Educational Studies in Mathematics (2022) 110:205–232
https://doi.org/10.1007/s10649-021-10128-y

Teacher noticing and its growth toward expertise: an expert–novice comparison with pre-service and in-service secondary mathematics teachers

Anton Bastian[1] · Gabriele Kaiser[2] · Dennis Meyer[3] · Björn Schwarz[4] · Johannes König[5]

Accepted: 8 November 2021 / Published online: 18 December 2021

Abstract
Although strong references to expertise in different theoretical approaches to teacher noticing have been made in the last decades, empirical knowledge about the development of teacher noticing from novice to expert level is scarce. The present study aims to close this research gap by comparing three different groups of mathematics teachers with different degrees of professional teaching experience—pre-service teachers at the master's level, early career teachers, and experienced teachers—using data sampled in the frame of the research program from the Teacher Education and Development Study in Mathematics (TEDS-M). Furthermore, the construct of teacher noticing is assessed in a differentiated way by analyzing different noticing facets. Findings confirm that three facets of teacher noticing can be empirically distinguished—perception of important classroom events, their interpretation, and decisions regarding further developments. The results reveal a considerable increase in professional noticing between master's students and practicing teachers. However, in contrast to other studies, among examples from East Asia, a stagnation or decrease in professional noticing between early career teachers and experienced teachers could be observed. Overall, the study highlights the cultural dependency of expertise development regarding teachers' noticing.

Keywords Teacher noticing · Expertise development · Expert–novice comparison · Mathematics education · Pre-service an in-service secondary mathematics teachers

During the last two decades, teacher noticing has increased in significance, particularly in mathematics teaching and mathematics teacher education (Schack et al., 2017; Sherin et al., 2011a). Although noticing is a common act, it is very specific to and within the teaching profession. As Mason (2002) notes, "every act of teaching depends on noticing: noticing what children are doing, how they respond, evaluating what is being said or done against expectations and criteria, and considering what might be said or done next" (p. 7). Thus, noticing poses the "primal questions of teaching" (Sherin et al., 2011b, p. 3).

✉ Gabriele Kaiser
gabriele.kaiser@uni-hamburg.de

Extended author information available on the last page of the article

Currently, noticing is considered integral to professional teacher competency and expertise (Kaiser et al., 2017) and seen as an essential component of teachers' expertise. It is the subject of numerous recent empirical studies, particularly in video-based research programs (overviews can be found in Santagata et al., 2021; Stahnke et al., 2016). Nevertheless, few studies have focused on the development of teacher noticing as part of their expertise development. Therefore, this paper aims to describe and analyze the development of pre-service and in-service mathematics teachers' expertise, focusing on teacher noticing by comparing master's students in initial teacher education with early career teachers and experienced teachers in the context of the Teacher Education Development Studies in Mathematics (TEDS-M) research program in Germany (Kaiser & König, 2019).

Despite increasing importance of teacher noticing in the educational discussion, consensus on the conceptualization of teacher noticing as a construct is lacking. While some theoretical approaches conceptualize noticing as a holistic concept (Mason, 2015), an analytical perspective on teacher noticing has gained importance in the last decade, particularly in empirical research distinguishing its different facets (for an overview, see Santagata et al., 2021). Although Sherin et al. (2011b) identify "attending to particular events" and "making sense of events" as two central aspects of noticing, disagreement surrounds the inclusion of a third aspect—"decision-making"—as described in Jacobs et al. (2010) and the conceptualization of our research team (Kaiser et al., 2015). However, not only are certain facets of noticing controversial but so are efforts to expand the focus of its attention beyond an exclusive focus on children's mathematical thinking and their strategies (Jacobs et al., 2010) to encompass whole lessons, including attending to important classroom incidents and decision-making regarding possible continuations of lessons and alternative student-teacher interactions (König et al., under review). Thus, to enrich this discussion with empirical arguments, we use our broadened conceptualization of noticing to examine the measurement of this construct comprising three facets—perception, interpretation, and decision-making—and test its empirical validity.

Although the necessity of teachers' expertise in teaching is emphasized in the current discussion, only a few empirical studies have analyzed the development of teachers' expertise and compared the noticing of novice and expert teachers. For example, early empirical studies of expertise by Berliner (1988) and colleagues (Carter et al., 1988; Sabers et al., 1991) demonstrated that expert teachers noticed issues concerning simultaneity, multidimensionality, and immediacy at higher levels than early career novice teachers or pre-service teachers in initial teacher education. Furthermore, expert teachers focused more on students and important content-related aspects of teaching. A more recent study by Jacobs et al. (2010) comparing various expertise groups of (pre-service) primary teachers indicated that teachers' growth and expertise development were decisive factors for the development all three noticing facets with teaching experience (see also Yang et al., 2021b).

Overall, despite this important work, the question of whether these three noticing facets of perception, interpretation, and decision-making can be separated empirically based on quantitative data from large-scale studies remains open. Furthermore, how these three facets differ empirically across various teachers' groups with varying professional expertise requires further analysis. Finally, the influence of the conceptualization of teacher noticing merits investigation with regard to how teachers' noticing differs across various groups if the focus of noticing is broadened from children's mathematical thinking to activities encompassing entire lessons. The present study aims to close this gap by empirically examining how different pre- and in-service teachers' groups notice, adopting a more analytic perspective by empirically differentiating the noticing construct into three different facets.

1 Literature review, theoretical framework, and research questions

1.1 Literature review on expertise development in noticing

Research on teacher noticing has mainly been conducted within the last two decades and has gained increased attention due to its focus on students and their mathematical thinking and its relevance to quality-oriented teaching. Despite the consensus regarding the central characteristics and aspects of teachers' noticing, the construct has largely been conceptualized heterogeneously (Dindyal et al., 2021; Schack et al., 2017; Sherin et al., 2011a; Stahnke et al., 2016). Given the extensive body of research on teacher noticing, this review includes work that directly relates to the study described in this paper. First, theoretical perspectives on teacher noticing and its conceptualizations and foci are displayed, followed by research on the development of teacher noticing.

As previously mentioned, the current discussion on teacher noticing can be characterized by a certain inconsistency concerning the theoretical orientation and conceptualization of teacher noticing. To present a systematized view on the current discussion, we refer to a classification proposed in a systematic literature survey on noticing education among mathematics teachers with a focus on video-based programs (Santagata et al., 2021). Santagata et al. (2021) described four theoretical perspectives on teacher noticing.

The first perspective—called the cognitive-psychological perspective—characterizes noticing as the underlying, subconscious mental processes in which teachers engage during teaching or teaching observation. In their seminal paper on the promotion of teacher noticing by video clubs, van Es & Sherin (2002) distinguished the following three facets of teacher noticing:

> (a) identifying what is important or noteworthy about a classroom situation; (b) making connections between the specifics of classroom interactions and the broader principles of teaching and learning they represent; and (c) using what one knows about the context to reason about classroom interactions. (p. 573)

This approach has been further developed by several scholars. For example, Jacobs et al. (2010) refocused the first two facets of noticing as "attending to children's strategies" and "interpreting children's mathematical understandings" and included a unique third facet of noticing, "deciding how to respond on the basis of children's understanding" (pp. 172-173). A similar conceptualization has been developed within the TEDS-M research program, in which decision-making as the third facet of noticing plays an important role. However, in contrast to Jacobs et al. (2010), the focus of the conceptualization of the TEDS-M research program is broadened by focusing on the whole lesson under mathematics educational and pedagogical perspectives. This construct will be elaborated in the theoretical framework section (Kaiser et al., 2015; Yang et al., 2021a). The role played by specific mathematical content domains such as fractions, ratios, multiple representations, or functions is only emphasized in a few studies (see, e.g., Dreher & Kuntze, 2015; Friesen & Kuntze, 2021; Ivars et al., 2020). Another stance is taken by Stockero and Van Zoest (2013) and Stockero (2021), who foreground "in-the-moment noticing" by zooming in on pivotal teaching moments or interruptions in the flow of lessons, which indicate opportunities to modify instructions in order to enhance student mathematical understanding. In their most recent paper, van Es & Sherin (2021) further developed their framework by rephrasing and regrouping these three facets as "attending" and "interpreting." They included a new third facet—shaping—defined as the "act of creating interactions that provide increased

opportunities to attend to and interpret noteworthy mathematical interactions" (p. 17); this new facet reinforces a knowledge-driven perspective on noticing.

In the second perspective, noticing is conceptualized as situated and socially constructed. This socio-cultural perspective draws on the seminal work of Goodwin (1994) on professional vision and is intertwined with the first perspective. Goodwin (1994) described professional vision as "socially organized ways of seeing and understanding events that are answerable to the distinctive interests of a particular social group" (p. 606), emphasizing the socio-cultural aspects of professional ways of seeing in certain professions (e.g., law or archaeology). Although the term "professional vision" has been adopted in several approaches and studies, its usage does not always involve socio-cultural framing, and some instead employ a cognitive-psychological perspective that also takes a quality-oriented focus on noticing (e.g., Seidel & Stürmer, 2014). More recently, theoretical conceptualizations emphasizing power and equity have been developed. These are more strongly connected with the original understanding of the professional vision of Goodwin. These conceptualizations challenge the focus on cognition by emphasizing the material and reciprocal nature of noticing (Dominguez, 2019) or introducing a sociopolitical perspective in which anti-deficit noticing shall be developed (e.g., Louie, 2018; Louie et al., 2021).

A third theoretical perspective on noticing—discipline-specific—was developed by Mason (2011) and conceptualizes teacher noticing as a "collection of practices designed to sensitize oneself so as to notice opportunities in the future in which to act freshly rather than automatically out of habit" (p. 35). Although this conceptualization of noticing attends to mental processes, differences in the cognitive-psychological perspective can be identified as the construct of sensitized awareness of the teachers is put in the foreground, which should "be methodical without being mechanical" (Mason, 2002, p. 61). The perspective by Mason (2002) is characterized by its focus on the professional development of teachers, for which Mason (2002) developed a set of practices, including systematic reflection and recognizing, which makes this approach different than the cognitive-psychological perspective.

