Title: Novice and expert teachers' noticing of classroom management in whole-group and partner work activities: Evidence from teachers' gaze and identification of events

Abstract

This study investigates how novice and expert teachers' noticing of classroom management events differs in two formats of instruction. 20 novices and 20 experts participated in the study, watching short video clips of whole-group and partner work teaching situations. Their retrospective verbal reports were analyzed for events identified as note-worthy along with their allocation of visual attention as indicators of their noticing. Experts noticed more classroom management events in the partner work format than novices. Furthermore, their noticing was characterized by a focus on student-related events. Similarly, their gaze prioritized students more than novices', particularly in the partner work format. In contrast, novice teachers' attention was more drawn to the teacher in both formats of instruction. The results show that expertise in teachers' noticing of classroom management is characterized by a focus on students with the partner work format being more challenging for novice teachers.

Keywords: Teacher Expertise, Classroom Management, Professional Noticing, Visual Expertise, Eye-Tracking

1 Introduction*

Teaching is a domain that is characterized by multidimensionality, simultaneity and immediacy, thus teachers must respond quickly to various demands in the classroom (Doyle, 2006; Sabers et al., 1991). In order to react adaptively in a teaching situation, they do not only need relevant knowledge but also situated skills to be able to transform their knowledge into practice (Blömeke et al., 2015). One important situated skill is teacher noticing, which is their knowledge-based ability to selectively attend to and to notice relevant events in a classroom situation (Star & Strickland, 2008; van Es & Sherin, 2002). Many studies about teachers' content-related noticing revealed insights into teachers' skills, in particular, remarkable differences between novice and expert teachers noticing (Stahnke et al., 2016).

Yet, concerning classroom management (CM) as an important generic aspect of teaching, recent research on teachers' noticing is less comprehensive. Being able to notice relevant events during instruction is particularly relevant with regard to CM, because it poses situated and spontaneous challenges to teachers that call for immediate action (Doyle, 1986, 2006). Initial results indicate a key role of such skills, as they seem to predict teachers' CM performance better than their pedagogical knowledge about CM (König & Kramer, 2016).

To identify characteristics of expertise, novice-expert comparisons are an established research approach in several domains (Chi, 2006), often also investigating participants' allocation of visual attention to relevant areas (Gegenfurtner et al., 2011). Although such comparisons have the potential to provide new insights into the characteristics of teachers' skills and their development, this approach has only recently been applied to teachers' noticing of CM. Analyzing teachers' visual attention to and verbal analysis of classroom video clips, these recent studies yielded partly contradictory results (van den Bogert et al., 2014; Wolff et al., 2016; Yamamoto & Imai-Matsumura, 2013). Prior research focused particularly on behavioral management (i.e. preventing and dealing with student misbehavior) in sequences of whole-group instruction. While this is a core aspect of teachers' CM practice, for other dimensions of CM - e. g. instructional management or teacher-student relationships - as well as other formats of instruction with distinct demands on teachers' CM (Doyle, 2006) further research is needed.

* Abbreviations: CM = Classroom management; AOI = Areas of interest; MR = Mean rank

Against this background, the objective of our study is to investigate novice and expert teachers' noticing of CM events in video clips displaying two different instructional formats (whole-group instruction and partner work) and including CM events that go beyond behavioral CM. Thereby, we want to provide insights into format-specific expertise differences in teachers' noticing, thus expanding the current state of research focusing on behavioral management and whole-group formats with regard to teachers' noticing of diverse CM events in different formats of instruction.

2 Theory

2.1 Novice and expert teachers' noticing of classroom events

One way to learn more about the nature and development of skills are expert-novice comparisons (Chi, 2006). Early studies on general characteristics of teacher expertise showed that expert teachers can deal better with the simultaneity, multidimensionality and immediacy that characterizes a classroom than novice teachers (Sabers et al., 1991). They are able to monitor events, to integrate information fast and to interpret what happens in a classroom (Carter et al., 1988; Copeland et al., 1994; Sabers et al., 1991). Overall, Berliner (2001, 2004) described expert teachers as having a better perception of classroom events and having faster and more accurate recognition in their domain than novices. Recently, studies started to use eye tracking methods in order to analyze novice and expert teacher general allocation of attention in the classroom (Beach & McConnel, 2019; Jarodzka et al., 2017). With regard to teachers' visual attention during teaching (as measured with eye tracking glasses), experts prioritized students with their gaze (McIntyre et al., 2019; McIntyre & Foulsham, 2018) and distributed their attention more evenly between individual students than novice teachers (Cortina et al., 2015). To what extent these results also apply to CM events, in particular to different formats of instruction, is not clear yet.

Superior performance of experts may not only be based on teachers' knowledge, but also on their situations-specific skills (Blömeke et al., 2015; Lachner et al., 2016). During the development from novice to expert through deliberate practice, the initially isolated and explicit knowledge base of novices is restructured and evolves towards more integrated and organized scripts (Boshuizen et al., 2020; Lachner et al., 2016; Wolff et al., 2020). Teachers' noticing, reasoning and acting skills help to apply these scripts to situations that are not pre-structured (Lachner et al., 2016). Thus more generally speaking, the three skills of perception, interpretation and decision-making play an important role when teachers' knowledge needs to be put into performance in a specific classroom situation (Blömeke et al., 2015). The skills are nevertheless knowledge-based as teachers' knowledge and scripts guide their noticing or interpretation of important events as well as their decisions on how to act (Lachner et al., 2016).

The present study particularly focusses on teachers' *noticing* as the first of these three skills which is defined as attending to and identifying what is important in a classroom situation (Star & Strickland, 2008; van Es & Sherin, 2002). One important aspect of teachers' noticing is their visual

attention to relevant areas of a classroom scene, as this allocation of attention is necessary in order to be able to identify important aspects of a scene. Thus, analyzing teachers' gaze as an operationalization of teachers' noticing in addition to their identification of note-worthy events is a promising approach (Seidel et al., 2020).

2.2 Classroom management

CM is an important indicator of instructional quality (Charalambous & Praetorius, 2018) and has positive effects on students' academic, social and emotional learning (Korpershoek et al., 2016). It has also an impact on the well-being of teachers, as CM is one of the most common concerns of pre-service (Kaufman & Moss, 2010) as well as beginning teachers (Chaplain, 2008).

Since Kounin's early studies (1970) it has been evident that a good classroom manager not only reacts appropriately to student misbehavior or disengagement but also prevents such behavior from occurring or spreading (Bear, 2015; Brophy, 1986; Doyle, 2006; Kounin, 1970). Expanding this focus on *student discipline* with *student learning*, Martin and Sass (2010) proposed two dimensions of CM: Behavioral management includes teachers' reaction to student misbehavior and efforts to prevent it; instructional management focusses on teachers' instructional aims and methodologies. Furthermore, recent research proposed that more attention should be given to affective-motivational learning (e. g. motivation of students or teacher-student relationships) (Bear, 2015; Schwab & Elias, 2015; Wubbels et al., 2015). Thus, in this study we apply such a comprehensive understanding of CM as the "actions teachers take to create an environment that supports and facilitates both academic and social-emotional learning" (Evertson & Weinstein, 2006, p. 4).

Teachers' CM behavior is often classified as either *reactive* (following misbehavior or disengagement) or *preventive* (preventing misbehavior and supporting student learning with e. g. routines, monitoring, or building of student-teacher relationships) (Clunies-Ross et al., 2008; Piwowar et al., 2013). Both types of strategies are essential for managing a classroom (Korpershoek et al., 2016; Simonsen et al., 2008). Yet, novices seem to be mostly concerned about student discipline and behavior control (Kaufman & Moss, 2010) and report to use reactive strategies more frequently than preventive strategies (Reupert & Woodcock, 2010). However, reactive CM is correlated with higher teacher stress as well as a decrease in students' on-task behavior (Clunies-Ross et al., 2008).

Being able to apply a broad repertoire of CM strategies adaptively is particularly important because different formats of instruction pose different challenges (Doyle, 2006). In whole-group instruction, the teacher needs to monitor the flow of the lesson as well as student learning and student behavior on the group level. In contrast, during periods of partner work or small group work the teacher has to observe many individual students or student groups, determine their learning progress and be available for individual student questions (Doyle, 2006). Thus, these formats can place high demands on teachers' noticing.

