# Vorschlag 1:

All Eyes on the Classroom – How Teacher Expertise Shapes Gaze Behavior and Classroom Management

# Vorschlag 2:

Who Masters Classroom Disruptions - Expertise Differences in Classroom Management and Gaze Behavior in a Laboratory Setting

# Abstract

*Keywords:* classroom management, professional vision, eye-tracking, expertise differences, managing classroom disruptions

# Introduction (Level 1)

Classroom disruptions, such as students chatting with peers or nervously clicking their pens, can significantly interfere with learning by consuming valuable instructional time, reducing student engagement, and impairing academic achievement (Keller, 2014; Chow et al., 2024; Kraft & Monti-Nussbaum, 2021; Marder et al., 2023). In particular, managing classroom disruptions poses a major challenge for beginning teachers, who often feel ill-prepared to cope with such situations – a finding consistently highlighted in international studies (e.g., Friedman, 2006; Stokking et al., 2003; Veenman, 1984; Christ, 2004) and described as a key stressor contributing to the so-called “reality shock” among novice teachers (Dicke et al., 2015). Effectively managing such disruptions is, therefore, a key component of successful classroom management (Helmke, 2022) and plays a significant role in maintaining teachers’ health and well-being (Braun et al., 2022; Dicke et al., 2015).

A central element in enabling teachers to manage such situations effectively is their *professional vision* – the ability to notice, interpret, and respond to relevant events in the classroom based on expert knowledge (Goodwin, 1994; Sherin & van Es, 2009). It links strategic knowledge with concrete action, allowing teachers to anticipate and address disruptions proactively (Gold et al., 2016).

Expert teachers differ from novices not only in the amount of classroom experience but also in how they perceive and respond to complex instructional situations. They demonstrate more refined professional vision and possess a broader range of strategies for managing challenges such as disruptions (Berliner, 2001, 2004; Gold & Holodynski, 2015; Emmer & Gerwels, 2006; Barth, 2017).

To gain a deeper understanding of the cognitive processes underlying such expertise, process-based methods such as eye-tracking have become increasingly important. Eye-tracking enables researchers to analyze where and how teachers focus their visual attention during instruction (Dessus et al., 2016; van den Bogert, 2016). Results indicate that expert teachers demonstrate more efficient gaze patterns, reflecting more accurate situational assessment (Gegenfurtner et al., 2020; Wolff et al., 2016).

This study examines differences in gaze behavior, disruption noticing, and classroom management between experienced and inexperienced teachers in a controlled micro-teaching unit in which teachers were confronted with a series of scripted classroom disruptions. Using eye-tracking and self-report measures, it investigates how teachers distribute their attention, how quickly they notice different types of disruptions, and whether expertise is reflected in more efficient gaze patterns, higher confidence, lower perceived disruptiveness, and better strategic knowledge. The study aims to contribute to a deeper understanding of professional vision and its relevance for effective classroom management, especially as it is learnable and can be systematically fostered (Sherin, 2007; Tucholka & Gold, 2025).

**Classroom Disruptions (Level 2)**

As described by Lohmann & Meyer (2003), disruptions are events that interfere with the essential conditions necessary for effective instruction and student participation. These behaviors can generally be categorized into four primary types, based on their nature and impact on classroom dynamics (Lohmann & Meyer, 2003). *Verbal disruptions* include spoken interruptions such as chatting, whispering, or heckling, which can disturb lesson flow and diminish focus. *Physical disruptions* refer to motor restlessness or unnecessary physical activity, like drumming on desks, snipping fingers, or clicking pens, which distract both teachers and students. Indicators of a *lack of eagerness to learn* manifest through disengagement behaviors, such as drawing, resting one’s head on the desk, or using a phone, actions that reduce participation and hinder comprehension. Finally, *aggressive behavior* encompasses hostile actions or emotional outbursts, including yelling, defiance, or physical confrontations, all threatening the classroom’s safety.

A crucial concept in understanding classroom disruptions is salience, which refers to how noticeable a behavior is within the classroom context (Kilbury et al., 2024). Salience directly affects teachers’ ability to notice disruptions – a key prerequisite for effective management. Highly salient behaviors, such as loud outbursts, tend to capture immediate attention, whereas more subtle behaviors may go unnoticed yet can be equally detrimental if left unaddressed over time.

Before any management strategy can be applied, a disruption must first be perceived and interpreted as requiring intervention. This process is complicated by the subjective nature of disruption, shaped not only by the behavior itself but also by how it is evaluated by both teachers and students (Eckstein et al., 2016). This dual perspective underscores the importance of recognizing both overt and subtle signs of disruption in order to respond appropriately and effectively.

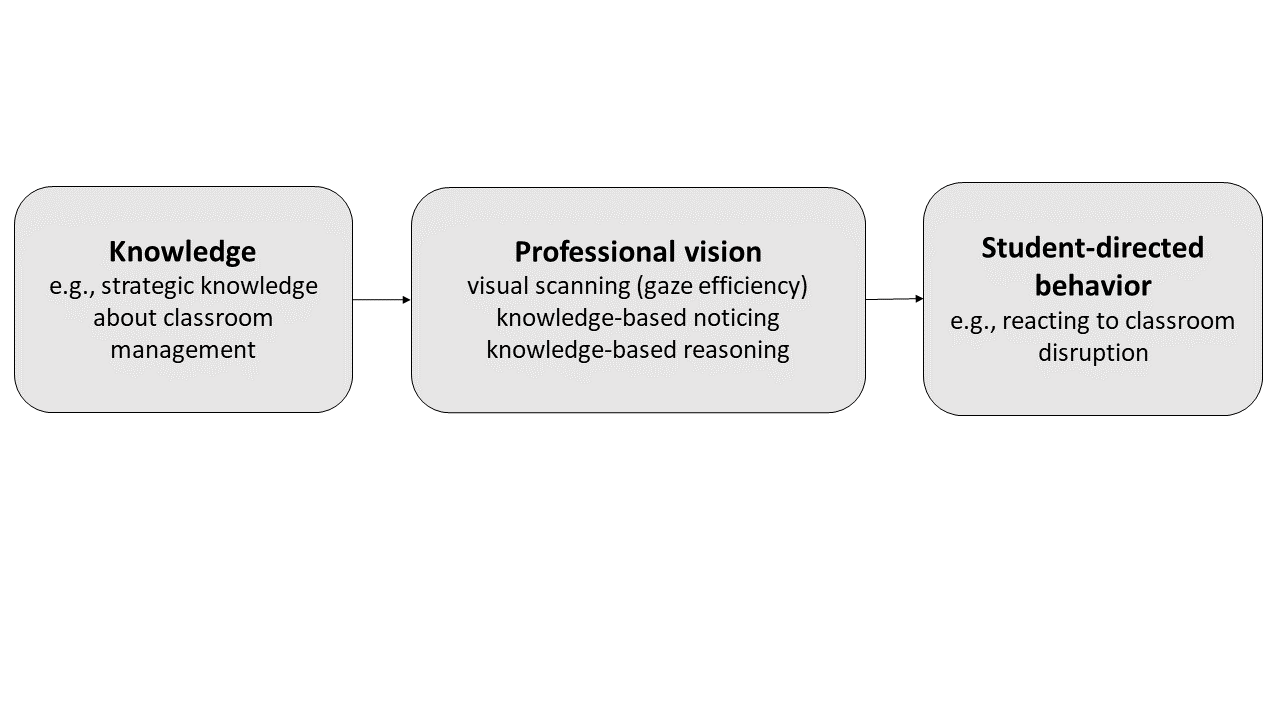
**Competence Development in Classroom Disruption Management (Level 2)**

The management of classroom disruptions is a central component of effective classroom management and requires teachers to develop context-sensitive strategies for responding to disruptive situations (Helmke, 2022). Effective disruption management depends on the ability to deal with the simultaneity and complexity of classroom events (Doyle, 1980), which demands a high level of knowledge and situational awareness.

Effectively addressing classroom disruptions requires teachers to continuously monitor their environment for behaviors that may interfere with the learning process. Figure 1 presents a simplified model that illustrates how knowledge interacts with perception processes – conceptualized as *professional vision* – and leads to observable teaching behavior. This model is adapted from the competence-to-performance transformation framework proposed by Blömeke et al. (2015) and further developed by Grub (2023).