The fourth perspective corresponds to the expert–novice paradigm developed by Berliner (1988), which can be described as a kind of precursor to research on teacher noticing (Lachner et al., 2016), although the noticing construct is not explicitly used in this perspective. The approach focuses similarly on the cognitive-psychological perspective on teacher ability to attend to important elements of teaching events, create coherent interpretations, and ignore distracting elements (Sabers et al., 1991). Overall, the processes of interpreting classroom situations, making sense of important events, and developing connections are described in this perspective as individual cognitive mental processes carried out by teachers, revealing connections to the discourse of teacher noticing. This perspective will be revisited during the discussion of the development of teachers noticing.

More recent work aims to overcome these different normative orientations (e.g., Scheiner, 2021); however, it thus far remains open just how strongly these new approaches will influence the discussion.

Concerning the relevance of these four theoretical perspectives, Santagata et al. (2021) report in their systematic literature survey that most studies examined (about 90%) referred to the cognitive-psychological perspective using the term "noticing," while the other three perspectives played more minor roles. In line with the dominance of the cognitive-psychological perspective is a more analytical view on noticing that advocates distinguishing different facets of noticing. As Santagata et al. (2021) point out, most examined studies in the systematic literature survey focus on the facets of attending/perceiving and interpreting/reasoning (about 80%), while fewer studies (about 37%) included responding/

decision-making. Concerning the object of the study, most of the examined studies analyzed teacher noticing of student thinking (83%), followed by a focus on instructional practices and classroom discourse (57%). Mathematical topics were addressed in fewer studies (20%). However, despite the dominance of the analytic approach, which is separating the different facets of noticing, empirical evaluation and analysis is often based on a holistic approach concerning the facets of noticing; that is, even if facets are distinguished, they are not assessed separately. Only a few instruments have measured the facets of noticing separately in compliance with high psychometrical standards, namely the observer/observer extended tool (Seidel & Stürmer, 2014), pre-service teachers' professional vision of inclusive classrooms (PVIC) (Keppens et al., 2019), the video-vignettes developed within the TEDS-M research program (Kaiser et al., 2015), and the instrument by Steffensky et al. (2015).

Another stance is taken in a recent literature survey by Amador et al. (2021), who analyzed methodological approaches to support and analyze pre-service teacher noticing by focusing on the two most currently influential theoretical frameworks on noticing development within the US discussion. The first framework by van Es (2011) on learning to notice mathematical thinking distinguished four levels of expertise in noticing, providing descriptors for each. The second framework, developed by Jacobs et al. (2010) on the professional noticing of children's mathematical thinking, empirically described the development of teacher noticing across different expertise groups using limited, lacking, or robust evidence as indicators for noticing. The results of the literature survey by Amador et al. (2021) indicated that more studies using the learning to notice framework reported positive results regarding progress in the development of noticing compared to studies using the professional noticing framework. However, as the effects reported in the included studies cover different participant groups and different designs, the question of how teacher noticing develops empirically remains open.

In the following, the research on the development of teacher noticing is delineated, and the fourth theoretical perspective is described in more detail. As alluded, earlier research on expertise and expert–novice comparisons identified important results regarding the development of expertise in noticing, which must be taken up in current discussions, although the construct of noticing was not explicitly used in these studies. In his seminal work, Berliner (1988) distinguished five different stages in teacher expertise development—novices, advanced beginners, competent teachers, proficient teachers, and expert teachers. While novices are characterized by the need for context-free rules and inflexible teaching and competent teachers act rationally but only with minor flexibility, expert teacher behavior can be described as "arational" (Berliner, 1988, p. 5), which is fluid, holistically, and unconsciously flexible. Experience plays a crucial role in expertise development and should be regarded as a necessary yet insufficient condition for expertise (Berliner, 2004; Palmer et al., 2005). In empirical studies comparing novice and expert teachers, Berliner (1988) highlighted the difficulties that novices face in interpreting classroom phenomena. Novice teachers were unable to predict teaching and student behavior, particularly concerning student errors. Further work from this research group shows a strengthened relationship to the noticing discussion. For example, the study by Sabers et al. (1991) described the differences between novice and expert teachers regarding the perception and interpretation of important classroom events.

These results are confirmed more recently by the literature survey by Stahnke et al. (2016), which suggests that novices tend to have difficulties perceiving or interpreting student work, which apparently can be improved by video-based professional development programs. Decision-making appears to be the most ambitious facet of noticing for novices,

as pre-service teachers seem to have difficulties understanding student mathematical thinking and their solution processes in a constructivist sense.

Nonetheless, a key area of concern is that although expertise is an important reference point in the theoretical frameworks of many studies on noticing, most current studies are restricted to pre-service teachers and do not include practicing teachers, which means that limited knowledge is available concerning the development of teacher expertise in noticing, as was noted in the systematic literature survey by Santagata et al. (2021). In a more extensive literature survey focusing on teacher noticing, König et al. (under review) indicated that only a few of the papers reviewed included both pre-service and in-service teachers and allowed examination of the development of teacher noticing, either in a cross-sectional or a longitudinal way, although the researchers used different constructs and definitions of novice and expert. Most of these few papers offered quantitative analyses of the differences between novice and expert teachers, mainly described as pre-service and in-service teachers. Only these few papers included in the literature survey focused explicitly on and reported noticing progress between novice and expert teachers. The subjects covered ranged from mathematics education (Jacobs et al., 2010) to elementary science (Meschede et al., 2017) or physical education (Reuker, 2017) or focused on classroom management (Gold & Holodynski, 2017). The studies clearly indicate that experts outperform novices in professional noticing.

One of the rare exceptions concerning comparisons of experts and novices is the seminal study by Jacobs et al. (2010) that provided insight into expertise development. They compared four different expertise groups ranging from pre-service elementary school teachers and three groups of practicing elementary school teachers that differed in terms of their years of teaching experience and were at different stages of a professional development program concerning children's mathematical thinking. Video clips and written student work were used to elicit teacher noticing of student mathematical thinking. The participants answered prompts concerning the three identified facets of noticing, attending to children's strategies, interpreting children's understanding, and deciding how to respond based on children's understanding. Within the study, an extensive coding manual with indicators for the adequacy of the analysis was developed using the categories of robust, limited, or lack of evidence. Jacobs et al. (2010) primarily identified a monotonic development across the four groups for all three noticing facets, "indicating that increased experience with children's thinking was related to increased engagement with children's thinking on the professional-noticing tasks" (p. 181). Their findings suggest that professional teaching experience and development programs facilitate expertise development in attending to children's strategies and interpreting children's understanding; however, they could not substantiate the same hypothesis for deciding how to respond. They also reported that "professional development seems to provide support for developing expertise in all three component skills" (p. 182), which provides evidence for the effectiveness of continued long-term professional development in this area, especially regarding leadership activities.

The results are confirmed by Yang et al. (2021b), who compared Chinese pre-service and two groups of in-service mathematics teachers with different degrees of teaching experience with respect to growth in their professional noticing. Although the study differentiated three facets of noticing, the facets interpretation and decision-making are joined in the results due to the low number of items on decision-making. The study reports mean differences in noticing across the three expertise groups that can be interpreted as a "nearly linear growth…with significant differences identified between pre-service and experienced teachers and only small differences between pre-service and early career teachers" (p. 29). Furthermore, the results using differential item functioning (DIF) analyses point out that

pre-service and early career teachers showed strengths in more reform-oriented Western topics, such as cooperative learning or mathematical modeling, which they may have learned during their university study, whereas experienced teachers showed strengths in analyzing student mathematical thinking. They found that the three noticing facets develop differently between the three expertise groups, with perception being better developed at the pre-service and beginning stages of teaching in contrast to interpretation and decision-making (which are more difficult to develop), which is aligned with the results by Jacobs et al. (2010).

Further results regarding expert–novice comparisons of teacher noticing are reported by other studies. For example, Dreher and Kuntze (2015) indicated that in-service teachers more strongly appreciated the usage of multiple representations and the role of changes of representations for fostering student understanding than pre-service teachers. Huang and Li (2012) emphasized qualitative differences between early career teachers and highly experienced teachers concerning the development of mathematical thinking and knowledge, student participation, and teacher motivation, with more experienced teachers attaining higher achievements.

Overall, more empirical research is needed to better understand the development of teacher noticing and the factors influencing this development, especially the role of teaching experience and expertise, its relation to the development of the different noticing facets, and how these factors manifest in a Western context.

1.2 Context of the study and theoretical framework

As previously explained, in the last decade, the paradigmatic differences between cognitively oriented frameworks on teacher competence that emphasize teacher knowledge and situated approaches putting social and situated aspects of teaching in the foreground have been challenged. In order to overcome these paradigmatic differences between cognitive and situated frameworks, Blömeke et al. (2015) developed the new theoretical approach of competence as continuum, which has become a prominent framework for teacher competencies. This approach called for situation-specific skills mediating between teacher knowledge and beliefs as dispositions and teacher performance.

The current study, which is carried out within the TEDS-M research program, departs from this discussion around these paradigmatic differences. To facilitate a better understanding of the context of the current study, this section briefly describes the various studies carried out within this program (for details of the development, see Kaiser et al. (2017)). The TEDS-M research program was initiated by the international TEDS-M study 2008, which took place across 17 countries from 2008 to 2010 and evaluated the professional knowledge of pre-service teachers for primary level and mathematics teachers for secondary level at the end of their study (Tatto et al., 2012). TEDS-M can be characterized as a cognitively oriented study with a focus on the professional knowledge of pre-service teachers referring to the well-known classification of teacher knowledge by Shulman (1986). Within the study, mathematical content knowledge (MCK), mathematics pedagogical content knowledge (MPCK), and general pedagogical knowledge (GPK) in a sub-sample of countries, including Germany, were evaluated; these knowledge domains were complemented by affective-motivational characteristics, such as beliefs about mathematics and its teaching and epistemological beliefs. The standardized testing instruments consisted mainly of multiple-choice items and a few open-ended items and were carried out as paper-and-pencil tests. The evaluation presented strong country differences, with the

participating pre-service teachers from East Asia achieving the best results and those from South America and Africa receiving the lowest results in all knowledge facets. Further analyses indicated, among others, the strong influence of opportunities to learn and background variables (Blömeke et al., 2012).