2.3 Teachers' Noticing of CM

Managing a classroom poses unpredictable and spontaneous challenges to teachers that call for immediate action (Doyle, 1986, 2006). Thus, noticing relevant CM events is a particularly important skill in this context. Addressing this area of research with standardized test instruments, studies revealed that teachers' situated skills were significantly positively related to teachers' level of expertise (Gold & Holodynski, 2015; König & Kramer, 2016). Furthermore, teachers' CM skills predicted their CM performance better than their pedagogical knowledge (König & Kramer, 2016). However, standardized test instruments lack the immediate and spontaneous character that makes noticing with regard to CM challenging for teachers (Doyle, 1986). Studies analyzing teachers' spontaneous noticing of CM events without directed questions along with their gaze address this challenge in an ecologically more valid way.

Concerning teachers' spontaneous noticing of events, expert teachers generally focus more on student learning while novice teachers talk more about student discipline (Wolff et al., 2015, 2017). However, the video clips used in these studies focused on behavioral CM and did not investigate how many and which specific events were noticed. On the contrary, analyses were limited to teachers' comments about those events that were frequently noticed by both novice and expert teachers (Wolff et al., 2015, 2017). Focusing on one event only, Yamamoto and Imai-Matsumura (2013) found no differences between novices' and experts' noticing of student misbehavior (two students not closing their textbook after being instructed to do so) in a video clip of whole-group instruction where the teacher was not visible.

Regarding teachers' gaze as one aspect of noticing, studies differed considerably in their methodology, yet also mainly focused on behavioral problems. In the study by Yamamoto & Imai-

Matsumura (2013), no expertise effect was found for teachers' gaze. In contrast, choosing video clips "representative of typical classroom behavior" that "require(s) teacher intervention" (p. 210), van den Bogert and colleagues (2014) found that expert teachers tended to distribute their visual attention more evenly across student groups in those segments of video clips where many CM-related events were frequently noticed by both groups. However, only few differences between both groups were found for segments where no events were happening or only experts noticed CM events. Wolff and colleagues (2016) compared novice and expert teachers' visual processing of CM in whole-group instruction without the teacher being visible. Experts attended more to areas showing students and classroom activity than novices. Wolff and colleagues (2020) concluded in a recent theoretical model on teachers' CM scripts that novice teachers' classroom perception is more image-driven (i.e. bottom-up processing) while experts' is more knowledge-driven (i.e. top-down processing), thus allowing them to direct their attention to informative areas. While novices consciously monitor classroom activity and engagement focusing on student behavior, experts monitor classroom activity automatically in terms of engagement and student learning based on their CM knowledge and scripts (Wolff et al., 2020). If such differences can also be observed beyond behavioral CM and in different formats of instruction has not been investigated yet.

Regarding the format of instruction, a recent study on teachers' diagnostic skills found more expertise effects in teachers' visual attention to different student profiles for a seatwork sequence in comparison to a whole-group sequence (Seidel et al., 2020). The authors argue that bottom-up drivers of visual attention are more salient in a whole-group setting where teacher-student interactions take place and students raise their hands. In seatwork scenes, such salient drivers are absent, thus allowing expert teachers' knowledge and scripts to guide their perception top-down (Seidel et al., 2020).

3 Research questions

The current study aims to expand the state of research on teachers' noticing of CM, which has so far focused on whole-group instruction and behavioral management. More specifically, this study investigates how novice and expert teachers' noticing of CM events differs regarding whole-group instruction and partner work in teaching situations that display events beyond behavioral management. Three aspects of teachers' noticing are of particular interest: their *identification* of note-worthy CM events, their visual *attention to student groups or the teacher* as well as their *visual attention to specific events*. Thus, the three research questions are:

RQ1: Do novice and expert teachers differ in their identification of note-worthy CM events in whole-group instruction vs. partner work?

Based on prior research, we assume experts, firstly, to generally notice more CM events than novices, particularly in the partner work scene as it is more demanding with respect to teachers' monitoring (Doyle, 2006) and allows top-down processing (Seidel et al., 2020). Secondly, novices are expected to notice more events relating to reactive CM and student discipline, while experts notice more events focusing on preventive CM and student learning (Kaufman & Moss, 2010; Wolff et al., 2020). Since partner work requires a broader range of CM strategies (Doyle, 2006; Reupert & Woodcock, 2010), we expect that these expertise differences show up more clearly in this format.

RQ2: Do novice and expert teachers differ in their gaze directed at student groups and the teacher in whole-group instruction vs. partner work?

Against the background of prior research, expertise effects are expected to be generally weaker for the whole-class setting than partner work, as visual processing should be more bottom-up in the first setting for both groups (Seidel et al., 2020). Nevertheless, in both formats novice teachers can be expected to pay more attention to the teacher than experts, who in turn can be assumed to attend more to students (McIntyre et al., 2019; Wolff et al., 2020).

RQ3: Do novice and expert teachers differ in their gaze directed at specific CM events in whole-group instruction vs. partner work?

Based on the few results available (Yamamoto & Imai-Matsumura, 2013), there are no differences expected between novices and expert teachers' gaze directed at specific CM events. As formulated in RQ2, expertise is expected to guide experts' attention to potentially relevant broader areas (i.e. student groups), yet not at the granular level of specific CM events.

4 Methods

4.1 Participants

40 German pre-service and in-service secondary school teachers voluntarily participated in this study. Although CM is considered to be generic, this study focused on mathematics and biology teachers to reduce the possible impact of teachers' familiarity with typical contents or formats of instruction. Participants were recruited via multiple channels (e. g. professional development networks, university courses). Novices were defined as pre-service teachers in their master studies who had no teaching experience beyond the short practice phases included in their teacher education program. Expertise was defined by professional membership and experience, thus following recommendations of teacher expertise research (Caspari-Sadeghi & König, 2018; Palmer et al., 2005): Experts were required to have at least five years of teaching experience after finishing their teacher education program. Furthermore, they had to be selected for additional responsibilities and tasks in their schools (e.g. head of the biology department) or in teacher education (e.g. supervision of preservice teachers) as indicators of an external evaluation of their outstanding quality.

Data collection took place where teachers could arrange their participation best (at the lab in the university, at schools or at teachers' homes). Expert teachers were on average 20 years older (M_{age}= 45.10, *SD* = 9.69; 15 female, 5 male) than novices (M_{age}= 26.70, *SD*= 3.79; 12 female, 8 male). Novices were on average in their final semester of master studies ($M_{semester}$= 3.35, *SD* = 0.90) for becoming secondary school teachers for biology (*N* = 10) or mathematics (*N* =10). Experts had on average 18 years of teaching experience *(*$M_{experience}$= 18.30; *SD* = 10.89) after their teacher education program for either biology (*N* = 9) or mathematics (*N* = 11).

4.2 Material

Participants watched four short video clips (between 1 and 2 minutes long) from authentic biology and mathematics lessons in lower secondary classes in Germany. These lessons were taught by beginning to intermediate teachers who were expected to show both successful as well as more critical CM actions, thus resulting in many different observable CM events in the video clips. Video clips were selected in four steps: (1) First, video clips with low audio or video quality were excluded. (2) Eight clips showing CM events that display multiple CM aspects were selected by the first author (e.g. teachers' management of misbehavior, transparency and clarity, routines, motivation of students,

or teacher-student relationships (cf. Piwowar et al., 2013 for rating scheme used). (3) These clips were rated by five experts from research and practice with respect to the observability of different CM aspects as well as general authenticity and typicality. The raters' expertise was in video-based research, CM research and teacher education. High ratings in authenticity and typicality should ensure that video clips were selected that do not feel staged and show representative situations of teachers' jobs regarding CM (Blömeke et al., 2015). (4) Finally, four segments were selected based on experts' agreement on the occurrence of events, authenticity and typicality as well as the final set displaying multiple aspects of CM.

For the purpose of this paper, data analysis will focus on those two video clips where more CM events are visually observable (as opposed to audible events). In both video clips, nearly all students of the class and the teacher are visible. One video clip shows a whole-group activity: The teacher guides the comparison of solutions for math fraction problems. Students are taking turns at presenting at the smartboard, while the rest of the class should listen but is rather loud. The second video clip shows a partner work activity: The teacher walks through the classroom while students are working on an assignment on osmosis. Some students are distracted and not working on the assignment. Subject knowledge is not necessary for understanding what is going on in both sequences.