**Figure 1**

*Model of Professional Competence Development in Classroom Disruption Management (Adapted From Blömeke et al. (2015) and Grub (2023), Modified by The Authors)*



As shown in Figure 1, teachers’ knowledge, such as strategic knowledge of classroom management, is essential for managing classroom disruptions, as it involves professional judgment in selecting appropriate responses within complex and ambiguous situations (Gold & Holodynski, 2015; Fenstermacher, 1994). As a core element of instructional competence (Barth, 2017; D’Agostino & VanWinkle, 2007; Kunter et al., 2011; Voss et al., 2014), it enables teachers to weigh alternative courses of action and choose strategies that maintain instructional flow while effectively addressing disruptions. This form of knowledge evolves through practical experiences, where declarative knowledge about classroom management is gradually transformed into procedural and situationally applicable expertise (Blömeke et al., 2022; Rauner, 2005). An example of strategic classroom management knowledge is the decision to respond to a student quietly distracting a peer during a presentation not with a direct reprimand, but with a subtle, non-disruptive intervention, such as calling on the student to participate or using eye contact. According to Gold and Holodynski (2015), such a response maintains instructional flow while addressing the misbehavior, reflecting the teacher’s ability to select a context-sensitive strategy rather than applying generic rules. Thus, managing disruptions does not rely solely on knowing rules or routines but on the ability to apply them flexibly and appropriately in context.

*Professional vision* serves as a crucial mediating skill between teachers’ dispositions and their classroom behavior. Originally introduced by Goodwin (1994) and adapted to teacher education by Sherin (2007), professional vision encompasses the ability to selectively perceive classroom events and interpret them based on pedagogical knowledge (Van Es & Sherin, 2002). In the initial stage of professional vision, *gaze efficiency* refers to the teacher’s ability to *visually scan* the classroom efficiently and purposefully (Cortina et al., 2015; McIntyre, 2016; Dessus et al., 2016; Huang, 2018; Gegenfurtner, 2020). For example, teachers with high gaze efficiency may regularly shift their gaze between students and quickly detect a student in the corner disengaging by looking away and fidgeting. Because teachers’ attention is not overly fixated on one area or object, but flexibly distributed, they can respond before the behavior develops into a broader disruption. Thus, gaze efficiency reflects an important facet of professional vision that enables teachers to early perceive relevant classroom events, such as classroom disruptions, supporting proactive classroom management (Grub et al., 2020).

In the next step, *knowledge-based noticing* refers to the teacher’s ability to not only perceive this behavior but also to recognize its pedagogical significance. For instance, interpreting the student’s fidgeting not merely as restlessness, but as an early sign of cognitive disengagement from the task (van Es & Sherin, 2002; König et al., 2022). This process goes beyond passive observation and is guided by the teacher’s prior knowledge and experience (Sherin et al., 2011), allowing them to filter what is relevant in the moment.

Following this, *knowledge-based reasoning* describes how teachers draw on their professional knowledge to decide how to act upon what they have noticed (van Es & Sherin, 2002; Gibson & Ross, 2016). In the example, the teacher may infer that directly addressing the student could cause embarrassment or disruption, and instead opts to move closer to the student’s desk or make subtle eye contact – nonverbal strategies that aim to re-engage the student without interrupting the flow of the lesson.

Together, these three processes – visual scanning, knowledge-based noticing, and knowledge-based reasoning – form a dynamic perceptual cycle in which teachers continuously observe the classroom, interpret what they see through the lens of their professional knowledge, and decide how to act accordingly. They bridge the gap between (strategic) knowledge and visible classroom action and are therefore pivotal for the development and demonstration of professional teaching competence.

The third component of the model, *student-directed behavior*, refers to the visible application of classroom management strategies, such as the effective reaction to classroom disruptions. In this regard, reactive strategies play a central role, particularly in addressing Low-Level Classroom Disruptions (LLCDs), which refer to minor, non-threatening, and often passive behaviors, such as chatting, fidgeting, or engaging in off-task activities (Cogswell et al., 2020). Though individually low in severity, LLCDs are perceived by teachers as particularly burdensome (Wheldall & Merrett, 1988) and can cumulatively interfere with instructional flow and reduce students’ active learning time (Helmke, 2007; Levin, 2010; Meyer, 2018).

Reactive disruption management aims to eliminate disturbances swiftly and with minimal interference to instruction, allowing the lesson to continue with as little interruption as possible (Lohmann, 2014). However, as every intervention constitutes a disruption itself, responses must be minimal, non-intrusive, and context-sensitive (Levin, 2010; Wettstein & Scherzinger, 2022). Nonverbal strategies, such as proximity, eye contact, or gestures, are particularly effective in managing mild disruptions and preventing ripple effects (Kounin & Gump, 1958; Nolting, 2013; Keller, 2008). When verbal intervention is required, it should be brief, specific, and behavior-focused (Levin, 2010; Nolting, 2013; Wahl, Weinert & Huber, 1984). If these remain ineffective, logical consequences may follow, ideally tied to the misbehavior and proportionate in nature (Levin, 2010; Keller, 2008).

Yet, how teachers react to classroom disruptions is not solely a matter of strategic knowledge and timing – it is also shaped by individual perception and emotional regulation. Research shows that the subjective perception of disruptions varies among teachers and is influenced by personal and contextual factors, including their emotional state, goals, and prior experiences (Wettstein et al., 2016; Eckstein, 2018). This subjective lens plays a key role in how teachers interpret and react to the same classroom event. For example, teachers with lower emotional self-regulation may respond more harshly or impulsively, while others maintain calm and composure (Makarova et al., 2014).

An important moderating factor in this process is teachers’ self-efficacy. Teachers with higher self-efficacy perceive disruptions as less threatening and manage them more confidently and constructively (Große Siestrup, 2010). Defined as the belief in one’s own capacity to deal with challenging situations (Schwarzer & Jerusalem, 2002), self-efficacy supports professional judgment and prevents overreactions triggered by stress or negative expectations. Moreover, Eckstein et al. (2022) demonstrated that the frequency and intensity of disruptions, teachers’ well-being, and their classroom context (e.g., type of instruction or class behavior norms) further shape their perception and response behavior.

In summary, the student-directed behavior phase of the model illustrates how teachers’ professional knowledge, perception, and self-regulatory abilities converge in deliberate, context-sensitive classroom actions. For example, a teacher who notices a student in the back corner fidgeting and looking away during a presentation may interpret this behavior, based on prior experience and pedagogical knowledge, as an early sign of disengagement. Rather than reacting with irritation or issuing a public reprimand, the teacher regulates their emotional response, refrains from impulsive action, and calmly walks closer to the student while maintaining eye contact or subtly involving them by asking a low-threshold, content-related question. This response demonstrates not only strategic knowledge and perceptual accuracy but also a high degree of emotional regulation and self-efficacy: the teacher trusts their ability to de-escalate the situation effectively without disrupting the flow of the lesson (Große Siestrup, 2010; Schwarzer & Jerusalem, 2002; Makarova et al., 2014). Such low-profile, nonverbal interventions help maintain instructional momentum and prevent the misbehavior from spreading (Kounin, 2006), while also modeling calm and respectful classroom leadership.

## Expertise Differences in Classroom Disruption Management (Level 2)

Building on the outlined model of professional competence in classroom disruption management (Ref.), it is essential to examine how these competencies manifest across different levels of teaching expertise. Teachers’ ability to respond effectively to classroom disruptions is not only shaped by their professional knowledge and perceptual accuracy but also by the extent to which these competencies are integrated and applied in practice.

### Teachers’ Knowledge (Level 3)

Strategic knowledge refers to context-sensitive, action-oriented knowledge that enables teachers to make informed pedagogical decisions in complex and dynamic situations such as disruptions (Fenstermacher, 1994; Doyle, 2006). A foundational aspect of this knowledge is the ability to recognize what constitutes a classroom disruption in the first place. Teachers need conceptual clarity about which student behaviors might interfere with instruction and how such behaviors manifest across varying classroom contexts. This understanding forms the basis for effective perception processes such as scanning and noticing. For example, knowing that early signs of disengagement (e.g., fidgeting, off-task glances) can precede more overt disruptions allows teachers to attend to these subtle cues proactively.