A follow-up study of TEDS-M (TEDS-FU) by German scholars further developed the conceptualization and the instruments of TEDS-M towards a situated approach including situation-specific skills—called teacher noticing—to compare the development of teacher expertise from pre-service teachers to early career teachers with instruments that more closely resembled actual classroom practices. Theoretically, the cognitively oriented knowledge framework was enriched by additional situated facets covering perception, interpretation, and decision-making as facets of noticing (the so-called PID model). For the measurement of these newly developed noticing facets, a standardized measurement instrument was developed based on several short video-vignettes and extensive coding manuals using closed and open-ended items. The study was conducted from 2011 to 2014, with voluntary participants from the original sample of TEDS-M (for a description, see Kaiser et al. (2015)). Among others, the study analyzed the structure of mathematics teacher competence, distinguishing stable teacher cognitions consisting of subject-related knowledge facets (MCK and MPCK) with a facet of fast error recognition and situation-specific cognitions with general pedagogical knowledge (GPK) and noticing (PID) (Blömeke et al., 2016). Concerning the development of mathematical and mathematics pedagogical content knowledge, the study noted a significant decrease of MCK, which is strongly influenced by MCK measured at the end of teacher education, and no change of the knowledge level of MPCK with less influence of MPCK measured at the end of teacher education. These results point to the strong impact of school practice on the development of MPCK, which is aligned with results by the expertise research (Berliner, 1994). However, the amount of school practice appears inconclusive, which points to the relevance of a "deliberate practice" (Ericsson et al., 1993).

The theoretical framework and the instruments of TEDS-FU have been further enriched, among others by a situation-specific test on classroom management expertise (König, 2015) and instruments for the in vivo evaluation of instructional quality (Schlesinger et al., 2018). These instruments were used in further studies, namely TEDS-Instruct and TEDS-Validate, which included student learning gains within the entire impact model of the TEDS-M research program. The analysis of the link between pedagogical knowledge and situation-specific classroom management expertise, instructional quality, and secondary student mathematics achievement indicated influences of teacher competence on instructional quality and student mathematical progress (König et al., 2021). Further analyses concerning this important impact model are under review.

Finally, within the TEDS-East-West study, most of the instruments were transferred to an East Asian context, evaluating, among others, the relation between teacher knowledge and their professional noticing. Yang et al. (2021a) reported a stronger relation between teacher knowledge and interpretation/decision-making facets of noticing compared to perception. In another study (Yang et al., 2021b), the development of teacher noticing across different expertise groups has been explored, reporting a nearly linear development over the different expertise groups. This is aligned with the results by Jacobs et al. (2010). This will be revisited in the discussion section.

In the following, our own framework on teacher noticing is described, which was developed within the discourse outlined above. We conceptualize noticing as a set of situation-specific skills comprising three interconnected facets: "(a) **p**erceiving particular events in an instructional setting; (b) **i**nterpreting the perceived activities in the

instructional setting and; (c) **d**ecision-making, either as anticipating responses to students' activities or as proposing alternative instructional strategies" (Kaiser et al., 2015, p. 374); this construct is termed the "PID model" (Blömeke et al., 2015). This conceptualization of noticing has specific characteristics that make it distinct from other conceptions, such as the original noticing approach developed by van Es & Sherin (2002), the professional vision approach by Seidel and Stürmer (2014), and further developments described in the survey paper by Dindyal et al. (2021).

This distinct theoretical framework is deliberately not referring to the construct of attending, which is the typical terminology in the noticing discussion as described above, as this construct is strongly related to interpreting; both facets are described as cyclical process of perception and interpretation within this discussion (Sherin et al., 2011b). Departing from the expert–novice paradigm in which the construct perception is widely used to describe the first phase of teacher actions in an instructional setting and by restricting the construct to observable, discernable incidents (Berliner, 2001; Carter et al., 1988), we conceptualize perception more narrowly as "seeing something without or with only minor reference to interpretation." Apart from this different conceptualization of attending and interpreting in our framework, we include decision-making as a third facet of noticing and an indispensable part of noticing as teacher actions in classrooms are essential to their expertise (Erickson, 2011).

Our framework moves forward the focal point of many studies, which delineate as core of their study the focus on student mathematical thinking, strategies, or reasoning (e.g., Jacobs et al., 2010 or Sherin & van Es, 2009) or in-the-moment noticing (Stockero, 2021; Stockero & Van Zoest, 2013). Decisive for our broadened understanding of noticing is the connection between noticing and quality-oriented mathematics education with a clear content-related orientation besides the pedagogical perspective (similarly to Ivars et al., 2020). In detail, we consider various aspects as decisive for a comprehensive understanding of teacher noticing, among others the various dimensions of instructional quality, namely classroom management, potential for students' cognitive activation, individual learning support, and scaffolding, considering both pedagogical and mathematical perspectives as implemented in many empirical studies (Praetorius et al., 2018). Furthermore, in our framework, a broad range of classroom incidents is combined with aspects of instructional quality, such as preventing classroom disturbances, addressing heterogeneous groups of students, and designing individualized and collective mathematical teaching and learning trajectories. Finally, possibilities of continuing a lesson or developing a new teaching sequences are included considering among others subject-related misconceptions or errors of students, and possible alternative of teacher-students' interactions (Yang et al., 2021a).

With this comprehensive understanding of noticing, we model a broad range of teachers' situation-specific skills and, with reference to Blömeke et al. (2015), integrate teacher noticing as a complement to the former cognitive-oriented and knowledge-based framework of teacher competence, which characterizes teacher competence within the scope of the TEDS-M research program.

Shaped by our theoretical framework of noticing and based on the shortcomings in empirical studies on teacher noticing identified above, our study addresses the following research questions:

1. Can the three distinguished noticing facets of perception, interpretation, and decision-making be separated empirically using quantitative models and measured through reli-

able scales? If yes, how are the scales representing the facets intercorrelated in this model?
2. How do teacher noticing and its facets differ between different pre- and in-service teacher groups with varying degrees of professional teaching experiences?

2 Methodology

Due to the complexity of our research questions, we draw our data from several studies conducted between 2011 and 2020 that were embedded in the TEDS-M research program to obtain a rich sample of cross-sectional data.

2.1 Sample and participants

The sample consists of three groups of pre-service and in-service secondary level mathematics teachers ($N = 457$) who participated in a survey of one of the following studies of the TEDS-M research program:

- TEDS-FU: 2011–2014, early career teachers from all over Germany, former participants of the study TEDS-M
- TEDS-Instruct (TEDS-I): 2014–2016, practicing teachers from the federal state of Hamburg, Germany
- TEDS-Validate (TEDS-V): 2016–2019, practicing teachers from three federal states in Germany (Thuringia, Saxony, Hesse)
- TEDS-Validate-Transfer (TEDS-V-T): 2020–ongoing; students pursuing master's degrees in education from six German universities

To ensure that the groups were clearly distinguished by their years of teaching expertise, the participants were divided into three groups based on their teaching experience while also taking study affiliation into account. The master's students from TEDS-Validate-Transfer ($n = 110$) had no professional teaching experience[1] and, hence, are novices according to the stage model by Berliner (1988). Participants with about 4.6 years of teaching practice ($SD = 0.5$, range from 3.5 to 6 years) from TEDS-Follow-Up ($n = 146$) and taken in parts from TEDS-Instruct and from TEDS-Validate ($n = 47$)—to assure a better discriminability between the groups—constituted the group of early career teachers. According to the Berliner model (1988), they can be classified as being competent at least; some may already have attained proficiency. The third group ($n = 154$) comprised experienced teachers from TEDS-Instruct and TEDS-Validate, who had on average 19.6 years of teaching experience ($SD = 10.4$, range from 6.5 to 41.5 years) and therefore represented participants at the competent, proficient, and expert levels. It can be assumed that the latter group included more proficient and expert teachers based on their more extensive teaching practice and reflection. As mentioned, however, experience alone is insufficient to constitute expertise (Caspari-Sadeghi & König, 2018; Palmer et al., 2005; Stigler & Miller, 2018). Rather, experience constitutes an essential factor for expertise. The distinguished groups represent

[1] The master's students had not experienced any extensive teaching practice and internships in their teacher education yet. Whether some students participated in teaching on their own was not asked and could therefore not be considered.

Table 1 Characterization of the three expertise groups

Variable	Expertise groups						Test of group differences	
	Master's students		Early career teachers		Experienced teachers		χ^2 (*df*)/odds ratio	*p*
Sample size	110		193		154			
Gender (% female)	66		58		55		3.76 (2)	.152
Teaching type (% academic track)	40		52		46		6.72 (2)	.035
	M	*SD*	*M*	*SD*	*M*	*SD*		
Age (in years)	24.57	4.08	31.68	5.50	46.00	9.54	488.36	<.001
Teaching experience (in years)	---	---	4.61	.50	19.64	10.41	397.45	<.001
GPA (Abitur)	2.09	.63	2.11	.61	---	---	21.47	.742
Grade (University)	---	---	1.87	.54	1.95	.61	91.31	<.001
Grade (induction phase)	---	---	1.96	.61	1.98	.67	48.48	.022

GPA grade point average; German grades range from 1 (best) to 4 (passed). For the test of group differences in some cases, Fisher's exact test was conducted, since the assumptions for a χ^2 test were violated

at least three different stages in expertise development, and, therefore, they facilitate an expert–novice comparison. Moreover, the comparison between the two in-service groups may yield valuable insights into the possible impact of long-term teaching experience on expertise development. Table 1 presents further characterization of these expertise groups, including grades from university entrance examinations, initial teacher education and the second phase of teacher education (so-called induction phase), the types of schools in which the teachers were contemporarily teaching, and whether the schools were academically oriented with higher achieving students or a more comprehensively oriented track for students of all ability levels.