4.3 Procedure

The full experiment took between 45 and 75 minutes. After participants signed consent and release forms, the Miles Test (Holmqvist et al., 2011; Miles, 1929) was used to determine each participant's ocular dominance. Participants had normal or corrected to normal vision. A test trial followed to familiarize participants with the eye tracking equipment and the retrospective reporting method. The eye tracker was calibrated to participants' eyes before each video clip (9-point calibration). The order of video clips was incompletely counterbalanced.

While viewing a video clip for the first time, participants' eye movements were recorded. Participants were instructed to push a button every time they noticed a CM event they considered to be relevant (cf. van den Bogert et al., 2014). During the first viewing, the video could not be paused because we were interested in teachers' spontaneous noticing of and visual attention to classroom events.

Immediately after the first viewing participants saw the video clip again enriched by their own prior eye movements. The video was paused at each timestamp and participants were instructed to report what they had noticed at this specific moment in the video clip. We chose retrospective reports instead of concurrent verbalizations due to the complexity of the task. However, teachers' initial thoughts were cued with the help of time stamps and by displaying their own prior eye movements. When concurrent verbalizations are not suitable, such a procedure can help to elicit verbalizations (Hyrskykari et al., 2008; van Gog et al., 2005).

Video clips were presented on a 20-inch display (1650x1050 pixels) using Experiment Center 3.7 (SensoMotoric Instruments, 2016b). A SMI RED-m eye tracker recorded participants' eye movements with a temporal resolution of 120 Hz. Retrospective reports were recorded with a camera attached to the screen.

4.4 Measures

4.4.1 Classroom management events noticed

Retrospective reports were transcribed verbatim and coded consecutively by the first author using MAXQDA (VERBI Software, 2017). Since participants often reported more than one event per time stamp, the number of events identified in their reports was analyzed. Coding started with a list of noticeable events based on the expert rating. New codes were added when events were noticed that had not been reported before. Once verbal reports of all participants were coded, all codes were checked for consistency and refined if necessary. Smaller interrelated events were integrated into one event if they represented the same episode while codes were differentiated if they represented distinct events in close temporal or spatial proximity. For the first video clip 26 noticeable events were identified, and 30 events for the second video clip.

Each event noticed was categorized as one of four event types based on prior research on CM and CMS: Events where the teacher is the actor were categorized as a reactive teacher event (TR events: the teacher is or should be reacting to student disengagement or misbehavior), or a preventive teacher event (TP events: the teacher is or should be preventing student disengagement or misbehavior or supporting learning). Events where one or more students are the actor are either a student discipline event (SD events: i.e. students are misbehaving or disengaged) or a student

learning event (SL events: i.e. students are learning or something is supporting or hindering their learning) (see appendix A for codes).

Two independent raters coded whether novices and experts did notice or did not notice these events for 10 % of all verbal reports. Interrater reliability was strong with κ_1= 0.81 (92.31%) for the first video clip and κ_2=0.87 (94.17%) for the second video clip (McHugh, 2012).

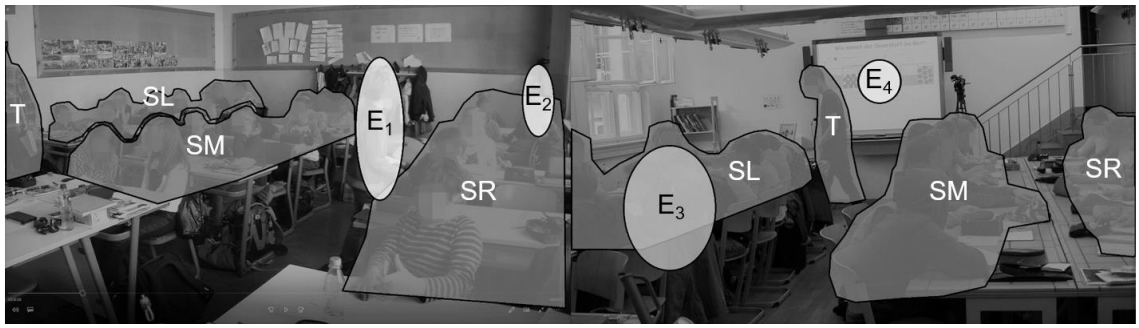
4.4.2 Teachers' visual attention to student groups and the teacher

Participants' eye movement data were analyzed with BeGaze 3.7 (SensoMotoric Instruments, 2016a) with regard to their proportion of gaze and fixation count on predefined areas of the classroom. Fixations, where the eye remains relatively motion-less (Holmqvist et al., 2011), were identified with a dispersion algorithm with a minimal duration of 80 ms and a maximum dispersion of 100 pixels.

To answer RQ2, proportions of gaze and fixation count were analyzed. For this purpose, areas of interest (AOIs) were defined for larger visually separated groups of students (three groups in each video on the right, in the middle and on the left side of the classroom) as well as the teacher (cf. figure 1: SL, SM, SR and T). The proportion of gaze represents a measure of participants' summarized dwell time (including fixations and quick scans) at an AOI relativized by the duration of the video clip. Areas with higher gaze proportions can be interpreted as more prioritized. Such measures have recently been used in teacher gaze studies (McIntyre et al., 2019). Similarly, we analyzed number of fixations on four AOIs as a second indicator of teachers' visual attention. A high number of fixations indicates that teachers' repeatedly allocated their attention to these areas (Holmqvist et al., 2011).

Figure 1

Areas of interest in the whole-group instruction (left) and the partner work format (right)



Note. SL = left student group, SM = middle student group, SR = right student group, T = teacher, E_1 = event 1 (student lingers and clowns around), E_2 = event 2 (student is raising hand and being ignored), E_3 = event 3 (two students fool around), E_4 = event 4 (timer on smartboard).

4.4.3 Teachers' visual attention to classroom management events

With respect to RQ3, we focused on events that were clearly linked to specific visual areas of the classroom and noticed by at least five teachers, because these events were thus identified as relevant to CM by a considerable proportion of teachers. Again, AOIs were created manually: For the first video clip, we identified two events (cf. figure 1): (a) a student lingers and clowns around while going back to his seat after presenting his solution (E_1); (b) a student in the back is raising her hand and is being ignored (E_2). For the second video clip, two events were identified as well: (c) two students fool around behind the back of the teacher (E_3); (d) a timer at the smart board shows the time remaining for the assignment (E_4). Teachers' proportion of gaze and fixation count on these AOIs were compared for novices and experts.

4.4.4 Statistical Analysis

Data analysis was conducted with SPSS 25. As the video clips were not controlled for complexity posed by the instructional format, effects of the level of expertise on dependent measures are investigated separately for whole-class instruction and partner work. Separate t-tests for independent samples or non-parametric Mann-Whitney tests were carried out. Dependent variables were inspected for outliers. There were no extreme outliers (data more than three interquartile ranges above the 75% quartile or under the 25% quartile). Shapiro-Wilk tests were carried out for all dependent variables within groups to find out whether the data were normally distributed. If a variable was not normally distributed, non-parametric tests were applied to examine our research questions[1]. If the homogeneity of variances was not given, adjusted values are reported.

An *a priori* power analysis was conducted using G*Power (Faul et al., 2007) with $\alpha = .05$ and power $(1-\beta) = .80$. There are only few studies available concerning expertise differences regarding

[1] Due to the relatively small sample size, non-parametric tests were also carried out for those dependent variables that fulfilled all preconditions for parametric testing to check the robustness of our results. The non-parametric tests revealed similar results. Thus, they are not reported additionally.

teachers' gaze to meaningful areas that could guide this analysis. A recent study on teachers' skills with a similar design reported moderate to large effect sizes for the number of fixations on different student groups (Seidel et al., 2020). However, it should be noted that no prior effect sizes are available for proportions of dwell times. Assuming similar effect sizes, the sufficient sample size for independent t-tests (d = 0.80) is N = 42. Our sample size of 40 teachers is slightly smaller yielding an acceptable post-hoc power of (1-β) = .799 for t-tests.