This connection between knowledge and perception is supported by findings from Casale et al. (2016), who emphasized that classroom management expertise relies heavily on situation-specific cognitive requirements, most notably the accuracy and holistic nature of perception as well as the justification of pedagogical decisions. In this sense, strategic knowledge includes not only the capacity to act effectively in response to disruptions but also the awareness of where to direct attention and how to interpret what is seen. Accordingly, gaze behavior itself, such as knowing how to distribute one’s visual attention efficiently across the classroom, is considered an element of strategic knowledge that supports professional vision and, ultimately, expert classroom action. Casale et al. (2016) demonstrated that expert teachers are more accurate and comprehensive in their perception and more able to provide pedagogically grounded justifications for their actions.

Furthermore, expertise research has shown that expert teachers differ significantly from novices in how their knowledge is structured and accessed. Experts possess highly organized and interconnected knowledge networks, which allow them to retrieve relevant information quickly and to apply it efficiently to novel situations (Livingston & Borko, 1989; Goldman, 2007). These knowledge structures are shaped through experience and enable teachers to integrate pedagogical concepts with situational demands. In contrast, novices tend to rely on fragmented, less structured knowledge that may lead to rigid or less adaptive decision-making (Shulman, 1987; Ophardt & Thiel, 2013).

Strategic classroom management knowledge develops through the interplay of theoretical learning and practical experience. While novice teachers often view classroom management primarily as discipline and control, expert teachers draw on a repertoire of proactive strategies aimed at establishing norms, preventing disruptions, and maintaining instructional flow (Helmke, 2022; Tynjälä et al., 1997). This development is cumulative and tied to growing professional judgment in navigating complex classroom dilemmas (Berliner, 2001).

Previous efforts to assess strategic knowledge of classroom management have either embedded it as a subdimension within broader pedagogical knowledge tests (e.g., COACTIV-R, Kunter et al., 2013; TEDS-M, Tatto et al., 2008) or used licensure-based assessments such as the Praxis Series (Educational Testing Service, 2011). However, these instruments are either not publicly available for research, not specifically focused on classroom management, or lack contextual richness in their scenarios. Moreover, they often emphasize propositional knowledge over strategic reasoning, limiting their utility for expertise research (Gold & Holodynski, 2015).

To address this gap, Gold and Holodynski (2015) developed a model-based SJT specifically designed to assess teachers’ strategic knowledge of classroom management. Their instrument includes realistic classroom scenarios that reflect the three major facets of classroom management (monitoring, managing momentum, and establishing rules and routines). Validation studies have demonstrated the SJT’s sensitivity to expertise: in-service teachers significantly outperform preservice teachers, especially in scenarios requiring proactive and nuanced management strategies (Gold & Holodynski, 2015). In contrast to dichotomous formats, their SJT asks participants to rate the effectiveness of each response, allowing for a more nuanced assessment. These findings support the claim that strategic knowledge is not only a core component of classroom management competence but also a reliable indicator of professional expertise.

### Teachers’ Professional Vision (Level 3)

~~To assess~~ Researchers studying expertise in professional vision typically rely on eye-tracking technology as it captures attentional behavior in real time and enables inferences about underlying cognitive mechanisms such as noticing and attention control, which typically occur unconsciously and cannot be verbalized reliably (Grub et al., 2020; Seidel & Stürmer, 2014; van den Bogert et al., 2014). Eye-tracking data are commonly analyzed using fixation-based metrics, as fixations are considered behavioral indicators of cognitive processing (Just & Carpenter, 1976). A fixation is defined as the moment when the eyes remain still over a visual object or person for a brief period, indicating focused attention (Holmqvist et al., 2011). Fixations are characterized by their frequency (number of fixations) and duration (average time in milliseconds), both of which can be used to infer how visual information is processed (Gegenfurtner et al., 2018; Negi & Mitra, 2020).

**Visually Scanning – Gaze Efficiency.** (Level 4) The Gaze Relational Index (GRI) summarizes the relationship between fixation duration and fixation count (Lowe & Boucheix, 2016) with lower GRI values indicating faster and more exploratory gaze behaviors (Gegenfurtner et al., 2020). In the educational, . Grub et al. (2022) conducted a lab-based, stationary eye-tracking study with scripted video vignettes, comparing 29 novice and 35 experienced teachers. Contrary to expectations, their study found no significant differences in GRI between groups. The authors suggest that the high salience of classroom disruptions in the videos may have elicited attentional responses across expertise levels.

In contrast, Kosel et al. (2023) applied mobile eye-tracking in authentic classroom settings, analyzing the gaze behavior of two novice and two experienced teachers. Their findings showed clear expertise effects: experienced teachers exhibited lower GRI values, consistent with faster, top-down controlled gaze patterns. This suggests that the GRI is better able to reflect visual expertise in ecologically valid, dynamic classroom environments. Notably, these teachers distributed their gaze across more students with shorter fixations, indicating more efficient scanning strategies aligned with expert classroom monitoring​. However, the interpretability of the findings is limited by the small sample size, which restricts the generalizability of the results.

Additionally, Gabel et al. (2023) conducted a stationary eye-tracking study with 135 pre-service teachers who viewed classroom videos under three instructional conditions (general instruction, specific task instruction, prompts). While the setting was lab-based, participants who received specific instructions or prompts showed slightly lower GRI values, reflecting more frequent and shorter fixations, a pattern associated with more efficient, top-down scanning behavior. This suggests that the GRI can not only reflect stable expertise differences but also capture instructionally induced shifts in visual scanning​. However, even in this controlled setting, GRI differences were modest, highlighting the need for more immersive environments to fully capture expertise-related effects​.

Thus, the reviewed studies suggest that the GRI is a promising metric for assessing visual scanning efficiency, particularly in mobile, in-situ settings where dynamic classroom complexity engages authentic teacher gaze behavior. While lab-based, video vignette studies (e.g., Grub et al., 2022; Gabel et al., 2023) showed mixed results, real-world applications (Kosel et al., 2023) demonstrated the GRI’s potential to meaningfully differentiate expert and novice scanning strategies.

**Knowledge-based Noticing.** (Level 4) In the context of classroom management, noticing is considered a foundational skill for effective teaching. It refers to teachers’ ability to selectively attend to relevant events while ignoring distractions (Sherin & van Es, 2009; Blömeke et al., 2015). A growing body of research uses eye-tracking to operationalize and compare visual attention in novice and expert teachers, with findings converging on more efficient, targeted, and distributed gaze behavior among experts (Cortina et al., 2015; McIntyre, 2016; Dessus et al., 2016; Huang, 2018; Biermann et al., 2023; Gegenfurtner et al., 2020).

For example, Cortina et al. (2015) used mobile eye-tracking in authentic mathematics classrooms to observe 12 mentor–mentee dyads. Mentors, as the more experienced teachers, distributed their gaze more evenly across students and focused more strongly on task-relevant cues. These patterns were positively associated with ratings of instructional support, suggesting that gaze distribution reflects aspects of teaching quality.

Dessus et al. (2016) studied two expert and two novice primary school teachers using mobile eye-tracking during full-length math lessons. Experts showed lower attentional lability and monitored student behavior more systematically. Novices, by contrast, focused narrowly on a few students, regardless of behavioral relevance.

Moreover, McIntyre (2016) conducted a cross-cultural study using mobile eye-tracking glasses with 40 secondary teachers from England and Hong Kong (10 novices and 10 experts per country). Experts showed more frequent and sustained gaze toward students, while novices fixated more often on instructional materials and irrelevant stimuli. The study supports the assumption that expert gaze is driven by instructional relevance and less influenced by peripheral distractions.

Huang (2018) examined gaze patterns in 25 expert–novice teacher pairs using mobile eye-tracking during real classroom instruction. Experts demonstrated broader gaze distribution, shorter and more targeted fixations, and reduced attentional bias toward salient but irrelevant objects. The findings suggest that expert teachers apply more efficient scanning strategies that help maintain situational awareness under complex conditions.

Time to first fixation (TTFF) has also emerged as a relevant metric in classroom management research (e.g., Yamamoto & Imai-Matsumura, 2013; Keskin et al., 2023). TTFF captures the latency between the onset of a salient classroom event (e.g., a disruption) and the first fixation on the relevant target (e.g., the disruptive student). Although the noticing speed is often interpreted as a marker of situational awareness (Van den Bogert, 2016; Wyss et al., 2021), findings are mixed. In a video-based eye-tracking study with 43 Japanese in-service teachers, Yamamoto and Imai-Matsumura (2013) found that teachers who noticed student misbehavior did not fixate earlier on the misbehaving students than those who failed to notice it, suggesting that TTFF alone may not reliably reflect awareness of disruptions.