2.2 Assessment instrument

The video-based test instrument used to measure teacher professional noticing limited to secondary education was developed within the TEDS-FU study (Kaiser et al., 2015; Kaiser et al., 2017). It comprises three scripted video-vignettes—*Frog King*, connected to a German fairy tale, *Box*, and *Solids*—of around three and a half minutes in length. The video-vignettes cover a wide range of mathematical topics (e.g., functions, surface, volume calculations, and modeling), different teaching phases, and different school types and showed lessons in the 9th and 10th grades (student age: 14–15 years). During the test, the participants first received some background information about the learning setting, the students, the previous lessons, and the mathematical topic before watching each video-vignette. For instance, the covered mathematical task and its solution were presented. The participants were then permitted to watch the video-vignette simulating real classroom situations once. Directly after watching the video-vignette, participants received open-response and closed (Likert-type) questions, which tested their abilities in all three noticing facets—perception, interpretation, and decision-making—with a mathematics pedagogical or general pedagogical focus. In total, 77 items were included in the tests; seven of these were excluded during scaling owing to their weak psychometric quality that rendered them unsuitable for the

Table 2 Number of items per noticing facet

Noticing facet	Number of rating scale items	Number of open-response items	Number of items used
Perception	17	5	22
Interpretation	19	18	37
Decision-making	0	11	11

samples that were combined specifically for use in this comparative study. Thus, the test instrument used for this study contains 70 items, 36 closed items with Likert scales, and 34 open-response items. The items focused on all three facets of noticing (see Table 2), with the lowest number of items focusing on decision-making. The low number of decision-making items is due to their complexity in the construction, as they must be formulated with strong restrictions to be unambiguous. Overall, the testing lasted about 60 minutes in total, with 15–20 minutes per video-vignette.

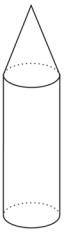
The test used video-vignettes rather than live observations of the participants in classroom for two reasons. First, this way of data collecting would not have been manageable for the intended sample size, and, second, the use of video-vignettes allows to create situations that are similar to practice, give comparable and standardized test results, and simultaneously present a controllable test environment and a strong connection to performance in classroom (Hughes & Huby, 2001; Piwowar et al., 2018). Moreover, this sort of instrument is partially analagous to situational judgement tests, whose functioning and validity are well examined (Lievens et al., 2008). The research team decided to use scripted videos instead of real teaching footage to assure a density of noteworthy events and provide impressions of a whole teaching lesson in a feasible test duration time.

An expert rating was conducted during test instrument development to determine which answer could be regarded as correct with respect to the rating scales. A coding manual was developed to assess the open-response items and piloted to improve its reliability and validity before it was used in TEDS-FU. For each item, the manual detailed which elements must be present in the response to be coded as correct as well numerous example-answers for correct and incorrect answers as well as border cases gathered in the pilot study. Furthermore, the coding rules left space for minor interpretation, which facilitates a low-inferent coding. Various approaches, including curricular analyses of the mathematical content and comprehensive expert workshops, were employed to ensure the validity of the instrument content and authenticity of the video-vignettes (for details, see Hoth et al., 2016; Kaiser et al., 2015).

The following presents three items and their coding to provide a prototypical insight on the items used and the coding process. In the video-vignette called *Solids*, the lesson concerns the calculation of the surface and volume of a church tower comprising a cylindrical tower and conical roof.[2] The students in the video-vignette calculated the surface and volume, first individually and then in pairs, while the teacher walked through the class helping them. An example for a closed rating scale item concerning this video-vignette asks the participants whether most students participated in the lesson, testing the selective perception of participants regarding classroom events. As given by the expert rating, the correct answer is *fully correct*, since nearly every student is involved in the lessons. All

[2] Only one video-vignette is described here, as the instrument and its video-vignettes have already been characterized in earlier publications in detail (Kaiser et al., 2015).

A church tower is composed of a cylinder and a cone sitting on it. Both have a radius of 5 m. The cylinder has a height of 25 m, the height of the cone is 12 m. Determine the surface and volume of this composite body.



The task used (see left) can be varied in such a way that on the one hand the question becomes more realistic and on the other hand modeling skills are required to complete the task.

Enter a variation of the task in question according to these demands in keywords.

Fig. 1 Open-response example item concerning the facet decision-making

other options—*partially correct*, *partially incorrect*, and *not correct at all*—were coded as incorrect. Every rating-scale item used these same four response options.

At the end of the video-vignette, the calculation of one student, which contained typical student errors connected to the unreflective usage of the algorithms for surface calculations of solids, is shown. Among other tasks, the study participants had to identify and describe these errors in items concerning the noticing facet of interpretation. Two other open-response items (see Fig. 1) challenged them to vary the task set by the teacher in such a way that it has a stronger connection to the real world and at the same time requires modeling skills to be completed. For a coding as correct for the real-world item, participants had to change the task in such a way that it obtained extra-mathematical relevance, like speaking about refurbishing the church tower. For the item concerning modeling, it was necessary to include parts that deal with the transition between reality and mathematics, like a needed simplified mathematical model, before starting to calculate. Both items would correspond to the noticing facet of decision-making, as the participants had to design a suitable task variation to foster certain competencies. Asking for task variation relates the tasks to mathematics pedagogy and school practice. Moreover, mathematics pedagogical content knowledge is required in relation to realistic task setting and mathematical modeling.

2.3 Scaling and data analysis

The data analysis comprised the following steps for all samples. First, the open-response items were coded according to the coding manual's rubrics. Two independent raters coded 20% of the test booklets, which means they randomly selected 20% of participants' response to all items, to establish intercoder agreement and ensure reliability. As Table 3 indicates, Cohen's kappa values for intercoder agreement were good for all items (Landis & Koch, 1977): between the four studies, the average mean ranged from .80 to .91, and the kappa values ranged from .42 to 1.00. However, few items (four in TEDS-V and seven in TEDS-V-T) scored an intercoder reliability below 0.61, and at least substantial agreement ($\kappa > 0.61$) was achieved for nearly every item. The displayed kappa values for TEDS-V and TEDS-V-T were adjusted. For these studies, one (TEDS-V) and five values (TEDS-V-T) were excluded, because these items showed poor intercoder reliability, particularly due to low incidence of correct answers. These items were again discussed by the raters and then coded by consensus. The final calculation of Cohen's kappa, for which descriptive results are provided in Table 3, excluded these items.

Table 3 Intercoder reliability

Study	κ_{mean}	κ_{min}	κ_{max}
TEDS-FU	.86	.76	1.000
TEDS-I	.91	.70	1.000
TEDS-V	.85	.42	1.000
TEDS-V-T	.80	.47	1.000

κ=Cohen's kappa. Intercoder agreement was investigated for all 34 open-response items and each study. Values above 0.4/0.6/0.8 indicate moderate/substantial/almost perfect agreement (Landis & Koch, 1977). Values for TEDS-V and TEDS-V-T were adjusted

The data were scaled using an item response theory (IRT) approach with the software *ConQuest Version 4.5.2* (Wu et al., 1997). First, a missing response was counted as not administered to estimate item parameters only on the basis of valid answers. Second, missing responses were recoded as wrong answers for ability estimation.[3] To validate the hypothesis that three scales for teacher noticing facets were required, a corresponding three-dimensional IRT model was calculated and compared with a one-dimensional IRT model that would not differentiate into noticing facets. To facilitate reading and interpretation, ability estimates obtained through scaling analyses were linearly transformed to a mean of 50 and a standard deviation of 10. Further analysis was conducted using chi-square tests, Fisher's exact test, and Scheffé post hoc tests to investigate significant differences between expertise groups. Measurement-error-free correlation analyses were carried out within the IRT model to examine associations between the facets.

3 Results

We will now present the results in two sections, where each section addresses one of our research questions.

3.1 Empirical separation of the facets of noticing and its association

A one-dimensional Rasch model was compared with a three-dimensional model to investigate the hypothesis of three separate scales. Two established indices of fit—the ratio of chi-square to degrees of freedom and the Akaike information criterion (AIC)—were then calculated (Table 4). They revealed a significantly better fit of the three-dimensional model for the data and confirmed the validity of separating the noticing construct into three facets.[4]

As Table 5 indicates, the three-dimensional model estimates the three facets with at least acceptable reliability. The item-total correlation of this model varied between .12 and

[3] Missing amount per item ranged from 2 to 16% with a mean of 5%. Open-response items showed significantly ($p < .001$) more missings than rating-scale items. Participants had 5% missings on average with a range from 0 to 48%. Master's students displayed 10% missings on average, and early career teachers and experienced teachers both displayed 3%. Hence, master's students produced significantly ($p < .001$) more missings than teachers.

[4] For χ^2, χ^2/df, and AIC, the better model is shown by a lower value. To deepen the model comparison, additional calculations using structural equation modeling were computed and are displayed in the Appendix.

Table 4 Fit indices for both model approaches

Model	χ^2	χ^2/df	AIC	Δ AIC
One-dimensional	5,047.27	73.15	39,247	
Three-dimensional	4,698.45	70.13	39,192	−55

AIC difference is relative to the one-dimensional model. The difference of the models' deviances is significant with $p < 0.001$. *AIC* Akaike information criterion

Table 5 Scale reliability for all three facets

Reliability measure	Reliability		
	Perception	Interpretation	Decision-making
EAP	0.704	0.800	0.621
WLE	0.663	0.796	0.489

EAP (expected as posteriori) and WLE (weighted likelihood estimates) are reliability measures provided by the software *ConQuest* (Wu et al., 1997) and can be interpreted similar to Cronbach's alpha

.50 with an acceptable mean of .31. The weighted mean square, which is part of the fit statistics and has an expected value of 1.0, ranged from .90 to 1.10 with a mean of 1.00. Values between .50 and 1.50 indicate a productive fit (Linacre, 2002).

The variation of item difficulty in relation to the individuals' abilities (Fig. 2) showed good estimation of teacher noticing abilities, although difficult items were somewhat underrepresented.

The data illustrate that the three facets of noticing conceptualized as perception, interpretation, and decision-making can be empirically separated and measured with at least acceptable reliability for the different reliability measures (except WLE for decision-making) using separate scales (Table 5).