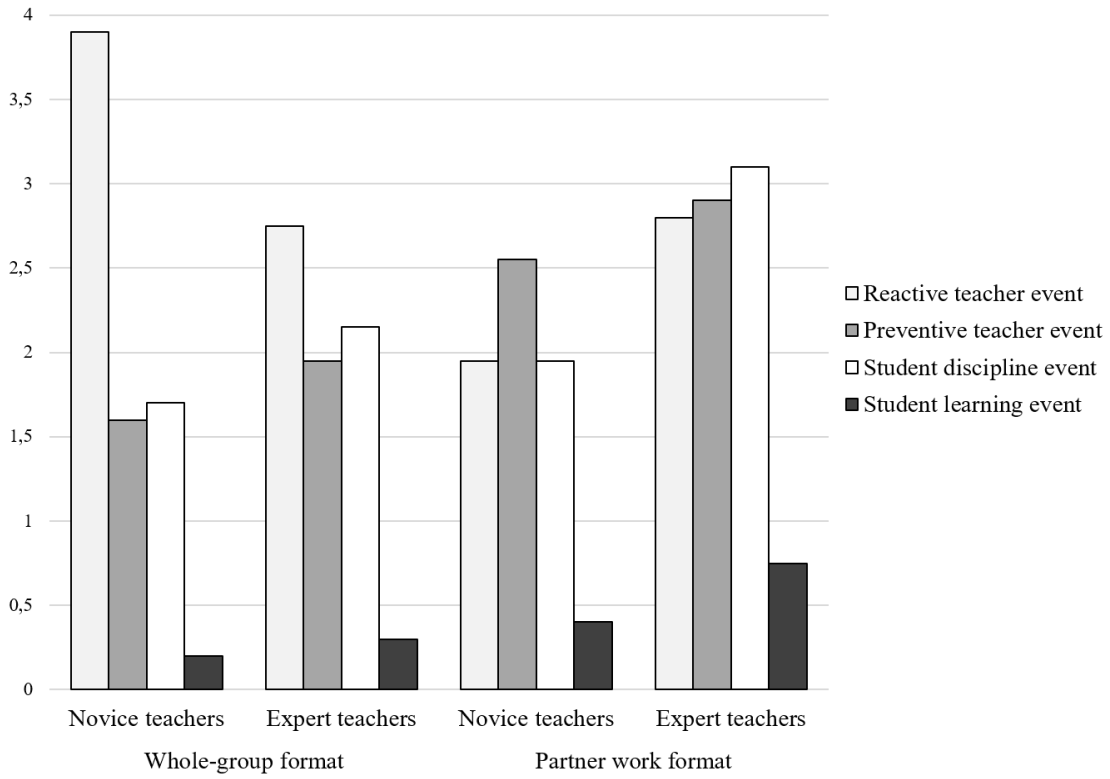
5 Results

5.1 Noticing of CM events

To answer the first research question, the number and type of events noticed by novice and expert teachers was analyzed for both formats of instruction. Figure 2 shows the results for both groups of teachers (cf. Appendix for details on specific events).

Figure 2

Number of events noticed by novice and expert teachers in the whole-group and the partner work format by type of event



Descriptive results and results of independent t-tests or Mann-Whitney-U-tests are reported in table 1 for the whole-group instruction format and in table 2 for the partner-work format. For the whole-group format, there were neither significant differences between novices and experts for the overall number of events noticed nor for the types of CM events (cf. table 1). Novices tended to notice more reactive teacher events than experts with a moderate effect size. Yet, the difference failed the level of significance ($t(38) = 1.91$, 95% CI [-0.07, 2.37], $p = .064$, $d = 0.60$).

Table 1

Expertise differences in teachers' noticing of CM events in the whole-group instruction format

	Novice teachers ***(n =20)***		**Expert teachers** ***(n = 20)***				
	M (SD)	*Mdn*	*M (SD)*	*Mdn*	Inferential statistics	*p*	*d*
Number of events noticed	7.40 (3.33)	6.50	7.15 (3.08)	7.50	*t* (38) = 0.25, 95% CI [-1.80, 2.30]	.807	0.08
Teacher events noticed	5.50 (2.52)	5.50	4.70 (2.70)	5.00	*t* (38) = 0.97, 95% CI [-0.87, 2.47]	.339	0.31
Reactive	3.90 (1.74)	4.00	2.75 (2.05)	2.50	*t* (38) = 1.91, 95% CI [-0.07, 2.37]	.064	0.60
Preventive	1.60 [a] (1.39)	1.50	1.95 [a] (1.70)	1.00	MR_{no} = 19.65; MR_{ex} = 21.53, *U* =217.00	.659	-0.15
Student events noticed	1.90 (1.30)	2.00	2.45 (1.23)	2.50	*t* (38) =-1.38, 95% CI [-1.36, 0.26]	.177	-0.44
Discipline	1.70 (1.17)	2.00	2.15 [a] (1.04)	2.00	MR_{no} = 17.88; MR_{ex} = 23.12, *U* =252.50	.157	-0.46
Learning	0.20 [a] (0.41)	0.00	0.30 [a] (0.47)	0.00	MR_{no} = 19.50; MR_{ex} = 21.50, *U* =220.00	.602	-0.17

Note. **[a]** = Group data deviates significantly from a normal distribution ($p < 0.05$, Shapiro-Wilk test). Thus, results of Mann-Whitney-U tests are reported instead of t-tests.

For the partner work format, expert teachers noticed significantly more relevant CM events overall than novices (*t*(38) = -2.49, 95% CI [-4.89, -0.51], *p* = .017, *d* = -0.79). Further analysis of the type of events showed that experts identify significantly more events focusing on students as note-worthy with regard to CM than novices (*t*(38) = -2.97, 95% CI [-2.52, -0.48], *p* = .005, *d* = -0.94), especially events focusing on student discipline (MR_{no} = 15.65; MR_{ex} = 25.35, *U* =297.00, *p* = .008, *d* = -0.91). Experts also tentatively noticed more preventive teacher events than novices (MR_{no} = 16.98; MR_{ex} = 24.02, *U* =270.50, *p* = .056, *d* = -0.63).

The hypotheses regarding our first research question were only partly supported by our data. As assumed, expertise effects were more prominent in the partner work format than in the whole-group format. However, against our assumptions a stronger focus of novices on student discipline, of experts on noticing student learning events as well as on preventive teacher events could not be confirmed. Furthermore, novices only tentatively noticed more reactive events in the whole-group format, while in the partner work format such events were more often noticed by experts.

Table 2

Expertise differences in teachers' noticing of CM events in the partner work format

	Novice teachers ***(n =20)***		**Expert teachers** ***(n = 20)***				
	M (SD)	*Mdn*	*M (SD)*	*Mdn*	Inferential statistics	*p*	*d*
Number of events noticed	6.85 (3.70)	7.00	9.55 (3.12)	9.50	*t* (38) = -2.49, 95% CI [-4.89, -0.51]	**.017**	-0.79
Teacher events noticed	4.50 (2.63)	5.00	5.70 (2.08)	6.00	*t* (38) = -1.60, 95% CI [-2.72, 0.32]	.117	-0.51
Reactive	1.95 (1.43)	2.00	2.80 [a] (1.24)	2.50	MR_{no} = 16.98; MR_{ex} = 24.02, *U* =270.50	.056	-0.63
Preventive	2.55 (2.04)	2.50	2.90 (1.71)	3.00	*t* (38) = -0.59, 95% CI [-1.56, 0.86]	.560	-0.19
Student events noticed	2.35 (1.60)	2.00	3.85 (1.60)	4.00	*t* (38) =-2.97, 95% CI [-2.52, -0.48]	**.005**	-0.94
Discipline	1.95 (1.40)	2.00	3.10 [a] (1.30)	3.00	MR_{no} = 15.65; MR_{ex} = 25.35, *U* =297.00	**.008**	-0.91
Learning	0.40 [a] (0.50)	0.00	0.75 [a] (0.77)	1.00	MR_{no} = 18.10; MR_{ex} = 22.90, *U* =248.00	.201	-0.42