In a video-based eye-tracking study with 135 pre-service teachers, Keskin et al. (2023) found that participants in the prompting condition fixated significantly faster on disruptive students than those in the specific task or control conditions, indicating that prompts can enhance early attention to critical classroom events.

Biermann et al. (2023) combined stationary eye-tracking, keystroke-based detection tasks, and stimulated recall interviews to examine how 52 novice and expert teachers noticed classroom disruptions in standardized video vignettes. While no significant differences in TTFF emerged, participants who successfully detected critical incidents showed more frequent fixations and revisits to behaviorally relevant areas. Experts also drew on a broader range of interpretive cues, highlighting the value of combining gaze, behavioral, and verbal data in the assessment of noticing. These results suggest that TTFF alone may not reliably reflect expertise but should be interpreted alongside complementary measures such as fixation frequency and distribution.

However, many eye-tracking studies in teacher research suffer from small sample sizes (e.g., Pouta et al., 2021) or methodological limitations, such as low ecological validity due to stationary setups (e.g., Stahnke & Blömeke, 2021) or low experimental control in authentic classrooms (e.g., McIntyre et al., 2020).

**Knowledge-based Reasoning.** (Level 4) Knowledge-based reasoning is a key element of teachers’ professional vision, describing how teachers interpret classroom situations based on their professional knowledge (Sherin & van Es, 2009; Seidel & Stürmer, 2014). In Blömeke et al.’s (2015) model of competence, interpretation is central: it links perception to decision-making and involves the evaluation of classroom events with regard to instructional relevance and behavioral significance. Especially in the context of classroom disruptions, interpretation is shaped not only by what teachers see but by how disruptive they perceive an event to be and how confident they feel in responding to it.

Empirical studies show that expert teachers differ from novices in both the content and focus of their interpretations. Reuker (2016) and Gegenfurtner et al. (2020) found that expert teachers generated more theoretically grounded and anticipatory reasoning, even under time pressure or when dealing with complex classroom situations. These experts did not react impulsively but evaluated situations more reflectively, suggesting that interpretations of disruptiveness are tied to professional knowledge structures, not just surface cues.

Wolff et al. (2017) specifically studied the interpretation of classroom disruptions. Using eye-tracking and concurrent think-alouds, they found that expert teachers made more coherent and instructionally focused interpretations, even when faced with subtle (“fuzzy”) disruptions. In contrast, novice teachers paid more attention to overt behavioral cues and tended to label situations as disruptive more quickly, often without linking them to instructional implications. These findings suggest that the perceived disruptiveness of a situation is not objective, but filtered through expertise and interpretation.

In parallel, teachers’ subjective confidence in handling situations appears to be closely linked to interpretive depth. Muhonen et al. (2021) found that more experience did not necessarily lead to more elaborated reasoning, challenging assumptions that confidence grows linearly with teaching experience. Gegenfurtner et al. (2020) showed that experts used metacognitive strategies during reasoning, indicating greater interpretive control and likely more situational confidence.

These studies highlight that both the perceived disruptiveness of classroom events and the confidence in dealing with them reflect deeper interpretive processes shaped by knowledge, attention, and expertise. Structured ratings of these aspects allow for a systematic investigation of how teachers evaluate challenging classroom situations, complementing verbal and visual data in the study of professional vision.

### Teachers’ Student-directed Behavior (Level 3)

Teachers’ observable student-directed behavior in managing classroom disruptions represents the final, outwardly visible stage of a complex internal process that begins with professional knowledge and continues through perception processes conceptualized as professional vision (Blömeke et al., 2015; Grub, 2023). These earlier stages – particularly knowledge-based visual scanning and noticing – are crucial for enabling timely, appropriate, and minimally intrusive interventions. Without perceiving and interpreting a situation correctly, any visible action is at risk of being ineffective or even counterproductive.

Expert teachers benefit from highly integrated knowledge structures that support swift interpretation and context-sensitive reasoning. They are more likely to detect subtle cues that precede disruptions, such as early signs of disengagement or restlessness, and to differentiate between types of disruptive behavior in terms of salience, intent, and instructional relevance (Kilbury et al., 2024; van Es & Sherin, 2002). This allows them to respond flexibly, with either silent proximity, subtle gestures, or verbal redirects, without disturbing the instructional flow (Thiel et al., 2012; Emmer & Gerwels, 2006).

By contrast, novice teachers often struggle with these processes. Their perception is less differentiated, and they are more likely to miss low-salient but pedagogically significant behaviors, leading to delayed or inappropriate reactions (Barth, 2017; Swanson et al., 1990). When responding, they tend to use direct, verbal reprimands that can disrupt the lesson and risk escalating student behavior (Westerman, 1991). Their actions may be driven more by momentary irritation or uncertainty than by pedagogical reasoning.

Importantly, not all disruptions are equally salient or easily recognizable. While verbal or physical disruptions tend to attract attention, more subtle indicators of disengagement may go unnoticed, especially by novice teachers (Kilbury et al., 2024). This underlines the role of visual attention and situation-specific knowledge in detecting and interpreting low-level, but pedagogically relevant, behavior.

In addition to perceptual skill and strategic knowledge, affective and motivational components also shape how teachers respond to disruptions. As Busse (2021) highlights, teachers’ self-perceived competence in classroom management significantly influences their emotional regulation and choice of action. Teachers with a strong sense of competence are more confident, less reactive, and more able to maintain a constructive stance, even in challenging situations.

Korevaar’s (1998) research adds another dimension to this picture by focusing on teachers’ reaction intentions and their underlying causal attributions in problematic student-teacher interactions. Her findings showed that experienced teachers displayed more adaptive and complex response profiles, while novices were more likely to respond impulsively or defensively. This highlights that behavioral differences are not only driven by perceptual and strategic differences but also by deeper cognitive-affective mechanisms.

In sum, teachers’ observable classroom actions are not isolated events but rather the outcome of an integrated competence structure: knowledge, perception, interpretation, and execution. However, while prior research has yielded valuable theoretical insights, many studies rely on self-reports, video-based simulations, or hypothetical scenarios. As a result, little is known about the real-time perceptual and behavioral processes during teaching itself, especially when it comes to noticing and responding to actual classroom disruptions.

## Present Study (Level 2)

The present study addresses this gap by employing mobile eye-tracking in micro-teaching units within a highly standardized learning environment, using a comparatively large sample. This approach balances authenticity and experimental control by capturing real teaching-learning interactions while enhancing standardization through scripted behavioral instructions.

The study examines differences in gaze behavior, self-reported classroom management, and strategic knowledge between experienced and inexperienced teachers. Data were collected in the laboratory-based study *Professional Vision of Novice and Expert Teachers* (ProVisioNET), where both in-service (experienced) and pre-service (inexperienced) teachers participated in a controlled micro-teaching unit involving classroom disruptions. Eye-tracking technology, self-report measures, and strategic knowledge assessments were used to explore how expertise manifests in visual attention, disruption noticing, and management strategies.

Participants individually attended a laboratory session, where they conducted a brief micro-teaching unit lasting approximately 15 minutes. The “class” consisted of three trained actors who portrayed students and simulated typical classroom disruptions. During the micro-teaching unit, teachers’ gaze patterns were recorded using eye-tracking technology. Additionally, self-reports on classroom management and strategic knowledge were assessed through a questionnaire, an interview, and a test.

The study addressed five primary aims:

One aim was to examine how teachers distribute their visual attention during classroom interactions in a lab-based micro-teaching unit. We expected students to be the primary focus of teachers’ attention. This exploration was descriptive.

Another aim was to investigate differences in gaze behavior related to noticing abilities between experienced and inexperienced teachers. We hypothesized that, compared to inexperienced teachers, experienced teachers would demonstrate more efficient gaze patterns throughout the micro-teaching unit (**Hypothesis 1a**), focus their attention more frequently on students (**Hypothesis 1b**), and notice disruptions more quickly (**Hypothesis 1c**).

Additionally, we examined whether the type of disruption (i.e., verbal disruptions, physical disruptions, and indicators of lack of eagerness to learn) influenced the speed at which teachers noticed disruptions. We hypothesized that verbal and physical disruptions would be more salient and, therefore, noticed more quickly than indicators of a lack of eagerness to learn (**Hypothesis 2a**). Furthermore, we explored whether teaching expertise had an impact on the disruption noticing speed.