Overall, the empirical results indicate substantial associations between the three facets of teacher noticing. The latent correlations between perception and interpretation ($r = 0.814$) and those between interpretation and decision-making ($r = 0.815$) are relatively strong and can be evaluated as a large positive effect or association (Döring & Bortz, 2016). A medium correlation between perception and decision-making ($r = 0.462$) was also identified (see Table 6). Thus, all three measured facets can be understood as associated, which empirically supports statements regarding the interrelated character of the noticing construct in the noticing discussion (e.g., Sherin et al., 2011b). The high correlation between perception and interpretation emphasizes the strong connection between these two noticing facets as perception is a cognitive, sense-based process, which can only be done in a limited way without interpretation. Although efforts were made to construct items that measured only one facet, namely perception or interpretation, it is hardly possible to adequately differentiate between them. In addition, the strong correlation between interpretation and decision-making can be interpreted as further justification for the inclusion of an action-related facet in the concept of teacher noticing and the theoretical intertwining of both facets.

Overall, this means that the three facets are indeed empirically separable, although the correlations between them are relatively high, particularly between perception and interpretation as well as between interpretation and decision-making. Whether group dependency

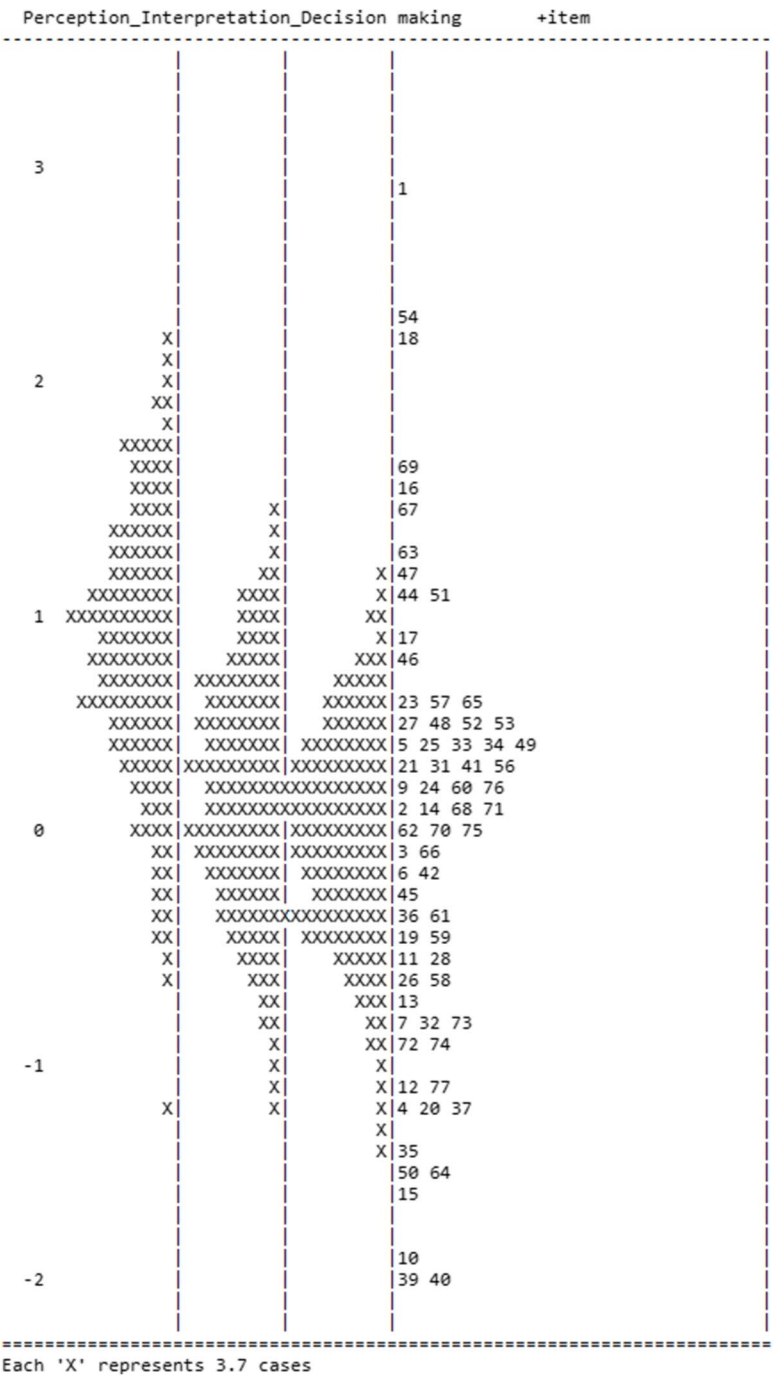


Fig. 2 Variation of the item difficulty in relation to participants' ability for all three scales. Figure is presented here as given by *ConQuest* (Wu et al., 1997). Personal ability values were transformed to $M = 50$ and $SD = 10$ afterwards. One X represents 3.7 cases. Numbers 1 to 77 represent the respective items.

Table 6 Variance (on the diagonal), covariance (above the diagonal), and correlation (below the diagonal) for all three facets

Facet	Variance/covariance/correlation		
	Perception	Interpretation	Decision-making
Perception	.455	.329	.159
Interpretation	.814	.359	.248
Decision-making	.462	.815	.258

Correlations were effectively corrected for any attenuation caused by measurement error

or group-specific patterns characterize the correlations was not investigated in the analyses of this study and needs to be subject of further studies.

3.2 Development of teacher noticing across different teacher groups of expertise

To address the second research question on the differences in teacher noticing and its facets between groups of pre- and in-service teachers, a cross-sectional comparison was conducted for each facet across the groups.

Significant patterns can be identified for all three noticing facets (Table 7). For all facets, significant mean differences in scores emerged as indicated by the Scheffé post hoc test from the group of master's students to the early career teachers, indicating a possible increase in teacher noticing average scores. However, the mean score of the experienced teacher group is slightly lower than that of the early career teachers but clearly higher than that of the master's student mean.

These results suggest that noticing skills progressed significantly from the group of master's students to early career teachers, demonstrating the importance of teaching practice for the development of expertise in noticing. Furthermore, stagnation at least or even small regression is noted in the comparison of early career teacher noticing and experienced teacher noticing. After nearly five years of professional teaching practice, early career

Table 7 Mean scores by expertise group and facet

Variable	Expertise groups						Test of group differences		
	MS (n = 110)		ECT (n = 193)		ET (n = 154)		Odds ratio	p	Scheffé post hoc test[b]
	Mean	SD	Mean	SD	Mean	SD			
Perception[a]	45.57	11.08	51.84	9.19	50.07	8.90	101.69	<.001	ET > MS, ECT > MS
Interpretation[a]	43.58	12.01	52.62	8.44	50.99	8.45	127.90	<.001	ET > MS, ECT > MS
Decision-making[a]	44.04	11.21	53.11	8.36	50.27	9.29	94.46	<.001	ET > MS, ECT > MS, ECT > ET

MS master's students, *ECT* early career teachers, *ET* experienced teachers, *SD* standard deviation. For the test of group differences, a Fisher's exact test was conducted, since the assumptions for a χ^2 test were violated

[a]M = 50 and SD = 10

[b]All given differences in the Scheffé post hoc test are statistically significant ($p < .05$). More statistical information to the Scheffé post hoc test is given in the Appendix

Table 8 *Linear regression of perception, interpretation, and decision-making on teaching experience*

Variable	Perception			Interpretation			Decision-making		
	b	*β*	*p*	*b*	*β*	*p*	*b*	*β*	*p*
Teaching experience	−.14	−.16	<.01	−.18	−.22	<.001	−.18	−.21	<.001
Intercept	51.06	5.62	<.001	51.90	6.13	<.001	51.85	5.85	<.001
R^2		.025			.048			.044	

teachers seem to perform better than experienced teachers with 20+ years of professional experience. However, as shown by the Scheffé post hoc test, these results only were significant for the decision-making facet. To investigate the connection between teaching experience and teacher noticing even further, the team conducted linear regression analyses with teaching experience as independent and each of the three noticing facets as dependent variables using the software *Mplus Version 8.6* (Muthén & Muthén, 2021). Given that the master's students had no regular extensive teaching practice yet, only the data of the in-service teachers, that is, the group of early career and experienced teachers, were included. As shown in Table 8, the regression reveals a small, statistically significant ($p < .01$) decline with increasing teaching experience for all three facets ($\beta_{\text{perception}} = -.16$, $\beta_{\text{interpretation}} = -.22$, $\beta_{\text{decision-making}} = -.21$)[5]. The group of teachers with low experience was somewhat overrepresented, as participants of the TEDS-FU study with low experience represented a distinct part of the sample, which violated the assumptions of the regression. A regression excluding the group of early career teachers revealed similar significant findings.

4 Summary, discussion, and conclusions

The present study aimed to investigate two questions—whether the concept of noticing can be assessed analytically using three facets and whether there are mean differences in teacher noticing among different groups of mathematics teachers with different amounts of professional teaching experience.

Concerning the first research question, the study findings highlight that noticing can be measured analytically and it is possible to measure the three facets of noticing—perception, interpretation, and decision-making—separately through reliable scales based on the rigorous standards of quantitative educational research. However, the results highlight that the three facets of teacher noticing are highly intercorrelated, as was suggested theoretically in the noticing discussion. In particular, the facets of perception and interpretation are strongly connected, which is not unexpected as it is difficult to perceive without interpretation. Furthermore, the strong association between interpretation and decision-making was anticipated as noticing has been described as a knowledge-based construct in the literature. Decision-making has been conceptualized as developing proposals for further continuations of the lesson or alternative ideas, which should be based on a sound interpretation of the noticed situation. The construct of noticing can therefore be divided into three different facets that are strongly connected as theoretically expected and explainable. Furthermore, a new facet, decision-oriented, emerged, which has the

[5] Scatter plots are displayed in the Appendix.

potential to enrich the theoretical construct of noticing and offers an empirical underpinning for our theoretical framework.

The second research question focused on differences in noticing among different pre- and in-service teacher groups with different professional teaching experience and likely also different levels of expertise. Comparisons between the different groups, ranging from master's students to practicing teachers with different degrees of teaching experience, allowed us to assume a progression, which can be interpreted as evidence for the relevance of professional teaching experience. However, the data only show that master's students were significantly outperformed by the two other groups, whereas experienced teachers did not outperform early career teachers. This pattern was identified for all three facets, indicating an underlying pattern.