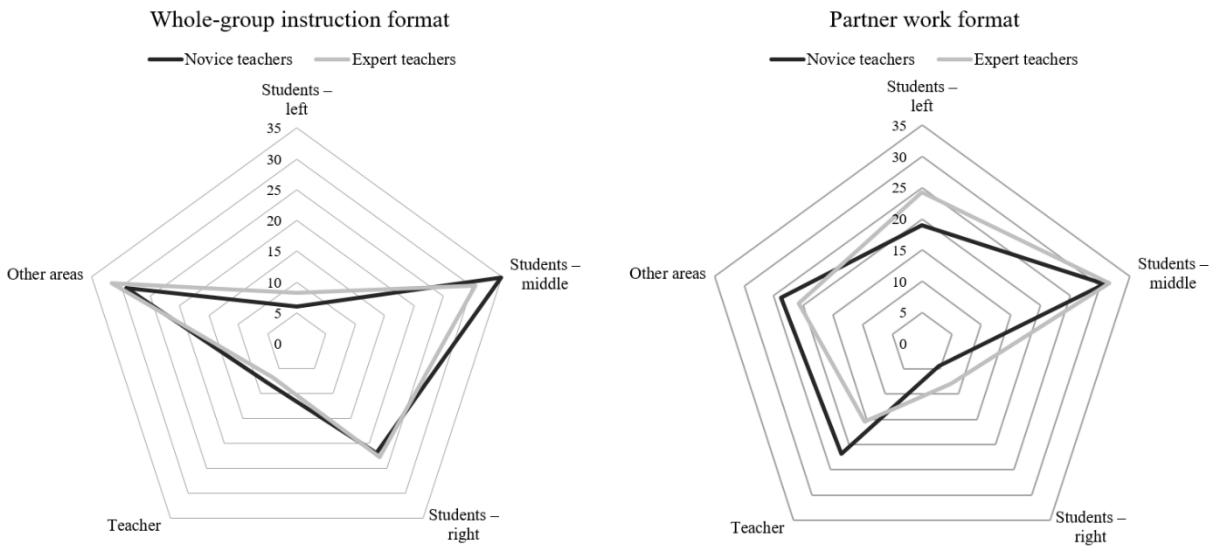
Note. [a] = Group data deviates significantly from a normal distribution ($p < 0.05$, Shapiro-Wilk test). Thus, results of Mann-Whitney-U tests are reported instead of t-tests.

5.2 Teachers' visual attention to student groups and the teacher

To answer our second research question, novice and expert teachers' proportion of gaze and the number of their fixations to three student groups or the teacher were compered for both formats of instruction. Descriptive results and inferential statistics are reported in table 3 and table 4. Figure 3 shows the distribution of novices' and experts' gaze.

Figure 3

Average proportion of gaze spent in the areas of interest by novice and expert teachers in the whole-group instruction and partner work format



In the whole-group format, both groups' proportions of gaze and number of fixations were highest for the middle and the right student group. Most gaze measure did not differ between novices and experts with the exception of the proportion of gaze to the left student group and the fixation count on the teacher. Experts' proportion of gaze to the left student group was significantly higher than novices' (t(38) = -2.13, 95% CI [-0.04, -0.0001], p = .040, d = -0.67). As hypothesized, novices allocated significantly more fixations to the teacher in the whole-group format than expert teachers (t(38) = 2.23, 95% CI [0.48, 11.32], p = .034, d = 0.71).

Table 3

Expertise differences in teachers' proportion of dwell time and fixation count on student groups and the teacher in the whole-group instruction format

		Novice teachers ***(n =20)***		**Expert teachers** ***(n = 20)***				
		M (SD)	*Mdn*	*M (SD)*	*Mdn*	Inferential statistics	p	d
Left student group	Prop. of gaze	0.06 (0.03)	0.05	0.08 (0.04)	0.08	t (38) = -2.13, 95% CI [-0.04, -0.001]	**.040**	-0.67
	No. of fixations	15.25[a] (9.22)	14.00	17.75 (7.15)	15.50	MR_{no} = 18.12; MR_{ex} = 22.88, U =247.50	.201	-0.42
Middle student group	Prop. of gaze	0.35[a] (0.09)	0.33	0.31 (0.08)	0.32	MR_{no} = 22.45; MR_{ex} = 18.55, U =161.00	.301	0.34
	No. of fixations	79.20 (19.99)	76.00	67.50 (21.54)	66.00	t (38) = 1.78, 95% CI [-1.60, 25.00]	.083	0.56
	Prop. of gaze	0.22 (0.09)	0.23	0.23 (0.10)	0.23	t (38) = -0.28, 95% CI [-0.07, 0.05]	.778	-0.09

Right student group	No. of fixations	50.95 (19.73)	52.50	48.60 (19.33)	46.00	*t* (38) = 0.38, 95% CI [-10.16, 14.86]	.706	0.12
Teacher	Prop. of gaze	0.08 (0.04)	0.08	0.07 [a] (0.07)	0.05	MR_{no} = 22.75; MR_{ex} = 18.25, *U* =155.00	.231	0.39
	No. of fixations	17.75 (8.12)	19.00	11.85 (8.81)	10.00	*t* (38) = 2.23, 95% CI [0.48, 11.32]	**.034**	0.71

Note. [a] = Group data deviates significantly from a normal distribution ($p < 0.05$, Shapiro-Wilk test). Thus, results of Mann-Whitney-U tests are reported instead of t-tests.

In the partner work format, significant expertise effects were found for the left and the right student group as well as the teacher which is in line with our assumption. Expert teachers' showed a significantly higher proportion of gaze than novices to the left (*t*(38) = -2.08, 95% CI [-0.104, -0.001], *p* = .044, *d* = -0.66) and the right student group (MR_{no} = 15.65; MR_{ex} = 25.35, *U* =297.00, *p* = .008, *d* = -0.91). Similarly, the number of fixations on the right student group was significantly higher for expert than for novice teachers (MR_{no} = 15.65; MR_{ex} = 25.35, *U* =297.00, *p* = .008, *d* = -0.91). In contrast and as hypothesized, novices allocated a significantly higher proportion of gaze towards (MR_{no} = 24.55; MR_{ex} = 16.45, *U* = 119.00, *p* = .028, *d* = 0.74) and fixated more often on the teacher than experts (MR_{no} = 24.80; MR_{ex} = 16.20, *U* = 114.00, *p* = .020, *d* = 0.79).

Our hypotheses for the second research question were generally supported by the data as more expertise differences were found in the partner work format and experts allocated more attention to student groups (in particular on the left and right side of the classroom) while novices paid more attention to the teacher in both formats of instruction.

Table 4

Expertise differences in teachers' proportion of dwell time and fixation count on student groups and the teacher in the partner work instruction format

		Novice teachers *(n =20)*		**Expert teachers *(n = 20)***				
		M (SD)	*Mdn*	*M (SD)*	*Mdn*	Inferential statistics	p	d
Left student group	Prop. of gaze	0.19 (0.07)	0.18	0.24 (0.09)	0.26	*t* (38) = -2.08, 95% CI [-0.10, -0.001]	**.044**	-0.66
	No. of fixations	33.80[a] (12.22)	30.00	40.45 (13.00)	43.50	MR_{no} = 17.40; MR_{ex} = 23.60, *U* =262.00	.096	-0.55
Middle student group	Prop. of gaze	0.31[a] (0.07)	0.31	0.32 (0.05)	0.32	*t* (38) = -0.38, 95% CI [-0.05, 0.03]	.703	-0.12
	No. of fixations	47.85 (15.27)	45.50	48.60 (7.96)	49.00	*t* (38) = -0.20, 95% CI [-8.55, 7.05]	.847	-0.06

Right student group	Prop. of gaze	0.05 [a] (0.03)	0.04	0.08 (0.04)	0.07	MR_{no} = 15.65; MR_{ex} = 25.35, *U* =297.00	**.008**	-0.91
	No. of fixations	8.45[a] (6.25)	7.00	14.35 (7.74)	14.00	MR_{no} = 15.65; MR_{ex} = 25.35, *U* =297.00	**.008**	-0.91
Teacher	Prop. of gaze	0.22 (0.09)	0.22	0.16[a] (0.11)	0.12	MR_{no} = 24.55; MR_{ex} = 16.45, *U* =119.00	**.028**	0.74
	No. of fixations	37.75 (12.90)	40.50	28.50[a] (19.16)	19.50	MR_{no} = 24.80; MR_{ex} = 16.20, *U* =114.00	**.020**	0.79

Note. **[a]** = Group data deviates significantly from a normal distribution ($p < 0.05$, Shapiro-Wilk test). Thus, results of Mann-Whitney-U tests are reported instead of t-tests.

5.3 Teachers' visual attention to classroom management events

Generally, teachers noticed a variety of events in both video clips (see Appendix for details on all events noticed). Some events were based on visual, others on audible perception. To examine the third research question, we focused on events grounded in visual perception. There were no significant differences between novice and expert teachers' visual attention to the four corresponding AOIs in terms of proportions of gaze or number of fixations (cf. table 5). Thus, our hypothesis was confirmed.