Another aim was to determine whether differences in teaching expertise were reflected in classroom management measures, including self-reports on classroom management (e.g., self-evaluated classroom management, disruptiveness, and confidence ratings) as well as strategic knowledge of classroom management. We expected experienced teachers to outperform inexperienced teachers (**Hypothesis 3a**).

Finally, we examined the relationship between gaze behavior and classroom management measures, expecting these variables to be correlated.

# Method (Level 1)

## Participants (Level 2)

We recruited a total of 84 teachers from Germany (42 pre-service teachers and 42 in-service teachers) through personal contacts, email lists, and flyers. Pre-service teachers were required to be actively enrolled in a teacher education program and to have completed their first internship, while in-service teachers needed to both have fully completed their teacher training (including the preparatory training year, the “Referendariat”) band to be currently working in the teaching profession. Data from two in-service teachers were excluded due to low-quality eye-tracking data, resulting in a final sample of 82 teachers, comprising 42 pre-service teachers and 40 in-service teachers.

The pre-service teachers (*n*= 29 women, *n*= 13 men) had a mean age of 22.80 years (*SD* = 1.90; range: 19–27). On average, they were in their 7th semester (*M* = 6.70, *SD* = 2.60; range: 3–11) and had an average of 9.60 hours (*SD* = 7.20; range: 1–36) of teaching experience through internships completed during their studies. The in-service teachers (*n*= 24 women, *n*= 16 men) had a mean age of 39.10 years (*SD* = 10.60; range: 26–60) and an average of 11.60 years (*SD* = 11.30; range: 1–38) of teaching experience.

The study adhered to the ethical guidelines put forth by the University of Leipzig and received approval from the University’s Institutional Review Board. Participants were fully informed about the study’s objectives before testing. Their participation was voluntary, not incentivized, and commenced only after written consent.

## Setting and Procedure (Level 2)

Participants individually attended the lab for approximately two hours, following a standardized procedure for which a seminar room was transformed into a classroom. Upon arrival, they were welcomed by the experimenter, introduced to the procedure, and asked to sign the data protection agreement. Participants were then fitted with a binocular Tobii Pro Glasses 2 eye-tracker, adjusted for comfort and vision. After performing an initial calibration of the glasses (for details of the eye-tracker and the calibration, see Appendix A), the experimenter activated and synchronized the recording devices (eye-tracking glasses, four cameras, and an audio recorder) using an auditory signal. This setup phase included a brief introductory game to acclimate participants to the eye-tracking equipment and the three students, which took approximately 10-15 minutes.

After the initial setup, a second calibration was done in a separate room. As soon as the teacher re-entered the classroom, the micro-teaching unit started. Participants were asked to prepare a 15-minute micro-teaching unit on a topic and grade level of their choice. The only requirement was that the unit had to be an introductory micro-teaching unit, and had to consist of supervised individual work and/or frontal teaching. During the unit, three trained actors (playing students) performed scripted classroom disruptions. The students followed prompts that appeared every 1.5 minutes on a screen only visible to them (e.g., chatting with a neighbor, heckling, looking at the phone; see Table A1 in the supplementary material for an overview and categorization of all disruptions; and Figure B1 and B2 in the supplementary material for a depiction of the laboratory setting of the micro-teaching unit). The order of the disruptions and the performing students were fully balanced using Latin Squares. The whole micro-teaching unit was recorded using eye-tracking glasses to capture teachers’ gaze patterns. The micro-teaching unit lasted about 15-20 minutes.

After the teaching session, participants underwent a third calibration. Following this, all recording devices were stopped, and participants filled out a brief computer-based questionnaire (~10-15 minutes) assessing sociodemographic data and a self-evaluation of their classroom management during the micro-teaching unit. A ten-minute break followed, concluding the first part of data collection.

In the study’s second part, participants engaged in a Stimulated Recall Interview (SRI). They watched a video of their own teaching session, recorded through the eye-tracking glasses, while the experimenter paused the video at each classroom disruption. Participants answered five open-ended questions and three rating questions for each disruption, including self-reported *disruptiveness* and *confidence ratings* (see Measures). The SRI lasted approximately 45-60 minutes. Finally, participants completed a Situational Judgment Test (SJT) online, assessing their *strategic knowledge of classroom management*. The questionnaire took approximately 15 minutes to complete, marking the end of the study.

## Measures (Level 2)

***Gaze Behavior Measures (Level 3)***

Gaze behavior was analyzed using predefined Areas of Interest (AOI) to examine how participants distributed their visual attention during the micro-teaching unit. To code and analyze the gaze behavior, we used the software Tobii Pro Lab Analyzer (Version 1.241.54542). A fixation filter was applied, with a threshold set at 30°/sec to identify fixations, as this default fixation filter is recommended for mobile eye-tracking data in the Tobii Lab Analyzer Software (Tobii AB, 2024). Two types of AOIs were defined to structure the data meaningfully: global AOI and event-based AOI. Global AOI were used to assess gaze behavior over the entire video duration and included the following: the *students*, representing gaze points focused on the group of three students; *teacher material*, capturing gaze points directed toward instructional materials such as the board, screen, or other teaching aids; *student desks*, representing gaze points on elements related to students’ desks, including name tags and student materials; and *classroom/others*, which encompassed gaze points directed toward other areas of the classroom that were not associated with the students or teacher. In contrast, the event-based AOI was coded only during specific classroom events, i.e., disruptions caused by students. The AOI labeled *disruptive student* captured gaze behavior directed toward one of the three students who performed a classroom disruption.

#### Average Fixation Duration on AOI. (Level 4)

The average fixation duration was calculated by dividing the total fixation duration (in milliseconds) by the total number of fixations. Fixation data for these calculations were aggregated across all AOI and summarized for both experienced and inexperienced teachers.

#### Fixation Number per Minute on AOI *Students*. (Level 4)

The fixation number per minute was calculated by summing all fixations within the AOI *students* and dividing by the session duration in minutes. This yielded an individual fixation rate per participant, which was then summarized for experienced and inexperienced teachers.

#### Gaze Relational Index. (Level 4)

The Gaze Relational Index (GRI) was calculated as a composite metric to evaluate the efficiency of participants’ scanning behavior during the micro-teaching unit. The GRI was derived by dividing the average fixation duration (in milliseconds) by the fixation number per minute.

#### Time to First Fixation on AOI *Disruptive Student*. (Level 4)

The time to first fixation (TTFF) measured the time (in seconds) for participants to fixate on the AOI *disruptive student* after a disruption onset. Values were extracted by identifying the first fixation timestamp relative to the disruption onset, excluding invalid cases (TTFF = 0[[1]](#footnote-2) or >30s[[2]](#footnote-3)). Raw values were converted from milliseconds to seconds and log-transformed for normalization. To assess overall responsiveness to classroom disruptions, log-transformed TTFF values were averaged across all disruptions per participant and then summarized for experienced and inexperienced teachers.

### Self-Evaluated Classroom Management (Level 3)

After the micro-teaching unit, teachers completed a questionnaire comprising five items from a validated instrument (Helmke et al., 2013) and eleven self-developed items derived from the research literature (Kiel et al., 2013; Kounin, 2006; Marzano, 2007) to assess their self-evaluated classroom management. The questionnaire was a 4-point Likert scale (1 = strongly disagree; 4 = strongly agree).

### Disruptiveness and Confidence Ratings (Level 3)

The disruptiveness and confidence ratings were assessed during the SRI on an 11-point rating scale, ranging from 0 (not at all disrupting/confident) to 10 (extremely disrupting/confident). Ratings were averaged across the nine classroom disruptions for each participant to capture a general sense of how disruptive the classroom disruptions were during the micro-teaching unit and how confident participants felt in handling them.

### Strategic Knowledge of Classroom Management (Level 3)

Teachers’ strategic knowledge of classroom management was assessed using a Situational Judgment Test (SJT; Gold & Holodynski, 2015) via an online questionnaire on SoSci Survey. Participants graded five to six action alternatives for twelve teaching scenarios in which classroom disruptions were discussed on a six-point Likert scale (grade 1 = “very good” to grade 6 = “unsatisfactory”). As the SJT was originally designed for primary schools, adjustments were made to enable the use of the SJT for all types of schools in the *ProVisioNET* study. Due to their general applicability, all twelve scenarios and answer options were adopted and only the names of the class levels were removed from the questions - except for scenario 6, where this information was essential.