One reason for the slight decrease in noticing skills from the cross-sectional evaluated groups of early career to experienced teachers may be a "falling-off of skill" (Forde & McMahon, 2019, p. 142) among experienced teachers, which might be explained by the fact that teachers with over 20 years of experience are less likely to participate in professional development (Pedder et al., 2008). Additionally, the expectation of linear growth could be challenged, given that, as mentioned by Berliner (1988), only some teachers attain proficiency and even fewer will attain expert status. The slight regression might indicate that teaching experience in years may contribute to expertise along with other factors, such as working in teacher education or the promotion of strong students' achievements (Palmer et al., 2005), which may vary across both practicing teacher groups. These findings may also substantiate that, as Patterson (2019, p. 4) hypothesizes, teacher personal professional development "may not be linear or truly cyclic but may have stops, starts, digressions, and regressions at various times." However, this may also provide evidence for a general, non-individual-linked pattern of development and raises questions about the long-term effects of teacher education (Liu & Phelps, 2020).

This result is, on the one hand, unexpected, referring to the findings of Jacobs et al. (2010) and Yang et al. (2021b) who interpret their findings as a monotonic or nearly linear growth in noticing between the various expertise groups covering similar teacher groups. The difference in the results reported by Jacobs et al. (2010) may be attributed to the fact that in this study, teacher educators were included in the final stage. The study by Yang et al. (2021b) is embedded in a cultural background—Chinese teachers coming from an Eastern educational paradigm—that differs from this study of German teachers reflecting a Western educational paradigm. The strong focus on content-related and pedagogical aspects has led to the description of newly graduated teachers in China as "semi-finished products" (Paine et al., 2003, p. 216), who will need to learn and develop their educational skills when they enter schools as recent graduates. New graduates are accompanied by mentors, whose role is to impart lessons to the new teachers and work to improve individual lessons or teaching units (Lu et al., 2020). By contrast, the seminar-based second phase of teacher education in Germany (the so-called induction phase) introduces university graduates to teaching on a theoretical basis, observing joint lessons, and theory-oriented phases and topics and aims to induct the candidates into teaching practice from a reflective perspective (Landesinstitut für Lehrerbildung und Schulentwicklung, 2015, 2019). Overall, these different ways of inducting new teachers into teaching may lead to different growth patterns across the three expertise groups.

The findings of this study also align with other empirical studies of teacher knowledge. Evidence suggests a strong increase in GPK during initial teacher education (e.g., König, 2013) and a flattening of the growth after entering the second phase of teacher education (Germany) or the teaching profession as early career teachers (Austria) (König et al., under review). Furthermore, the results of a study by Kleickmann et al. (2013) show a significant increase in MCK and MPCK for teachers from the beginning of initial teacher education to the end of the second phase of teacher education (induction phase). However, their study results describe stagnation or decrease tendencies for MCK and also only slight progression or stagnation for MPCK in both groups when early career teachers and experienced teachers are compared. Both studies report longitudinally interpreted cross-sectional data. Similar results are presented by Blömeke et al. (2014), referring to longitudinal data from the TEDS-M and TEDS-FU Study.

To summarize, the global stagnation or small decrease effects between both groups of practicing teachers indicate that teaching practice is not the only decisive factor in the development of teacher noticing, and the mere extension of teaching experience does not automatically yield higher achievements in noticing. The results provided clear evidence for the necessity of professional development, which would allow more senior teachers to learn more about recent didactical approaches (e.g., Lipowsky, 2019).

Finally, the study limitations must be acknowledged. The analyses presented here were based on convenience samples and cross-sectional data, and caution is urged regarding the development of teacher noticing. Furthermore, the homogeneous TEDS-FU sample, which constitutes most of the early career teachers, may have advantaged this group. One reason for the regression between early career and experienced teachers may be that the sample of early career teachers from TEDS-FU was relatively homogenous and possibly cognitively more selective than the other groups (Blömeke et al., 2014).

Furthermore, apart from the linear regressions, which regarded more variance, only three different expertise groups were compared, making it difficult to establish differences by the levels of expertise outlined by Berliner (1988). Overall, further studies with more differentiated groups are needed to better understand the development of expertise in teacher noticing. In particular, the comparison of groups with different cultural backgrounds is necessary, as the results suggest cultural influences on the development of teachers professional noticing. These differences must be addressed if Eastern and Western cultures wish to cooperate and learn from one another as has occurred in the past (Kaiser & Blömeke, 2013).

Appendix

Examination of the noticing conceptualization with structural equation modeling (in addition to Table 4)

To deepen the model comparison, additional calculations using structural equation modeling were computed. We compared a one-dimensional model that used all 70 items as indicators for the latent trait teacher noticing with a three-dimensional model, where the respective items were indicators for the three latent traits perception, interpretation and decision-making. The calculation was implemented with *Mplus Version 8.6* (Muthén &

Table 9 *Fit indices for both model approaches using MLR estimators*

Fit indices	One-dimensional model	Three-dimensional model	Δ
AIC	36,531	36,470	−61
BIC	37,133	37,085	−48
Adjusted BIC	36,670	36,612	−58

Differences are relative to the one-dimensional model. For AIC, BIC, and adjusted BIC, the better model is shown by a lower value. *AIC* Akaike information criterion, *BIC* Bayesian information criterion. Adjusted BIC were adjusted by sample size

Table 10 *Fit indices for both model approaches using WLSMV estimators*

Fit indices	One-dimensional model	Three-dimensional model	Δ
χ^2	3,269.868	3,234.898	−45.483[a]
RMSEA	.029 [.027, .032]	.029 [.026, .031]	---
CFI	.646	.658	.012

Differences are relative to the one-dimensional model. χ^2 chi-square test of model fit, *RMSEA* root mean square error of approximation, *CFI* comparative fit index. For RMSEA the 90% confidence interval is given in brackets. Cut-off for RMSEA is .05

[a]Chi-square test for difference testing used the diff-test option. Therefore, the difference does not result from both individual values

Muthén, 2021) using the robust maximum likelihood estimators MLR (Table 9). It states the better fit of the three-dimensional model for all common information criterions.

Since the chi-square test could not be computed with MLR estimators due to a too large frequency table, weighted least square mean and variance adjusted estimators (WLSMV) were also applied to conduct a chi-square test for differences and calculate the root mean square error of approximation (RMSEA) and other fit indices to investigate the model fit (Table 10). Both models fit the data, but the chi-square test for model differences shows clear advantages of the three-dimensional model. Comparative fit index (CFI) shows only poor fit, which was expected since it punishes complex models with large number of indicators relative to sample size such as were used in our IRT approach. Besides, as we examined our data in a confirmatory context, for evaluating the models, we prefer RMSEA following recommendations by Rigdon (1996).

Linear regression for perception, interpretation, and decision-making by teaching experience (in addition to Tables 7 and 8)

For all three noticing facets, linear regressions with the noticing facet as dependent variable and teaching experience in years as independent variable were conducted. Since the master's students had no regular extensive teaching practice yet, only the data of the in-service teachers, i.e., the group of early career and experienced teachers, were included.

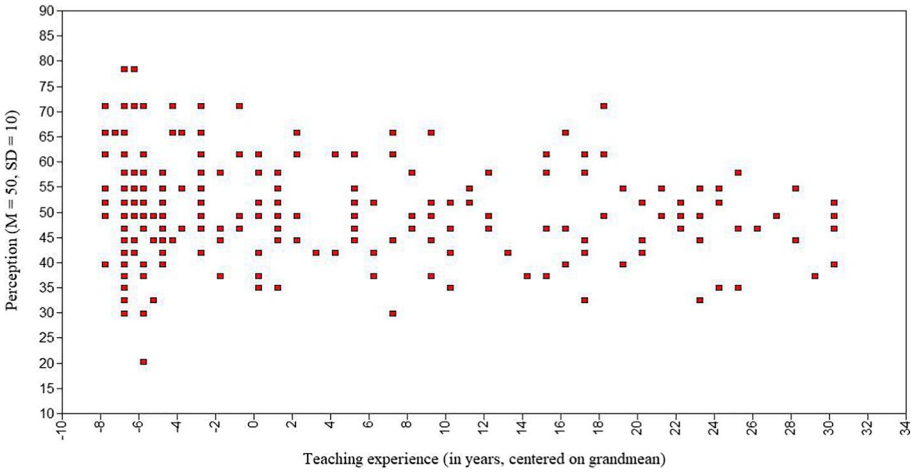


Fig. 3 Scatter plot of perception (Y) vs. teaching experience (X). Teaching experience is centered on the mean ($M = 11.28$), which, therefore, functions as zero in years of teaching experience.

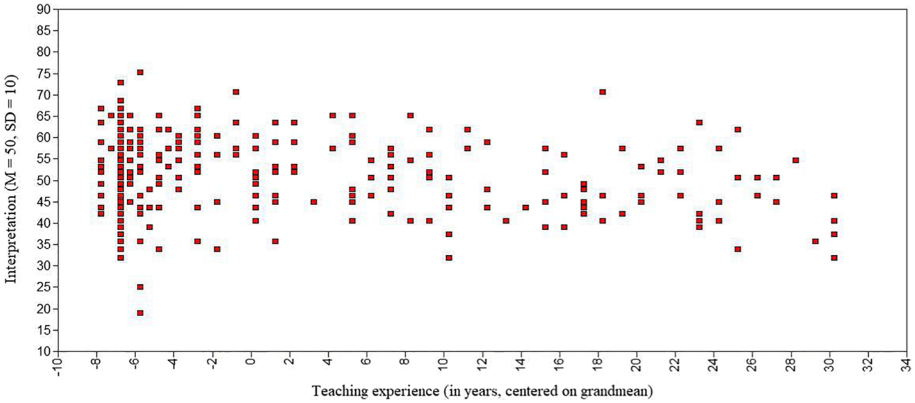


Fig. 4 Scatter plot of interpretation (Y) vs. teaching experience (X). Teaching experience is centered on the mean ($M = 11.28$) which, therefore, functions as zero in years of teaching experience.