Table 5

Expertise differences in teachers' proportion of dwell time and fixation count on student groups and the teacher in the partner work instruction format

		Novice teachers *(n =20)*		**Expert teachers *(n = 20)***				
		M (SD)	*Mdn*	*M (SD)*	*Mdn*	Inferential statistics	p	d
Student lingers and clowns around	Prop. of gaze	0.04 (0.01)	0.04	0.04 (0.02)	0.05	*t* (38) = -1.25, 95% CI [-0.005, 0.003]	.220	-0.40
	No. of fixations	8.60 (3.12)	8.00	9.45 (3.30)	9.50	*t* (38) = -0.84, 95% CI [-2.91, 1.21]	.408	-0.27
Student is raising her hand and being ignored	Prop. of gaze	0.07[a] (0.05)	0.07	0.06[a] (0.05)	0.06	MR_{no} = 22.70; MR_{ex} = 18.93, *U* =168.50	.398	0.27
	No. of fixations	17.60 (9.17)	18.00	13.30 (7.77)	13.50	*t* (38) = 1.60, 95% CI [-1.14, 9.74]	.118	0.50
Two students fool around and fight each other	Prop. of gaze	0.13 (0.06)	0.14	0.16[a] (0.06)	0.18	MR_{no} = 16.95; MR_{ex} = 24.05, *U* =271.00	.056	-0.64
	No. of fixations	21.50[a] (11.36)	20.50	26.10 (9.61)	29.50	MR_{no} = 17.18; MR_{ex} = 23.82, *U* =266.50	.072	-0.59
Timer on the smartboard as orientation for students	Prop. of gaze	0.03[a] (0.03)	0.02	0.02[a] (0.03)	0.02	MR_{no} = 21.88; MR_{ex} = 19.22, *U* =172.50	.461	0.24
	No. of fixations	4.20[a] (3.59)	4.50	3.50[a] (3.12)	3.00	MR_{no} = 22.15; MR_{ex} = 18.85, *U* =167.00	.383	0.29

Note. **[a]** = Group data deviates significantly from a normal distribution ($p < 0.05$, Shapiro-Wilk test). Thus, results on Mann-Whitney-U tests are reported instead of t-tests.

6 Discussion, limitations and conclusions

6.1 Summary and discussion

This study examined how novice and expert teachers' noticing of CM events differs with regard to whole-group instruction and partner work. In particular, teachers' identification of note-worthy events based on their verbal reports and their visual attention to broader areas and specific events in the classroom video were investigated. Furthermore, with the video clips selected for this study, the narrow focus of prior research on behavioral management was broadened. In summary, experts noticed more CM events in the partner work format than novices and were further characterized by a focus on student events. Analyses of teachers' gaze revealed a stronger focus of experts on student groups, again especially in the partner work format. Novices paid more attention to the teacher than experts in both video clips. Finally, we found no evidence for a relationship between expertise and teacher gaze to specific CM events.

Expanding prior research that focused on CM in a whole-group setting, our study added an examination of partner work. That these two formats may be associated with different demands on teachers' CM was already suggested by Doyle (2006). In the present study, we found indeed that experts noticed significantly more CM events than novices in the partner work format. Thus, these results support the assumption that partner work may be more challenging for novice teachers in terms of noticing CM events. This result is also consistent with novices' tendency to regard CM as a primarily behavioral issue that calls for reactive CM (Kaufman & Moss, 2010; Reupert & Woodcock, 2010): The partner work format requires teachers to master a broader repertoire of strategies and novices might not have developed the CM scripts (Wolff et al., 2020) yet, thus might fail to notice CM events in this format.

So far it has been reported that novices tended to focus more on reactive CM as well as student discipline (Reupert & Woodcock, 2010; Wolff et al., 2015, 2017). In this study, we were not able to replicate this focus. Only for the whole-group format, novices tentatively noticed more reactive teacher events than experts. However, the opposite tentative result was found for the partner work scene with experts noticing more reactive teacher events and significantly more student discipline events. It is possible, that novices fail at noticing these events in the partner work format as they are not as salient as in the whole-group format. Novices might have at least developed some CM knowledge about behavioral problems in whole-group settings as this is the aspects of CM that is

often stressed in training programs and also in teachers' worries (Bear, 2015). Thus, novices' lack of CM scripts that guide their noticing becomes particularly apparent in the partner work format.

Against our expectations, experts did not notice more preventive teacher events than novices. We can only speculate about the reason for this result which may be due to the fact that not only behavioral problems were displayed in the video clips, but also, for example, instructional management (e.g. seating arrangements, time management, and lesson flow). Maybe these CM events are less subtle thus easier to notice for novices. However, prior studies investigated teachers' verbal analysis of CM-related events which is not necessarily comparable to the number and type of events noticed. Thus, the reported focus of novices on discipline and experts on learning might also be found in this study when teachers' comments about single, specific events are analyzed more qualitatively (cf. Wolff et al., 2015, 2017).

For the second research question, teachers' proportion of gaze and number of fixations at student groups or the teacher supported our hypotheses. Overall, experts looked longer and more often at student groups on the left and right side of the classroom, thus prioritizing those areas were potentially important student learning or student discipline events took place. This result is consistent with findings on general visual teacher expertise showing that experts focus more on students than novices (Cortina et al., 2015; McIntyre et al., 2019; McIntyre & Foulsham, 2018; Wolff et al., 2016). Monitoring student learning is particularly demanding in partner work (Doyle, 2006). Thus, keeping an eye on students in partner work is crucial for effective CM. Experts' prioritizing of students with their gaze might be the reason for their noticing of significantly more CM events, particularly of student discipline events that are less salient in the partner work than in the whole-group format. Novice teachers more bottom-up processing of classroom scenes (Wolff et al., 2020) was further confirmed in both formats as they allocate more attention to the teacher than experts: The teacher is salient by guiding the whole-group activity or interacting with student groups in the partner work activity.

Regarding the third research question, we further analyzed novice and expert teachers' gaze at specific CM events. Our result that the level of expertise did not make a difference in terms of visual attention to individual CM events, was previously also found with respect to one student event (Yamamoto & Imai-Matsumura, 2013). With the complexity of teaching in mind, these results may not be surprising. Both groups of teachers noticed many different events in both short classroom scenes. Noticing one among these CM events is probably a too specific and fine-grained measure. However,

as shown in the second research question, expertise is characterized by an allocation of attention towards students, were such events can potentially take place.

6.2 Limitations

Limitations of the present study need to be discussed before we turn to conclusions. Due to the high demands of analyzing eye-tracking and verbal data, the sample size is rather small, but comparable to other eye tracking studies. Also, power analysis showed that the sample size is sufficient to uncover similarly large effects as previously reported. However, it would be desirable to increase the sample size in future studies in order to uncover possible smaller effects.

Since teachers volunteered to participate, a self-selection bias is possible: Teachers that felt more confident could have been more willing to participate. However, this could apply to both novices and experts and thus balance each other. Our selection of experts was based on domain-specific experience (at least five years after qualification) and external evaluation (indicated by being selected for additional responsibilities and tasks in school or teacher education). However, we could not use student achievement or peer nomination as a further criterion of expertise (Caspari-Sadeghi & König, 2018; Palmer et al., 2005) because there is no longitudinal student testing in Germany that would allow to relate student achievement to a specific teacher. Since German teachers do not often observe their colleagues' lessons either (Richter & Pant, 2016), peer nominations might be uninformed.

Our study was conducted in Germany and here in an urban context. While the results might be similar in other Western nations, different results could be expected in other cultures or to some extent even in rural areas. Daily practices and concerns of CM (Bear et al., 2016) and teachers' gaze patterns can differ across cultures (McIntyre et al., 2019). We analyzed teacher's noticing in video segments of another teachers' instruction. Results could be different if teachers' observed their own instruction (Seidel et al., 2011). Additionally, the ecological validity of our results might have been higher if teachers' eye movements were recorded during teaching.

Using video clips to represent authentic teaching situation with two different formats of instruction allowed for a situated investigation of novice and expert teachers' noticing. This choice may include limited generalizability of results though. The representativeness of the selected video clips was ensured by an expert rating. However, while recording authentic teaching situations the

complexity of the resulting video clips could only be controlled to some extent. Thus, it cannot be ruled out that idiosyncratic features of the video clips may affect the generalizability of our results, particularly regarding the eye tracking data. Further research is needed that uses multiple sequences of one instructional format yielding more generalizable results.

6.3 Conclusions

The present study yielded new insights into teachers' noticing of CM events in two different formats of instruction that expand the state-of-research beyond behavioral management in whole-group instruction. The format-specific effects found regarding teachers' identification of note-worthy CM events and teachers' gaze point to the relevance of instructional formats for generalizing research results. We cannot take it for granted that findings are valid across different formats of instruction.