## Data analysis (Level 2)

The data were analyzed using R (RStudio Team, 2020, Version 2024.12.0) and IBM SPSS Statistics (Version 29). Graphics were created using ggplot2 (Wickham, 2016).

To examine teachers’ visual attention distribution (Aim 1), fixation durations[[3]](#footnote-4) for different AOI (e.g., students, disruptive student, teacher material) were analyzed descriptively. Mean proportions were compared between experienced and inexperienced teachers using independent-sample *t*-tests, with effect sizes reported as Cohen’s *d* (Cohen, 1988).

To assess differences in gaze behavior related to noticing abilities (Aim 2), we calculated the GRI during the micro-teaching unit, fixation number per minuteon AOI *students*, and TTFF on AOI *disruptive student*. Group differences were tested with t-tests, and effect sizes were reported.

To investigate how disruption type affects noticing speed (Aim 3), a 2 × 3 repeated-measures ANOVA was conducted, with expertise (experienced vs. inexperienced teachers) as a between-subject factor and disruption type (verbal, physical, lack of eagerness) as a within-subject factor. Bonferroni-adjusted post-hoc comparisons were performed for significant effects.

To examine classroom management differences (Aim 4), independent-sample *t*-tests compared self-evaluated classroom management, disruptiveness and confidence ratings, and strategic knowledge.

Finally, to explore the relationship between gaze behavior and classroom management (Aim 5), Pearson correlations were computed between gaze efficiency metrics (GRI, fixation number per minuteon AOI *students*, TTFF on AOI *disruptive student*) and classroom management measures. Statistical significance was assessed using *p*-values, with effect sizes reported.

# Results (Level 1)

## Gaze Distribution for Areas of Interest (Level 2)

The descriptive results for experienced and inexperienced teachers to investigate teachers’ gaze distribution are displayed in Figure 1.

**Figure 1**

*Average Fixation Duration Percentages by Area of Interest (AOI) and Teacher Experience Group for the Entire Micro-Teaching Unit*

******

*Note.* The bar chart displays the percentage of fixation durations for five AOI (students, teacher material, student desk, classroom/others, and disruptive student) during the micro-teaching unit. Data are presented separately for experienced and inexperienced teachers, with error bars indicating 95% confidence intervals.

As suggested, both groups exhibited the highest percentages of fixation durations in the AOI *students*, while no statistically significant differences were observed between experienced and inexperienced teachers for any AOI (see Appendix C, Table C2 for *t*-test results and effect sizes).

## Gaze Behavior (Level 2)

Means, standard deviations, and range of experienced and inexperienced teachers’ gaze behavior measures are shown in Table 1.

**Table 1**

*Means, Standard Deviations, And Range of Experienced and Inexperienced Teachers’ Gaze Behavior Measures*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Gaze Behavior Measures | Experienced | | | | Inexperienced | | | |
| *M* | *SD* | Min | Max | *M* | *SD* | Min | Max |
| GRI for micro-teaching unit | 5.32 | 2.21 | 2.22 | 11.72 | 6.08 | 2.15 | 2.14 | 10.80 |
| Fixation Numbers in AOI *Students* | 98.53 | 19.21 | 49.07 | 135.81 | 91.21 | 14.52 | 68.99 | 123.37 |
| TTFF (log) in AOI *Disruptive Student* | 0.28 | 0.69 | –1.81 | 1.51 | 0.30 | 0.65 | –1.06 | 2.21 |
| *Note*. GRI = Gaze Relational Index, AOI = Area of Interest, TTFF = Time to First Fixation (log-transformed). Values represent the mean (*M*), standard deviation (*SD*), and range (minimum and maximum values) for each measure. Fixation numbers in AOI *Students* refer to the total number of fixations directed at students per minute. TTFF values represent the log-transformed time (in seconds) until teachers first fixated on a disruptive student. | | | | | | | | |

First, experienced teachers exhibited more frequent but shorter fixations, leading descriptively to a lower GRI than inexperienced teachers across the entire micro-teaching session (**Hypothesis 1a**). However, this difference was not statistically significant, *t*(80) = –1.57, *p* = .12.

Second, experienced teachers directed their gaze toward AOI *students* more often than inexperienced teachers (**Hypothesis 1b**), and this difference was statistically significant, *t*(80) = 1.96, *p* = .05, *d* = 0.43 (small effect).

Third, experienced teachers detected disruptions descriptively slightly faster, as reflected in their shorter noticing speed for the AOI *disruptive student*[[4]](#footnote-5) (**Hypothesis 1c**), but with no statistically significant difference, *t*(80) = –0.14, *p* = .89.

## Effect of Disruption Type and Expertise on Disruption Noticing Speed (Level 2)

Means, standard deviations, and ranges of experienced and inexperienced teachers’ log-transformed time to first fixation by disruption type are shown in Table 2.

**Table 2**

*Descriptive Statistics for Log-Transformed Time to First Fixation by Disruption Type and Group*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Disruption type | Experienced | | | | Inexperienced | | | |
| *M* | *SD* | Min | Max | *M* | *SD* | Min | Max |
| Lack of eagerness to learn | 1.10 | 1.58 | –4.61 | 3.34 | 1.21 | 1.63 | –3.91 | 3.34 |
| Physical disruptions | 0.26 | 1.54 | –3.91 | 3.21 | 0.14 | 1.44 | –4.61 | 3.25 |
| Verbal disruptions | –0.58 | 1.39 | –4.61 | 2.94 | –0.76 | 1.28 | –4.61 | 2.82 |
| *Note*. Values represent the mean (*M*), standard deviation (*SD*), and range (minimum and maximum values) for the log-transformed time to first fixation (TTFF) in seconds until teachers first fixated on a disruptive student. | | | | | | | | |

For **Hypothesis 2a**, a significant main effect of disruption type was found, *F*(1.94, 141.49) = 68.05, *p* < .05, *η²* = .34, indicating that the type of disruption influenced how quickly teachers noticed it. Consistent with our hypothesis, post-hoc comparisons revealed that verbal disruptions were detected significantly faster than both physical disruptions, *t*(73) = 6.33, *p* < .05, *d* = 0.62 (medium effect), and lack of eagerness disruptions, *t*(73) = 11.09, *p* < .05, *d* = 1.23 (large effect). Furthermore, physical disruptions were noticed faster than lack of eagerness disruptions, *t*(73) = -5.72, *p* < .05, *d* = 0.62 (medium effect). The main effect of expertise, *F*(1, 73) = 0.03, *p* = .86, and the interaction between expertise and disruption type, *F*(1.94, 141.49) = 1.28, *p* = .28, were not significant, suggesting that expertise did not influence noticing speed, nor did it interact with disruption type.

## Expertise Differences in Classroom Management Measures (Level 2)

Table 3 presents descriptive statistics and internal consistency reliabilities for all classroom management measures.

**Table 3**

*Means, Standard Deviations, Range, and Internal Consistency Reliability (McDonalds’ Omega, ω) of Experienced and Inexperienced Teachers’ Classroom Management Measures*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Classroom Management Measures | Experienced | | | | Inexperienced | | | | ω |
| *M* | *SD* | Min | Max | *M* | *SD* | Min | Max |
| Self-evaluated Classroom Management | 3.25 | 0.31 | 2.38 | 3.88 | 3.06 | 0.30 | 2.56 | 3.62 | .82 |
| Disruptiveness Rating | 4.81 | 2.90 | 0 | 10 | 5.55 | 2.81 | 0 | 10 | .84 |
| Confidence Rating | 8.44 | 1.68 | 2 | 10 | 7.18 | 2.04 | 0 | 10 | .88 |
| Strategic Knowledge | 0.75 | 0.08 | 0.54 | 0.88 | 0.72 | 0.13 | 0.10 | 0.91 | .89 |
| *Note*. Values represent the mean (*M*), standard deviation (*SD*), and range (minimum and maximum values), along with McDonald’s Omega (ω) coefficient, which indicates the internal consistency reliability for each measure. *Self-evaluated classroom management* refers to participants’ assessment of their classroom management competencies. The *disruptiveness rating* indicates the extent to which participants judged the disruptions as disruptive, while the c*onfidence rating* measures their confidence in managing them. S*trategic knowledge* measures teachers’ knowledge of classroom management strategies. | | | | | | | | | |

Independent-sample *t*-tests examined whether teaching expertise influenced classroom management measures (**Hypothesis 3a**). As expected, experienced teachers evaluated their classroom management competencies significantly higher than inexperienced teachers, *t*(80) = 2.78, *p* < .05, *d* = 0.62 (medium effect). They also reported disruptions as less disruptive, *t*(80) = –2.57, *p* < .05, *d* = –0.57 (medium effect), and greater confidence in managing them, *t*(80) = 5.63, *p* < .05, *d* = 1.24 (large effect). However, strategic knowledge of classroom management did not differ significantly between groups, *t*(80) = 1.00, *p* = .32, suggesting expertise was not associated with higher scores on the situational judgment test (SJT).