As reported in the paper, small, significant degression could be identified for each facet. Consult Figs. 3, 4, and 5 for the scatter plots of the regressions.

Additional data for the Scheffé post hoc test (in addition to Table 7)

We present all statistical data for the Scheffé post hoc test in Table 11.

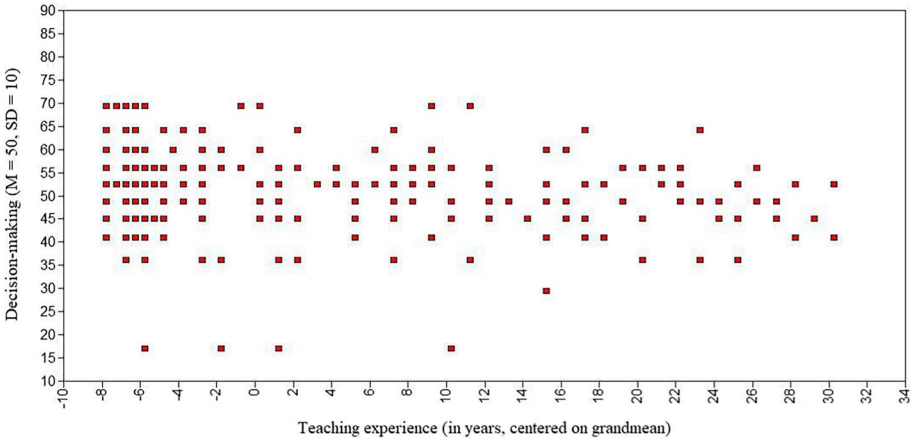


Fig. 5 Scatter plot of decision-making (Y) vs. teaching experience (X). Teaching experience is centered on the mean ($M = 11.28$) which, therefore, functions as zero in years of teaching experience.

Table 11 *Additional data for the Scheffé post hoc test*

Group a	Group b	Mean difference (a-b)	SE	*p*	95 % CI
Perception					
MS	ECT	−6.27	1.15	<.001	[−9.08, -3.46]
MS	ET	−4.5	1.20	.001	[−7.44, −1.56]
ECT	ET	1.77	1.04	.233	[−.77, −4.31]
Interpretation					
MS	ECT	−9.03	1.13	<.001	[−11.80, −6.27]
MS	ET	−7.41	1.18	<.001	[−10.30, −4.52]
ECT	ET	1.62	1.02	.282	[−.88, 4.12]
Decision-making					
MS	ECT	−9.07	1.13	<.001	[−11.84, −6.31]
MS	ET	−6.24	1.18	<.001	[−9.13, −3.35]
ECT	ET	2.83	1.02	.021	[.33, 5.33]

MS master's students, *ECT* early career teachers, *ET* experienced teachers, *SE* standard error, *CI* confidence interval

Literature

Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *3*(4), 369-379.

Funding Open Access funding enabled and organized by Projekt DEAL. TEDS-Follow-Up: German Research Foundation (grant number: KA 797/18-1)
TEDS-Validate: Federal Ministry of Education and Research (grant number: 01PK15006A/B)
TEDS-Validate-Transfer: Federal Ministry of Education and Research (grant number: 01PK19006A/B)

Declarations

Conflict of interest The authors declare no competing interests.

References

Amador, J. M., Bragelman, J., & Superfine, A. C. (2021). Prospective teachers' noticing: A literature review of methodological approaches to support and analyze noticing. *Teaching and Teacher Education*, *99*, 103256. https://doi.org/10.1016/j.tate.2020.103256

Berliner, D. C. (1988). *The development of expertise in pedagogy*. American Association of Colleges for Teacher Education.

Berliner, D. C. (1994). Expertise: The wonders of exemplary performance. In J. N. Mangieri & C. Collins (Eds.), *Creating powerful thinking in teachers and students* (pp. 141–186). Holt.

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). Expert teachers: Their characteristics, development and accomplishments. In R. Batllori i Obiols, A. E. Gomez Martinez, M. Oller i Freixa, & J. Pages i Blanch (Hrsg.), *De la teoria… a l'aula: Formacio del professorat ensenyament de las ciències socials* (pp. 13–28). Departament de Didàctica de la Llengua de la Literatura I de les Ciències Socials, Universitat Autònoma de Barcelona.

Blömeke, S., Busse, A., Kaiser, G., König, J., & Suhl, U. (2016). The relation between content-specific and general teacher knowledge and skills. *Teaching and Teacher Education*, *56*, 35–46.

Blömeke, S., Gustafsson, J. E., & Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift für Psychologie*, *223*(1), 3–13.

Blömeke, S., König, J., Busse, A., Suhl, U., Benthien, J., Döhrmann, M., & Kaiser, G. (2014). From teacher education into profession: Subject-related knowledge as prerequisite for perception, analysis and reaction in teaching [Von der Lehrerausbildung in den Beruf: Fachbezogenes Wissen als Voraussetzung für die Wahrnehmung, Analyse und Reaktion im Unterricht]. *Zeitschrift für Erziehungswissenschaft*, *17*, 509–542.

Blömeke, S., Suhl, U., Kaiser, G., & Döhrmann, M. (2012). Family background, entry selectivity and opportunities to learn: What matters in primary teacher education? An international comparison of fifteen countries. *Teaching and Teacher Education*, *28*, 44–55.

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

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

Dindyal, J., Schack, E. O., Choy, B. H., & Sherin, M. G. (2021). Exploring the terrains of mathematics teacher noticing. *ZDM Mathematics Education*, *53*(1), 1–16. https://doi.org/10.1007/s11858-021-01249-y

Dominguez, H. (2019). Theorizing reciprocal noticing with non-dominant students in mathematics. *Educational Studies in Mathematics*, *102*(1), 75–89.

Döring, N., & Bortz, J. (2016). *Research methods and evaluation in the social and human sciences [Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften]* (5th ed.). Springer. https://doi.org/10.1007/978-3-642-41089-5

Dreher, A., & Kuntze, S. (2015). Teachers' professional knowledge and noticing: The case of multiple representations in the mathematics classroom. *Educational Studies in Mathematics*, *88*(1), 89–114. https://doi.org/10.1007/s10649-014-9577-8

Erickson, F. (2011). On noticing teacher noticing. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Studies in mathematical thinking and learning. Mathematics teacher noticing* (pp. 17–34). Routledge.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363–406.

Forde, C., & McMahon, M. (2019). *Teacher quality, professional learning and policy: Recognising, rewarding and developing teacher expertise*. https://doi.org/10.1057/978-1-137-53654-9

Friesen, M., & Kuntze, S. (2021). How context specific is teachers' analysis of how representations are dealt with in classroom situations? Approaching context-aware measure for teacher noticing. *ZDM Mathematics Education*, *53*(1), 181–193. https://doi.org/10.1007/s1185.8-020-01204-3

Gold, B., & Holodynski, M. (2017). Using digital video to measure the professional vision of elementary classroom management: Test validation and methodological challenges. *Computers & Education*, *107*, 13–30. https://doi.org/10.1016/j.compedu.2016.12.012

Goodwin, C. (1994). Professional vision. *American Anthropologist*, *96*(3), 606–633.

Hoth, J., Schwarz, B., Kaiser, G., Busse, A., König, J., & Blömeke, S. (2016). Uncovering predictors of disagreement: Ensuring the quality of expert ratings. *ZDM Mathematics Education*, *48*(1-2), 83–95. https://doi.org/10.1007/s11858-016-0758-z

Huang, R., & Li, Y. (2012). What matters most: A comparison of expert and novice teachers' noticing of mathematics classroom events. *School Science and Mathematics*, *112*(7), 420–432. https://doi.org/10.1111/j.1949-8594.2012.00161.x

Hughes, R., & Huby, M. (2001). The application of vignettes in social and nursing research. *Journal of Advanced Nursing*, *37*(4), 382–386. https://doi.org/10.1046/j.1365-2348.200202100.x

Ivars, P., Fernández, C., & Llinares, S. (2020). A learning trajectory as a scaffold for pre-service teachers' noticing of students' mathematical understanding. *International Journal of Science and Mathematics Education*, *18*, 529–548. https://doi.org/10.1007/s10763-019-09973-4

Jacobs, V. R., Lamb, L. L. C., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education*, *41*(2), 169–202 www.jstor.org/stable/20720130

Kaiser, G., & Blömeke, S. (2013). Learning from the eastern and the western debate: The case of mathematics teacher education. *ZDM Mathematics Education*, *45*(1), 7–19.

Kaiser, G., Blömeke, S., König, J., Busse, A., Döhrmann, M., & Hoth, J. (2017). Professional competencies of (prospective) mathematics teachers—Cognitive versus situated approaches. *Educational Studies in Mathematics*, *94*(2), 1–22. https://doi.org/10.1007/s10649-016-9713-8

Kaiser, G., Busse, A., Hoth, J., König, J., & Blömeke, S. (2015). About the complexities of video-based assessments: Theoretical and methodological approaches to overcoming shortcomings of research on teachers' competence. *International Journal of Science and Mathematics Education*, *13*(2), 369–387. https://doi.org/10.1007/s10763-015-9616-7

Kaiser, G. & König, J. (2019). Competence measurement in (mathematics) teacher education and beyond: Implications for policy. *Higher Education Policy, 32*, 597–615. https://doi.org/10.1057/s41307-019-00139-z

Keppens, K., Consuegra, E., Goossens, M., de Maeyer, S., & Vanderlinde, R. (2019). Measuring pre-service teachers' professional vision of inclusive classrooms: A video-based comparative judgement instrument. *Teaching and Teacher Education*, *78*, 1–14. https://doi.org/10.1016/j.tate.2018.10.007

Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S. & Baumert, J. (2013). Teachers' content knowledge and pedagogical content knowledge. *Journal of Teacher Education*, *64*(1), 90–106. https://doi.org/10.1177/0022487112460398

König, J. (2013). First comes the theory, then the practice? On the acquisition of general pedagogical knowledge during initial teacher education. *International Journal of Science and Mathematics Education*, *11*(4), 999–1028. https://doi.org/10.1007/s10763-013-9420-1

König, J. (2015). Measuring classroom management expertise (CME) of teachers: A video-based assessment approach and statistical results. *Cogent Education*, *2*(1), 991178.