Furthermore, our results indicate that novice teachers may have different developmental needs regarding different formats of instruction. Therefore, further research should investigate format-specific differences in novice and expert teachers noticing. Also, teacher education and professional development programs may want to pay attention to CM in these different formats. The results of this study suggest that partner work is particularly challenging for novice teachers. In order to support novice teachers' in developing knowledge in this regard, using video-based or case-based activities could be promising approaches in teacher education (Boshuizen et al., 2020; Gaudin & Chaliès, 2015), as they could allow deliberate practice. How such interventions can change teachers' gaze priorities or identification of events is of particular interest with respect to the development of noticing: Can it be accelerated so it does not necessarily take many years of teaching experience?

In the present study, video clips were selected that showed CM events beyond behavioral CM. We regard these clips therefore to be more representative for a comprehensive understanding of CM than the narrower focus on behavioral management. With this approach, we did obtain some results that differed from previous studies. Therefore, further research on teachers' noticing is needed that accounts for such a understanding by paying attention to instructional management or social, emotional and motivational aspects of CM as well (Bear, 2015).

References

Beach, P., & McConnel, J. (2019). Eye tracking methodology for studying teacher learning: a review of the research. *International Journal of Research & Method in Education*, *42*(5), 485–501. https://doi.org/10.1080/1743727X.2018.1496415

Bear, G. G. (2015). Preventive and classroom-based strategies. In E. T. Emmer & E. J. Sabornie (Eds.), *Handbook of classroom management* (2nd ed., pp. 25–49). Routledge/Taylor & Francis.

Bear, G. G., Chen, D., Mantz, L. S., Yang, C., Huang, X., & Shiomi, K. (2016). Differences in classroom removals and use of praise and rewards in American, Chinese, and Japanese schools. *Teaching and Teacher Education*, *53*, 41–50. https://doi.org/10.1016/j.tate.2015.10.003

Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research*, *35*(5), 463–482. https://doi.org/10.1016/S0883-0355(02)00004-6

Berliner, D. C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology & Society*, *24*(3), 200–212. https://doi.org/10.1177/0270467604265535

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

Boshuizen, H. P. A., Gruber, H., & Strasser, J. (2020). Knowledge restructuring through case processing: The key to generalise expertise development theory across domains? *Educational Research Review*, *29*, 100310. https://doi.org/10.1016/j.edurev.2020.100310

Brophy, J. (1986). Classroom management techniques. *Education and Urban Society*, *18*(2), 182–194.

Carter, K., Cushing, K. S., Sabers, D. S., Stein, P., & Berliner, D. C. (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

Caspari-Sadeghi, S., & König, J. (2018). On the adequacy of expert teachers: From practical convenience to psychological reality. *International Journal of Higher Education*, *7*(5), 1–19. https://doi.org/10.5430/ijhe.v7n5p1

Chaplain, R. P. (2008). Stress and psychological distress among trainee secondary teachers in England. *Educational Psychology*, *28*(2), 195–209. https://doi.org/10.1080/01443410701491858

Charalambous, C. Y., & Praetorius, A. K. (2018). Studying mathematics instruction through different lenses: setting the ground for understanding instructional quality more comprehensively. *ZDM - Mathematics Education*, *50*(3), 355–366. https://doi.org/10.1007/s11858-018-0914-8

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffmann (Eds.), *The Cambridge handbook of expertise and expert performance* (1st ed., pp. 21–30). Cambridge University Press. https://doi.org/10.1017/CBO9780511816796.002

Clunies-Ross, P., Little, E., & Kienhuis, M. (2008). Self-reported and actual use of proactive and reactive classroom management strategies and their relationship with teacher stress and student behaviour. *Educational Psychology*, *28*(6), 693–710. https://doi.org/10.1080/01443410802206700

Copeland, W. D., Birmingham, C., DeMeulle, L., D'Emidio-Caston, M., & Natal, D. (1994). Making meaning in classrooms: An investigation of cognitive processes in aspiring teachers, experienced teachers, and their peers. *American Educational Research Journal*, *31*(1), 166–196. https://doi.org/10.2307/1163271

Cortina, K. S., Miller, K. F., 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*(2), 389–403. https://doi.org/10.1007/s10763-014-9610-5

Doyle, W. (1986). Classroom organization and management. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 392–431). Macmillan.

Doyle, W. (2006). Ecological approaches to classroom management. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practive and contemporary issues* (1st ed., pp. 97–125). Lawrence Erlbaum Associates.

Evertson, C. M., & Weinstein, C. S. (2006). Classroom management as a field of inquiery. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practive and contemporary issues* (1st ed., pp. 3–15). Lawrence Erlbaum Associates.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Gaudin, C., & Chaliès, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review*, *16*, 41–67. https://doi.org/10.1016/j.edurev.2015.06.001

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, *23*(4), 523–552. https://doi.org/10.1007/s10648-011-9174-7

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

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye Tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Hyrskykari, A., Ovaska, S., Majaranta, P. P., Räihä, K.-J., & Lehtinen, M. (2008). Gaze path stimulation in retrospective think-aloud. *Journal of Eye Movement Research*, *2*(4), 1–18. https://doi.org/10.16910/jemr.2.4.5

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017). Eye tracking in Educational Science: Theoretical frameworks and research agendas. *Journal of Eye Movement Research*, *10*(1), 1–18. https://doi.org/10.16910/jemr.10.1.3

Kaufman, D., & Moss, D. M. (2010). A new look at preservice teachers' conceptions of classroom management and organization: Uncovering complexity and dissonance. *Teacher Educator*, *45*(2), 118–136. https://doi.org/10.1080/08878731003623669

König, J., & Kramer, C. (2016). Teacher professional knowledge and classroom management: on the relation of general pedagogical knowledge (GPK) and classroom management expertise (CME). *ZDM - Mathematics Education*, *48*(1–2), 139–151. https://doi.org/10.1007/s11858-015-0705-4

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research*, *86*(3), 643–680. https://doi.org/10.3102/0034654315626799

Kounin, J. S. (1970). *Discipline and group management in classrooms*. Holt, Rinehart & Winston.

Lachner, A., Jarodzka, H., & Nückles, M. (2016). What makes an expert teacher? Investigating teachers' professional vision and discourse abilities. *Instructional Science*, *44*(3), 197–203.

https://doi.org/10.1007/s11251-016-9376-y

Martin, N. K., & Sass, D. A. (2010). Construct validation of the behavior and instructional management scale. *Teaching and Teacher Education*, *26*(5), 1124–1135. https://doi.org/10.1016/j.tate.2009.12.001

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, *22*(3), 276–282. https://doi.org/10.11613/BM.2012.031

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

McIntyre, N. A., Jarodzka, H., & Klassen, R. M. (2019). Capturing teacher priorities: Using real-world eye-tracking to investigate expert teacher priorities across two cultures. *Learning and Instruction*, *60*, 215–224. https://doi.org/10.1016/j.learninstruc.2017.12.003

Miles, W. R. (1929). Ocular dominance demonstrated by unconcious sighting. *Journal of Experimental Psychology*, *12*(2), 113–126.

Palmer, D. J., Stough, L. M., Burdenski, T. 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

Piwowar, V., Thiel, F., & Ophardt, D. (2013). Training inservice teachers' competencies in classroom management. A quasi-experimental study with teachers of secondary schools. *Teaching and Teacher Education*, *30*(1), 1–12. https://doi.org/10.1016/j.tate.2012.09.007

Reupert, A., & Woodcock, S. (2010). Success and near misses: Pre-service teachers' use, confidence and success in various classroom management strategies. *Teaching and Teacher Education*, *26*(6), 1261–1268. https://doi.org/10.1016/j.tate.2010.03.003

Richter, D., & Pant, H. A. (2016). *Lehrerkooperation in Deutschland: Eine Studie zu kooperativen Arbeitsbeziehungen bei Lehrkräften der Sekundarstufe I*. Deutsche Telekom Stiftung, Robert Bosch Stiftung, Bertelsmann Stiftung und Stiftung Mercator.

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

Schwab, Y., & Elias, M. J. (2015). From compliance to responsibility: Social-emotional learning and

classroom management. In E. T. Emmer & E. J. Sabornie (Eds.), *Handbook of Classroom Management* (2nd ed., pp. 94–115). Routledge/Taylor & Francis.