## Correlation Between Gaze Behavior Measures and Classroom Management Measures (Level 2)

Table 3 presents Pearson correlation coefficients between gaze behavior measures and classroom management measures, separately for experienced (below the diagonal) and inexperienced teachers (above the diagonal).[[5]](#footnote-6)

**Table 3**

*Correlations Between the Gaze Efficiency Measure and Classroom Management Measures for Experienced (Below Diagonal) and Inexperienced (Above Diagonal) Teachers*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| (1) GRI | – | –.44\* | .11 | –.01 | .16 | .15 | –.14 |
| (2) Fixation Number Per Minute on AOI *Students* | –.53\* | – | –.39\* | .39\* | –.49\* | –.22 | .14 |
| (3) TTFF (log) on AOI *Disruptive Student* | .05 | –.29 | – | –.20 | .32\* | –.01 | .01 |
| (4) Disruptiveness Rating | .03 | –.20 | .02 | – | –.02 | .01 | –.15 |
| (5) Confidence Rating | 0 | .06 | –.06 | –.51\* | – | .48\* | –.19 |
| (6) Self-evaluated Classroom Management | .11 | 0 | –.18 | –.11 | .33\* | – | –.21 |
| (7) Strategic Knowledge | –.36\* | .16 | –.34\* | –.19 | –.03 | .12 | – |
| *Note.* The table presents Pearson correlation coefficients between gaze behavior measures (GRI = Gaze Relational Index, fixation number per minute on AOI s*tudents,* and TTFF (= log-transformed time to first fixation on AOI d*isruptive student*), and classroom management measures (self-evaluated classroom management, disruptiveness and confidence ratings, and strategic knowledge) for inexperienced teachers (above diagonal) and experienced teachers (below diagonal).  Statistically significant correlations (*p* < .05) are marked with an asterisk (\*).  GRI = Gaze-Relational Index; TTFF = Time to First Fixation on AOI d*isruptive student* (log-transformed). | | | | | | | |

For experienced teachers, lower GRI values (indicating more efficient gaze behavior) were significantly associated with higher fixation frequency on students and greater strategic knowledge. Additionally, more frequent fixations on students correlated with higher confidence in handling disruptions. Longer TTFF on disruptive students was linked to lower strategic knowledge.

For inexperienced teachers, greater fixation numbers on students were negatively correlated with self-evaluated classroom management and confidence ratings. Additionally, higher confidence in handling disruptions was positively associated with self-evaluated classroom management.

# Discussion (Level 1)

## Key Findings (Level 2)

Our study investigated …

Overall, our findings indicate that …

Our findings are consistent with prior research that illustrates the …

## Limitations and future directions (Level 2)

While the laboratory setting of the study allowed for a controlled implementation of stressors and high internal validity, it was not an authentic classroom environment, raising questions about its external validity. Most importantly, the teacher and their students did not have a shared history, and only a very thin basis for establishing a positive teacher-student relationship, which is a core characteristic of effective classroom management (Beaty-O’Ferrall et al., 2010; Rüedi, 2014).

In addition, the micro-teaching unit was only about 15 minutes long, and thus much shorter than a regular school lesson, providing less opportunities for experienced teachers to build up an engaging lesson.

Finally, the onset of disruptive student behavior was scripted, following an experimental time schedule, which was not affected by the behavior of the teacher. Thus, the setting may have masked effects of teaching experience by providing too little opportunities of experienced teachers to demonstrate their true classroom management skills.

In subsequent studies, it would therefore be insightful to …

## Conclusion (Level 2)

This study investigated …

In summary, our study contributes to the understanding …

# References

Beaty-O’Ferrall, M. E., Green, A., & Hanna, F. (2010). Classroom Management Strategies for Difficult Students: Promoting Change through Relationships. *Middle School Journal (J1)*, *41*(4), 4–11.

Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift Für Psychologie*, *223*(1), 3–13. https://doi.org/10.1027/2151-2604/a000194

Braun, S. S., Kaihoi, C. A., McDaniel, H. L., & Bradshaw, C. P. (2022). Profiles of teachers’ occupational health: Associations with classroom management practices, gender, and race. *Teaching and Teacher Education*, *118*, 103819. https://doi.org/10.1016/j.tate.2022.103819

Cohen, J. (1988). Statistical power for the behavioural sciences. Hilsdale. *NY: Lawrence Erlbaum*, *58*(1), 7–19.

Dicke, T., Elling, J., Schmeck, A., & Leutner, D. (2015). Reducing reality shock: The effects of classroom management skills training on beginning teachers. *Teaching and Teacher Education*, *48*, 1–12. https://doi.org/10.1016/j.tate.2015.01.013

Gold, B., & Holodynski, M. (2015). Development and Construct Validation of a Situational Judgment Test of Strategic Knowledge of Classroom Management in Elementary Schools. *Educational Assessment*, *20*(3), 226–248. https://doi.org/10.1080/10627197.2015.1062087

Grub, A.-S. (2023). *PRONOEA - Professional vision of novice and expert teachers* [doctoralThesis, Saarländische Universitäts- und Landesbibliothek]. https://doi.org/10.22028/D291-39788

Helmke, A., Schrader, F.-W., Helmke, T., Lenske, G., Pham, G., Praetorius, A.-K., & Ade-Thurow, M. (2013). *Basisfragebogen EMU - Evidenzbasierte Methoden der Unterrichtsentwicklung*. Unterrichtsdiagnostik. http://www.unterrichtsdiagnostik.info/downloads/fragebogen/

Keskin, Ö., Gabel, S., Kollar, I., & Gegenfurtner, A. (2023). Relations between pre-service teacher gaze, teacher attitude, and student ethnicity. *Frontiers in Education*, *8*. https://doi.org/10.3389/feduc.2023.1272671

Kiel, E., Frey, A., & Weiß, S. (2013). *Trainingsbuch Klassenführung*. Verlag Julius Klinkhardt.

Kounin, J. S. (2006). *Techniken der Klassenführung*. Waxmann Verlag.

Lohmann, G., & Meyer, H. (2003). *Mit Schülern klarkommen: Professioneller Umgang mit Unterrichtsstörungen und Disziplinkonflikten*. Cornelsen-Scriptor.

Marzano, R. J. (2007). *The Art and Science of Teaching: A Comprehensive Framework for Effective Instruction*. ASCD.

Onkhar, V., Dodou, D., & de Winter, J. C. F. (2024). Evaluating the Tobii Pro Glasses 2 and 3 in static and dynamic conditions. *Behavior Research Methods*, *56*(5), 4221–4238. https://doi.org/10.3758/s13428-023-02173-7

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. RStudio, PBC.

Rüedi, J. (2014). Zur Bedeutung positive Beziehungen für die Klassenführung und den Umgang mit Unterrichtsstörungen. *Beziehungen in Schule Und Unterricht. Teil*, *3*, 105–126.

Sherin, M. G. (2007). The Development of Teachers’ Professional Vision in Video Clubs. In *Video Research in the Learning Sciences*. Routledge.