König, J., Blömeke, S., Jentsch, A., Schlesinger, L., Felske, C., Musekamp, F., & Kaiser, G. (2021). The links between pedagogical competence, instructional quality, and mathematics achievement in the lower secondary classroom. *Educational Studies in Mathematics*, *107*, 189–212.

König, J., Ligtvoet, R., Klemenz, S., & Rothland, M. (under review). Discontinued knowledge growth: On the development of teachers' general pedagogical knowledge at the transition from higher education into teaching practice.

König, J., Santagata, R., Scheiner, T., Adleff, A.-K., Yang, X., & Kaiser, G., (under review). Teacher noticing: A systematic literature review on conceptualizations, research designs, and findings on learning to notice.

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

Landesinstitut für Lehrerbildung und Schulentwicklung. (2015). *Curriculum for education in mathematics. [Ausbildungscurriculum Mathematik.]* https://li.hamburg.de/contentblob/5887082/698b68ec837e795ca8b5daa4364a7b50/data/pdf-ac-mathe.pdf

Landesinstitut für Lehrerbildung und Schulentwicklung. (2019). *Framework for teacher education and qualification. [Referenzrahmen Lehrkräfteausbildung und Anpassungsqualifizierung.]* https://li.hamburg.de/contentblob/12779726/f628678e4a1c59.b9614951ddc71d21c6/data/pdf-referenzrahmen-19-08.pdf

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159. https://doi.org/10.2307/2529310

Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgement tests: A review of recent research. *Personnel Review*, *37 (4)*, 426–441. https://doi.org/10.1108/00483480810877598

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, *16*(2), 878.

Lipowsky, F. (2019). How do findings from research get into classroom? – Impulses from professional development research [Wie kommen Befunde der Wissenschaft in die Klassenzimmer? –Impulse der Fortbildungsforschung]. In C. Donie, F. Foerster, M. Obermayr, A. Deckwerth, G. Kammermeyer, G. Lenske, M. Leuchter, & A. Wildemann (Eds.), *Grundschulpädagogik zwischen Wissenschaft und Transfer* (pp. 144–161). VS Verlag.

Liu, S., & Phelps, G. (2020). Does teacher learning last? Understanding how much teachers retain their knowledge after professional development. *Journal of Teacher Education*, *71*(5), 537–550. https://doi.org/10.1177/0022487119886290

Louie, N. L. (2018). Culture and ideology in mathematics teacher noticing. *Educational Studies in Mathematics*, *97*(1), 55–69.

Louie, N. L., Adiredja, A. P., & Jessup, N. (2021). Teacher noticing from a sociopolitical perspective: The FAIR framework for anti-deficit noticing. *ZDM Mathematics Education*, *53*(1), 95–107. https://doi.org/10.1007/s11858-021-01229-2

Lu, X., Kaiser, G., & Leung, F. (2020). Mentoring early career mathematics teachers from the mentees' perspective—A case study from China. *International Journal for Science and Mathematics Education*, *18*, 1355–1374.

Mason, J. (2002). *Researching your own practice: The discipline of noticing*. Routledge.

Mason, J. (2011). Noticing: roots and branches. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Studies in mathematical thinking and learning. Mathematics teacher noticing* (pp. 35–50). Routledge.

Mason, J. (2015). Responding in-the-moment: Learning to prepare for the unexpected. *Research in Mathematics Education*, *17*(2), 110–127. https://doi.org/10.1080/14794802.2015.1031272

Meschede, N., Fiebranz, A., Möller, K., & Steffensky, M. (2017). Teachers' professional vision, pedagogical content knowledge and beliefs: On its relation and differences between pre-service and in-service teachers. *Teaching and Teacher Education*, *66*, 158–170. https://doi.org/10.1016/j.tate.2017.04.010

Muthén, L. K., & Muthén, B. O. (2021). *Mplus 8 (Version 8.6) [computer software]*. Muthén & Muthén http://statmodel.com

Paine, L. W., Fang, Y., & Wilson, S. (2003). Entering a culture of teaching. In E. Britton, L. Paine, D. Pimm, & S. Raizen (Eds.), *Comprehensive teacher induction: Systems for early career learning (pp. 20–82)*. Kluwer.

Palmer, D. J., Stough, L. M., Burdenski, T. K., & JR., & 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

Patterson, C. (2019). *Enacted personal professional learning: Re-thinking teacher expertise with story-telling and problematics*. Springer. https://doi.org/10.1007/978-981-13-6007-7

Pedder, D., Storey, A., Opfer, V. D., & Wolfenden, F. (2008). *Schools and continuing professional development (CPD) in England - State of the Nation research project*. Cambridge University and The Open University.

Piwowar, V., Barth, V. L., Ophardt, D., & Thiel, F. (2018). Evidence-based scripted videos on handling student misbehavior: The development and evaluation of video cases for teacher education. *Professional Development in Education*, *44*(3), 369–384. https://doi.org/10.1080/19415257.2017.1316299

Praetorius, A. K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of three basic dimensions. *ZDM Mathematics Education*, *50*(3), 407–426. https://doi.org/10.1007/s11858-018-0918-4

Reuker, S. (2017). The noticing of physical education teachers: A comparison of groups with different expertise. *Physical Education and Sport Pedagogy*, *22*(2), 150–170. https://doi.org/10.1080/17408989.2016.1157574

Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modelling: A Multidisciplinary Journal*, *3*(4), 369–379. https://doi.org/10.1080/10705519609540052

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.(28(1))

Santagata, R., König, J., Scheiner, T., Nguyen, H., Adleff, A.-K., Yang, X., & Kaiser, G. (2021). Mathematics teacher learning to notice: A systematic review of studies of video-supported teacher education. *ZDM Mathematics Education*, *53*(1), 119–134.

Schack, E. O., Fisher, M. H., & Wilhelm, J. (2017). *Research in mathematics education. teacher noticing: Bridging and broadening perspectives, contexts, and frameworks*. Springer.

Scheiner, T. (2021). Towards a more comprehensive model of teacher noticing. *ZDM Mathematics Education*, *53*(1), 85–94. https://doi.org/10.1007/s11858-020-01202-5

Schlesinger, L., Jentsch, A., Kaiser, G., König, J., & Blömeke, S. (2018). Subject-specific characteristics of instructional quality in mathematics education. *ZDM Mathematics Education*, *50*(3), 475–490.

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

Sherin, M. G., Jacobs, V. R., & Philipp, R. A. (2011a). *Studies in mathematical thinking and learning: Mathematics teacher noticing*. Routledge.

Sherin, M. G., Jacobs, V. R., & Philipp, R. A. (2011b). Situating the study of teacher noticing. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Studies in mathematical thinking and learning: Mathematics teacher noticing* (pp. 3–13). Routledge.

Sherin, M. G., & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, *60*, 20–37.

Shulman, L. (1986). Those who understand: knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14.

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

Steffensky, M., Gold, B., Holdynski, M., & Möller, K. (2015). Professional vision of classroom management and learning support in science classrooms—Does professional vision differ across general and content-specific classroom interactions? *International Journal of Science and Mathematics Education*, *13*(2), 351–368. https://doi.org/10.1007/s10763-014-9607-0

Stigler, J. W., & Miller, K. F. (2018). Expertise and expert performance in teaching. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & M. A. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (2nd ed.pp. 431–452). Cambridge University Press. https://doi.org/10.1017/9781316480748.024

Stockero, S. L. (2021). Transferability of teacher noticing. *ZDM Mathematics Education*, *53*(1), 73–84. https://doi.org/10.1007/s11858-020-01198-y

Stockero, S. L., & Van Zoest, L. R. (2013). Characterizing pivotal teaching moments in beginning mathematics teachers' practice. *Journal of Mathematics Teacher Education*, *16*, 125147.

Tatto, M.T., Schwille, J., Senk, S., Ingvarson, L., Rowley, G., Peck, R., Bankov, K., Rodriguez, M. & Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. IEA.

van Es, E. A. (2011). A framework for learning to notice student thinking. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Hrsg.), *Studies in mathematical thinking and learning. Mathematics teacher noticing* (pp. 134–151). Routledge.

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*, 571–596.

van Es, E. A., & Sherin, M. G. (2021). Expanding on prior conceptualizations of teacher noticing. *ZDM Mathematics Education*, *53*(1), 17–27. https://doi.org/10.1007/s11858-020-01211-4

Wu, M. L., Adams, R. J., & Wilson, M. R. (1997). *ConQuest: Multi-aspect test software (Version 4.5.2) [computer program]*. Australian Council for Educational Research.

Yang, X., Kaiser, G., König, J. & Blömeke, S. (2021a). Relationship between Chinese mathematics teachers' knowledge and their professional noticing. *International Journal of Science and Mathematics Education*, *19*, 815–837. https://doi.org/10.1007/s10763-020-10089-3

Yang, X., König, J.,& Kaiser, G. (2021b). Growth of professional noticing of mathematics teachers: A comparative study of Chinese teachers noticing with different teaching experience. *ZDM Mathematics Education*, *53*(1), 29–42. https://doi.org/10.1007/s11858-020-01217-y

Authors and Affiliations

Anton Bastian[1] · Gabriele Kaiser[2] · Dennis Meyer[3] · Björn Schwarz[4] · Johannes König[5]

Anton Bastian
anton.bastian@uni-hamburg.de

Dennis Meyer
dennis.meyer@uni-hamburg.de

Björn Schwarz
bjoern.schwarz@uni-vechta.de

Johannes König
johannes.koenig@uni-koeln.de

[1] University of Hamburg, Von-Melle-Park 8, 20146 Hamburg, Germany

[2] University of Hamburg, Australian Catholic University, Von-Melle-Park 8, 20146 Hamburg, Germany

[3] University of Hamburg (ProfaLe), Bogenallee 11, 20144 Hamburg, Germany

[4] University of Vechta, Driverstraße 22, 49377 Vechta, Germany

[5] University of Cologne, Gronewaldstraße 2a, 50931 Cologne, Germany