulische Lernumgebungen für Schülerinnen und Schüler aus der Mittel- und Oberstufe*. Universitätsbibliothek Kiel: Kieler Dissertationen online. https://nbn-resolving.org/urn:nbn:de:gbv:8-diss-36693. http://eldiss.uni-kiel.de/macau/receive/dissertation_diss_00003669.

Peterson, P. L., & Comeaux, M. A. (1987). Teachers' schemata for classroom events: The mental scaffolding of teachers' thinking during classroom instruction. *Teaching and Teacher Education, 3*(4), 319–331. https://doi.org/10.1016/0742-051X(87)90024-2

Piwowar, V., Barth, V. L., Ophardt, D., & Thiel, F. (2017). Evidence-based scripted videos on handling student misbehavior: the development and evaluation of video cases for teacher education. *Professional Development in Education, 44*(3), 369–384. https://doi.org/10.1080/19415257.2017.1316299.

Rank, A., Gebauer, S., Hartinger, A., & Fölling-Albers, M. (2012). Situiertes Lernen in der Lehrerfortbildung [Situated learning in in-service teacher training]. *Lehrerbildung auf dem Prüfstand, 5*(2), 180–199. https://doi.org/10.25656/01:14737

Ryu, D., Kim, S. S., Abernethy, B., & Mann, D. L. (2013). Guiding attention aids the acquisition of anticipatory skill in novice soccer goalkeepers. *Research Quarterly for Exercise & Sport, 84*(2), 252–262. https://doi.org/10.1080/02701367.2013.784843

Sabers, D. S., Cushing, K. S., & Berliner, D. C. (1991). Differences among teachers in a task characterized by simultaneity, multidimensional, and immediacy. *American Educational Research Journal, 28*(1), 63–88. https://doi.org/10.3102/00028312028001063

Scholten, N., Höttecke, D., & Sprenger, S. (2020). How do geography teachers notice critical incidents during instruction? *International Research in Geographical & Environmental Education, 29*(2), 1–17. https://doi.org/10.1080/10382046.2019.1670915

Schulden, M., Schüürmann, A., Käter, C., & Hillenbrand, C. (2019). Alle im Blick?! Aufmerksamkeitserfassung im Klassenzimmer mittels Eye-Tracking. https://doi.org/10.13140/RG.2.2.25220.58246.

Seidel, T. T., Blomberg, G., & Stürmer, K. (2010). In E. Klieme, D. Leutner, & M. Kenk (Eds.), *Observer" – Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. Projekt OBSERVE*. Beltz (Hrsg.), *Zeitschrift für Pädagogik. Kompetenzmodellierung: Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes* (Bd. 56, S. 296–306.

Seidel, T. T., & Stürmer, K. (2014). Modeling and measuring the structure of professional vision in preservice teachers. *American Educational Research Journal, 51*(4), 739–771. https://doi.org/10.3102/0002831214531321

Sherin, M. G. (2001). Developing a professional vision of classroom events. In T. Wood, B. S. Nelson, & J. Warfiel (Eds.), *Beyond classical pedagogy: Teaching elementary school mathematics* (pp. 75–93). Erlbaum (Hrsg.).

Sherin, M. G., & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education, 60*(1), 20–37. https://doi.org/10.1177/0022487108328155

Stahnke, R., & Blömeke, S. (2021). Novice and expert teachers' noticing of classroom management in whole-group and partner work activities: Evidence from teachers' gaze and identification of events. *Learning and Instruction, 74*, Article 101464. https://doi.org/10.1016/j.learninstruc.2021.101464

Telgmann, L., & Müller, K. (2023). Training & prompting pre-service teachers' noticing in a standardized classroom simulation – a mobile eye-tracking study. *Frontiers in Education, 8*. https://doi.org/10.3389/feduc.2023.1266800

Theelen, H., van den beemt, A., & Pj, B. (2019). Using 360-degree videos in teacher education to improve preservice teachers' professional interpersonal vision. *Journal of Computer Assisted Learning, 35*. https://doi.org/10.1111/jcal.12361

Thiel, F. (2016). *Interaktion im Unterricht: Ordnungsmechanismen und Störungsdynamiken. UTB Schulpädagogik*. In *, 4571. Bd.* Verlag Barbara Budrich. http://www.utb-studi-e-book.de/9783838545714.

Thiel, F., & Böhnke, A.. *Analyse von und Umgang mit herausforderndem Schüler*innenverhalten im Unterricht*. https://tetfolio.fu-berlin.de/tet/1464668.

Thiel, F., Böhnke, A., Barth, V. L., & Ophardt, D. (2020). How to prepare preservice teachers to deal with disruptions in the classroom? Differential effects of learning with functional and dysfunctional video scenarios. *Professional Development in Education*, 1–15. https://doi.org/10.1080/19415257.2020.1763433.

Thiel, F., Richter, S. G., & Ophardt, D. (2012). Steuerung von Übergängen im Unterricht: Eine Experten-Novizen-Studie zum Klassenmanagement. *Zeitschrift für Erziehungswissenschaft, 15*(4), 727–752. https://doi.org/10.1007/s11618-012-0325-5.

Tsui, A. (2003). *Understanding expertise in teaching: Case studies of second language teachers* (1. Publ). *Cambridge applied linguistics*. Cambridge University Press.

van den Bogert, N. J. (2016). On teachers' visual perception and interpretation of classroom events using eye tracking and collaborative tagging methodologies. https://api.semanticscholar.org/CorpusID:217522725.

Van den Bogert, N., van Bruggen, J., Kostons, D., & Jochems, W. (2014). First steps into understanding teachers' visual perception of classroom events. *Teaching and Teacher Education, 37*, 208–216.

van Driel, S., Crasborn, F., Wolff, C. E., Brand-Gruwel, S., & Jarodzka, H. (2021). Exploring preservice, beginning and experienced teachers' noticing of classroom management situations from an actor's perspective. *Teaching and Teacher Education, 106*, Article 103435. https://doi.org/10.1016/j.tate.2021.103435

van Es, E., & Sherin, M. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education, 10*.

van Es, E. A., & Sherin, M. G. (2010). The influence of video clubs on teachers' thinking and practice. *Journal of Mathematics Teacher Education, 13*(2), 155–176. https://doi.org/10.1007/s10857-009-9130-3

Wolff, C. E., Jarodzka, H., van den Bogert, N., & Boshuizen, H. P. A. (2016). Teacher vision: Expert and novice teachers' perception of problematic classroom management scenes. *Instructional Science, 44*(3), 243–265. https://doi.org/10.1007/s11251-016-9367-z