Tobii AB. (2024). *Tobii Pro Lab User Manual v 24.21*. https://go.tobii.com/tobii\_pro\_lab\_user\_manual

Tucholka, I., & Gold, B. (2025). Analysing classroom videos in teacher education—How different instructional settings promote student teachers’ professional vision of classroom management. *Learning and Instruction*, *97*, 102084. https://doi.org/10.1016/j.learninstruc.2025.102084

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://doi.org/10.1007/978-0-387-98141-3

Yamamoto, T., & Imai-Matsumura, K. (2013). Teachers’ Gaze and Awareness of Students’ Behavior: Using An Eye Tracker. *Comprehensive Psychology*, *2*, 01.IT.2.6. https://doi.org/10.2466/01.IT.2.6

# Appendix

# Appendix A

**Eye-tracking apparatus and calibration**

Teachers wore a binocular Tobii Pro Glasses 2 eye-tracker during the micro-teaching unit to record eye-tracking data. The system consisted of a wearable head unit and a recording unit. The head unit was a measuring device with different sensors. A high-definition scene camera captured a full HD video of the teacher’s field of vision. An integrated microphone recorded surrounding sounds. Infrared light illuminators supported the eye-tracking sensors which recorded the eye orientation to capture the teacher’s gaze point. The videos were recorded with a sampling rate of 50 Hz in a video resolution of 1920 x 1080 at 25 frames per second. The scene camera had a field of view of 90 degrees in 16:9 format (82 degrees horizontal and 52 degrees vertical) and a frame dimension of 179 x 159 x 57 mm (width x depth x height). The recording unit is a compact computer that manages the head unit. It captures and saves eye-tracking data, audio, and scene camera footage on a removable SD memory card.

The evaluation of the calibration process followed the guidelines outlined in the Manual of Tobii AB (2024) and Onkhar et al. (2024) for assessing calibration quality. Participants’ gaze was calibrated using a bullseye card that the participant held at arm’s length. A successful calibration was achieved when the participant’s gaze marker sufficiently overlapped with the bullseye for a specified time, based on criteria internally determined by the manufacturer’s software (Tobii AB, 2024). All participants achieved successful calibration, and no participants were excluded due to calibration failure. The robustness of the calibration was further verified through a secondary nine-point calibration. During this step, participants were asked to read numbers from one to nine aloud and direct their gaze at specific fields corresponding to each number. The initial calibration was performed before and the verification calibration before and after each micro-teaching unit.

**Table A1**

*Classification of Nine Typical Classroom Disruptions According to Lohmann & Meyer (2003) Performed in The Micro-Teaching Unit by Actors*

|  |  |  |
| --- | --- | --- |
| Verbal disruptions | Physical disruptions | Lack of eagerness to learn |
| Heckling | Clicking pen | Looking at phone |
| Chatting | Snipping hands | Drawing |
| Whispering | Drumming hands | Head on table |

*Note.* Disruptions were classified based on the typology provided by Lohmann & Meyer (2003)*.* Categories include verbal, physical, and disengagement-related behaviors performed during the micro-teaching unit. The order of the performing actors and the disruptions was fully balanced using Latin squares.

# Appendix B

# Laboratory Setting of The Study

**Figure B1**

*Laboratory Setting of The Micro-Teaching Unit. Ein Bild, das Mobiliar, Stuhl, Kleidung, Schuhwerk enthält.

Automatisch generierte Beschreibung*

*Note*. The setting included three actors as the class (left) and a teacher (participant, right).

**Figure B2**

*Laboratory Setting of The Interview.*

Ein Bild, das Mobiliar, Zeichnung, Entwurf, Tisch enthält.

Automatisch generierte Beschreibung

*Note*. The experimenter and participant watched the previously taught micro-teaching unit on video.

# Appendix C

**Additional Results for Gaze Behavior Across AOI**

**Figure C1**

*Average Fixation Number Percentages by Area of Interest (AOI) and Teacher Experience Group for the Entire Micro-Teaching Unit with 95% Confidence Intervals*



*Note.* The bar chart illustrates the average fixation number percentage directed at five AOIs (Students, Teacher Material, Classroom/Others, Student Desk, and Disruptive Student) during the micro-teaching unit. Results are presented separately for experienced and inexperienced teachers, with error bars indicating 95% confidence intervals.

**Table C2**

t*-Test Results and Effect Sizes for Fixation Number Percentages (FNP) and Fixation Duration Percentages (FDP) Across AOIs Between Experienced and Inexperienced Teachers*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Measures | AOI | *t-*value | df | *p*-value | *M* Difference | Cohen's *d* | 95% CI |
| FNP a | Students | 0.01 | 80 | 0.99 | 0.03 | 0.00 | [–4.63, 4.69] |
| FNP | Disruptive Student | –1.32 | 80 | 0.19 | –0.27 | –0.29 | [–0.67, 0.14] |
| FNP | Teacher Material | 0.89 | 80 | 0.37 | 2.32 | 0.20 | [–2.84, 7.48] |
| FNP | Student Desk | –1.17 | 80 | 0.25 | –1.72 | –0.26 | [–4.65, 1.21] |
| FNP | Classroom/Others | –0.43 | 80 | 0.67 | –0.63 | –0.10 | [–3.53, 2.27] |
| FDP b | Students | –0.43 | 80 | 0.67 | –1.26 | –0.10 | [–7.1, 4.58] |
| FDP | Disruptive Student | –1.90 | 80 | 0.06 | –0.48 | –0.42 | [–0.98, 0.02] |
| FDP | Teacher Material | 1.55 | 80 | 0.13 | 4.45 | 0.34 | [–1.28, 10.18] |
| FDP | Student Desk | –1.67 | 80 | 0.1 | –2.89 | –0.37 | [–6.35, 0.56] |
| FDP | Classroom/Others | –0.28 | 80 | 0.78 | –0.30 | –0.06 | [–2.43, 1.84] |
| *Note*. This figure displays the results of *t*-tests and effect sizes (Cohen’s *d*) for fixation number percentages (FNP) and fixation duration percentages (FDP) across Areas of Interest (AOIs) between experienced and inexperienced teachers. AOIs included Students, Disruptive Student, Teacher Material, Student Desk, and Classroom/Others. Positive mean differences (*M* Difference) indicate higher percentages for experienced teachers. Confidence intervals (95%) for the mean differences are presented in brackets.  *p*-value < .05 is considered statistically significant.  a FNP = Fixation Number Percentages  b FDP = Fixation Duration Percentages | | | | | | | |

**Appendix D**

**Correlation Between Gaze Behavior Measures and Classroom Management Measures**

**Table D1**

*Correlations Between the Gaze Behavior Measures and Classroom Management Measures Across All Participants*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Variable | (2) | (3) | (4) | (5) | (6) | (7) |
| (1) GRI | –.50\* | .08 | .06 | –.02 | .07 | –.23\* |
| (2) Fixation Number Per Minute on AOI *Students* | – | –.34\* | 0 | –.13 | –.06 | .16 |
| (3) TTFF (log) on AOI *Disruptive Student* | –.34\* | – | –.07 | .12 | –.01 | –.12 |
| (4) Disruptiveness Rating | 0 | –.07 | – | –.36\* | –.13 | –.03 |
| (5) Confidence Rating | –.13 | .12 | –.36\* | – | .49\* | –.06 |
| (6) Self-Evaluated Classroom Management | –.06 | –.10 | –.13 | .49\* | – | –.05 |
| (7) Strategic Knowledge | .16 | –.12 | –.03 | –.06 | –.05 | – |
| *Note.* The table presents Pearson correlation coefficients between gaze behavior measures (GRI = Gaze Relational Index, fixation number per minute on AOI *students*, and TTFF = log-transformed time to first fixation on AOI *disruptive student*) and classroom management measures (self-evaluated classroom management, disruptiveness and confidence ratings, and strategic knowledge) across all participants. Lower GRI values indicate more efficient gaze behavior. Statistically significant correlations (*p* < .05) are marked with an asterisk (\*). | | | | | | |

1. As our focus was on gaze behavior directed toward the disruptive student, instances where the gaze was already fixated on the disruptive student, when the event started, were excluded from the analysis. [↑](#footnote-ref-2)
2. Fixation times beyond 30 seconds were excluded, as they exceeded the duration of the scripted disruptions. [↑](#footnote-ref-3)
3. The percentage of fixation duration was used as the primary measure, as it offered a more intuitive visualization of attention distribution over time (see Appendix C, Figure C1 for the percentage for fixation numbers). [↑](#footnote-ref-4)
4. Before conducting the analysis, 21.33% of data points were excluded because the participant either did not fixate on the disruption or had fixation times of zero or longer than 30 seconds, exceeding the defined threshold (see Data Analysis). Removing them ensured a focus on meaningful and interpretable gaze detection times. [↑](#footnote-ref-5)
5. See Appendix D (Table D1) for a correlation table including all gaze behavior measures and classroom management measures for all participants. [↑](#footnote-ref-6)