Seidel, T., Schnitzler, K., Kosel, C., Stürmer, K., & Holzberger, D. (2020). Student characteristics in the eyes of teachers: Differences between novice and expert teachers in judgment accuracy, observed behavioral cues, and gaze. *Educational Psychology Review*. https://doi.org/10.1007/s10648-020-09532-2

Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M., & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? *Teaching and Teacher Education*, *27*(2), 259–267. https://doi.org/10.1016/j.tate.2010.08.009

SensoMotoric Instruments. (2016a). *Be Gaze* (3.7).

SensoMotoric Instruments. (2016b). *Experiment Center* (3.7).

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children*, *31*(3), 351–380.

Stahnke, R., Schueler, S., & Roesken-Winter, B. (2016). Teachers' perception, interpretation, and decision-making: A systematic review of empirical mathematics education research. *ZDM - Mathematics Education*, *48*(1–2), 1–27. https://doi.org/10.1007/s11858-016-0775-y

Star, J. R., & Strickland, S. K. (2008). Learning to observe: Using video to improve preservice mathematics teachers' ability to notice. *Journal of Mathematics Teacher Education*, *11*(2), 107–125. https://doi.org/10.1007/s10857-007-9063-7

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. https://doi.org/10.1016/j.tate.2013.09.001

van Es, E. A., & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, *10*(4), 571–596.

van Gog, T., Paas, F., van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology. Applied*, *11*(4), 237–244. https://doi.org/10.1037/1076-898X.11.4.237

VERBI Software. (2017). *MAXQDA 2018*.

Wolff, C. E., Jarodzka, H., & Boshuizen, H. P. A. (2017). See and tell: Differences between expert and novice teachers' interpretations of problematic classroom management events. *Teaching and Teacher Education*, *66*, 295–308. https://doi.org/10.1016/j.tate.2017.04.015

Wolff, C. E., Jarodzka, H., & Boshuizen, H. P. A. (2020). Classroom Management Scripts: a Theoretical Model Contrasting Expert and Novice Teachers' Knowledge and Awareness of Classroom Events. *Educational Psychology Review*. https://doi.org/10.1007/s10648-020-09542-0

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*(2), 243–265. https://doi.org/10.1007/s11251-016-9367-z

Wolff, C. E., van den Bogert, N., Jarodzka, H., & Boshuizen, H. P. A. (2015). Keeping an eye on learning: Differences between expert and novice teachers' representations of classroom management events. *Journal of Teacher Education*, *66*(1), 68–85. https://doi.org/10.1177/0022487114549810

Wubbels, T., Brekelmans, M., Brok, P. den, Wijsman, L., Mainhard, T., & Tartwijk, J. van. (2015). Teacher–student relationships and classroom management. In E. T. Emmer & E. J. Sabornie (Eds.), *Handbook of Classroom Management* (2nd ed., pp. 363–386). Routledge/Taylor & Francis.

Yamamoto, T., & Imai-Matsumura, K. (2013). Teachers' gaze and awareness of students' behavior: using an eye tracker. *Comprehensive Psychology*, *2*(6). https://doi.org/10.2466/01.IT.2.6

Appendix

Table A.1

Events noticed in video clip 1 – Whole-group instruction format

Event	Novice teachers	Expert teachers	All teachers	Type of event
Individual events in temporal order				
Students take turns at presenting their solution at the smart board				
Teacher alternates between boys and girls	N = 5	N = 0	N = 5	TP
Student lingers and clowns around after his presentation *	N = 4	N = 10	N = 14	SD
Transition between students is not structured well	N = 2	N = 2	N = 4	TP
Teacher urges students to be quit for the first time				
Teacher calls students' names and urges them to be quiet	N = 16	N = 11	N = 27	TR
Teacher asks students to pull through because the break is close	N = 13	N = 11	N = 24	TR
Teacher is unmotivated and on edge	N = 4	N = 1	N = 5	TR
Anna presents her solution on the smartboard				
Teacher wants Anna to explain her solution	N = 6	N = 3	N = 9	TP
Students don't listen to Anna (uncomfortable for Anna)	N = 6	N = 8	N = 14	SD
Teacher presses Anna: „But you must be able to explain your solution!"	N = 3	N = 3	N = 6	TP
Teacher doesn't discipline the students during Annas presentation	N = 5	N = 4	N = 9	TR
Teacher urges students to be quiet again				
Teacher asks students to be quiet.	N = 13	N = 9	N = 22	TR
Teacher asks students to pull through because the break is close	N = 16	N = 10	N = 26	TR
Teacher asks students louder and more urgently to be quiet.	N = 9	N = 6	N = 15	TR
Teacher seems desperate.	N = 2	N = 3	N = 5	TR
Class is getting calmer.	N = 3	N = 3	N = 3	SD
Students not receptive anymore. Time for a break?	N = 1	N = 1	N = 2	SL
Lasting or repeated events				
Teacher doesn't keep an eye on students	N = 2	N = 3	N = 5	TP
Source of noise in the background	N = 1	N = 3	N = 4	TP
Whole class is unruly and loud	N = 15	N = 12	N = 27	SD
Student is raising her hand and being ignored*	N = 6	N = 9	N = 15	TP
Teacher's position in the room	N = 5	N = 6	N = 11	TP
Teacher's posture and presence	N = 1	N = 5	N = 6	TP
Missing structure and task	N = 0	N = 4	N = 4	TP
Wasted time, low time-on-task	N = 1	N = 1	N = 2	TP
Individual students are engaged and attentive	N = 3	N = 5	N = 8	SL
Individual students are disengaged and misbehaving	N = 6	N = 10	N = 16	SD
	N =148	N = 143	N = 291	

Note. TP = preventive teacher event, TR = reactive teacher event; SD = student discipline event; SL = student learning event; * = visible, frequently noticed event

Table A.2

Events noticed in video clip 2 – Partner work format

Event	Novice teachers	Expert teachers	All teachers	Type
Individual events in temporal order				
Whole class is loud at the beginning				
Students are loud and not paying attention	N = 3	N = 3	N = 6	SD
Teacher does not react to noisy class	N = 1	N = 3	N = 4	TR
Two students are fighting each other behind the teachers' back				
Two students fool around and fight each other*	N = 6	N = 12	N = 18	SD
Teacher does not notice the students fighting	N = 0	N = 3	N = 3	TR
Teacher does not react to students fighting each other	N = 2	N = 5	N = 7	TR
The boy with the hoody				
Student is hooded – against rules	N = 2	N = 7	N = 9	SD
Teachers does not react to hooded student (yet)	N = 1	N = 5	N = 6	TR
Student seems to be unmotivated and sad	N = 5	N = 7	N = 12	SL
Teachers talks briefly to hooded student	N = 5	N = 3	N = 8	TR
Teachers is hunched over and talks to hooded student (again)	N = 9	N = 8	N= 17	TR
Teacher goes through rows and monitors students	N = 7	N = 9	N = 16	TP
The boy with the hat				
Student puts on a hat – against rules	N = 10	N = 16	N = 26	SD
Teachers does not react to student with hat (yet)	N = 6	N = 13	N = 19	TR
Teacher pulls students' hat	N = 15	N = 16	N = 31	TR
Student takes off hat	N = 5	N = 6	N = 11	SD
Teacher talks to student in the left front corner	N = 1	N = 2	N = 3	TP
Lasting or repeated events				
Teacher doesn't keep an eye on students	N = 1	N = 3	N = 4	TP
Good student-teacher-relationship	N = 1	N = 3	N = 4	TP
Whole class is unruly and loud	N = 7	N = 14	N = 21	SD
Group work or partner work (phase and mode of instruction)	N = 13	N = 11	N = 24	TP
Teacher's position in the room	N = 1	N = 2	N = 3	TP
Teacher's posture and presence	N = 9	N = 6	N = 15	TP
Missing structure and task	N = 3	N = 3	N = 6	TP
Furnishing and architecture of the room	N = 2	N = 1	N = 3	TP
Seating arrangements	N = 4	N = 7	N = 11	TP
Right group and students in the front are attentive	N = 2	N = 4	N = 6	SL
Timer on smartboard as orientation for students*	N = 4	N = 4	N = 8	TP
Rule of no jackets or headdress in science rooms	N = 4	N = 8	N = 12	TP
Individual students are engaged and attentive	N = 1	N = 4	N = 5	SL
Individual students are disengaged and misbehaving	N = 6	N = 4	N = 10	SD
	N = 137	N = 191	N =328	

Note. TP = preventive teacher event, TR = reactive teacher event; SD = student discipline event; SL = student learning event; * = visible, frequently noticed event