ZDM - Mathematics Education
https://doi.org/10.1007/s11858-025-01745-5

ORIGINAL PAPER

Trends in expertise studies in the domain of teaching

Helen Jossberger[1] · **Henny P.A. Boshuizen**[2] · **Hans Gruber**[1] · **Stefan Krauss**[3]

Received: 1 November 2024 / Accepted: 23 September 2025

Abstract
Education permeates the entire society and without doubt what teachers do matters. Therefore, it is no surprise that researchers, practitioners, and policy makers share an interest in understanding outstanding performance and are eager to learn more about possibilities to support expertise development in the domain of teaching. Since the very beginning, expertise research is intrigued by the extraordinary cognitive processing of experts. In this contribution, we revisit the roots of expertise research and identify challenges when transferring this paradigm to the teaching profession. A literature review including 40 empirical studies is presented aiming at providing a historical overview and identifying trends in expertise research. Several commonalities, systematic differences and trends are apparent regarding the selection of participant groups, the role of the domain, the authenticity of setting and task, macro- and micro-level measurement, the development of tools and methods, and dynamics in research goals including merging of research paradigms. The discussion of the findings includes a plea for future directions of expertise research.

Keywords Expertise research · Expert teacher · Teaching · Expertise development

1 Introduction

The work and effectiveness of teachers play a crucial role in society and economy. No doubt, what teachers do matters, but they can vary in major ways having positive or negative influences on learners (Hattie, 2009). Many teachers work at an average level, some below average. What about those who perform at an outstanding level? Can they be identified? And what insights can be gained from their outstanding performance for learning and professional development purposes in teaching?

From the very beginning expertise research was intrigued by outstanding performance and tried to capture differences between experts and novices in a particular domain. The Cambridge Dictionary (2024) defines an expert as "a person with a high level of knowledge or skill relating to a particular subject or activity". According to Ericsson (2018), expertise refers to "the characteristics, skills, and knowledge that distinguish experts from novices and less experienced people" (pp. 3–4). If a domain has objective criteria to identify experts, and a person can show superior reproducible performance on representative tasks of that field, the term expert performance is used (Ericsson, 2018). Therefore, it is not surprising that expertise research has its roots in chess, a domain with very clear performance criteria (e.g., Elo rating scale; Elo, 1986). Obviously, the domains of chess and teaching differ enormously. Still, in the 1980s, two prominent research groups at the University of Pittsburgh (Leinhardt and colleagues) and University of Arizona (Berliner and colleagues) clearly departed from the expertise studies in chess and applied the study design to investigate teachers' cognitive processing. Instead of chess positions as stimulus material, they used photographic slides of classroom sessions. Their work marks an important starting point for expertise research in teaching.

In the last decades, research related to expertise has flourished in the domain of teaching and other professional fields. Interestingly, from around the 2000s, another line of research developed in parallel. Especially in the German literature, the concept of "expertise" has been severely pushed back by the concept of "professional teacher competence". Although Kunter et al. (2013) describe the professional competence approach as a further development of

✉ H. Jossberger
helen.jossberger@ur.de

1 Faculty of Human Sciences, University of Regensburg, Regensburg, Germany

2 Open Universiteit, Heerlen, The Netherlands

3 Faculty of Mathematics, University of Regensburg, Regensburg, Germany

Published online: 02 October 2025

the expertise paradigm, the approaches clearly differ. The quantitative, mainly correlational, large-scale studies such as COACTIV (Kunter et al., 2013) and TEDS-M (Blömeke et al., 2020, 2022; Kaiser et al., 2017) focus on the investigation of interindividual differences in professional competence of mathematics teachers with a variety of standardised psychometric instruments. Expertise research, on the contrary, rather examines individual cognitive processes and their development that enable a teacher to act competently in the classroom (Krauss, 2020; Lachner et al., 2016).

A discussion about expertise research in teaching adds to the scientific debate about the impact of long-lasting practice activities for intraindividual expertise development through case-based knowledge restructuring. In this contribution, we aim to explore how expertise studies in the domain of teaching have evolved over time and what trends can be identified. Before presenting our literature review, we start off with revisiting the roots of expertise research to highlight underlying premises and goals of this paradigm. Then we look more closely at the domain-specific characteristics of teaching to inform researchers and practitioners about some inherent challenges of expertise research in this domain. Based on the review, we identify trends and raise awareness of potential blind spots. Conclusions are drawn about the current relevance of the expertise paradigm for teachers' professional development. This includes a plea for future directions of expertise research.

2 Revisiting the roots of expertise research

Chess was the first intensively studied domain in expertise research. We believe that revisiting these roots is helpful to understand the decisions taken later and the associated challenges when transferring expertise research to the domain of teaching. In his seminal work, De Groot (1946/1978) analysed in much detail the professional action of the world-best chess players and related it to the performance of proficient chess players at a lower level. He used a think-aloud technique in a move choice task: The participants were asked to think aloud about their next move in a presented chess position. It turned out that the experts were able to immediately grasp a huge amount of information from the presented chess position, relate it to their own experience, embed the information in evaluative processes, and were able to make fast and valid decisions. As a result, their move choice was impressively superior to the move choice of good club players.

In a second part of his studies, de Groot explored in more detail the memory performance of experts. This part became even more influential in later research. It was partially replicated in the ground-breaking studies by Chase and Simon (1973) and Simon and Chase (1973). Those studies revealed that chess masters can memorise a chess position almost perfectly after viewing it for only a few seconds. This extraordinary recall performance inspired researchers to investigate its roots in more detail. It was revealed that due to years of extensively playing and reflectively analysing chess matches, complex and highly organised mental representations are constructed that allow one to access complex perceptual information fast and unconsciously. Chase and Simon (1973) phrase this complex combination of perception, pattern recognition and memory organisation "The Mind's Eye in Chess". They describe the mind's eye as "the meeting point where visual information from the external world is combined and coordinated with visual representations stored in short-term and long-term memory" (p. 277). Similar findings about the mind's eye were found in many other domains, including academic disciplines like physics or computer science. It became clear that what matters are not general search heuristics or an unspecific memory capacity that characterise an expert, but rather domain-specific knowledge and experience (Ericsson et al., 2007; Grabner et al., 2007; Gruber, 1994) that affect knowledge restructuring, retention, and retrieval (Boshuizen et al., 2020a). This is a crucial finding in expertise research and already points to several underlying premisses and goals that are essential when trying to grasp expertise and its development. We briefly resume them here.

At least two goals can be distinguished in expertise research. One is to capture and explain expert performance (e.g., how do they think, perceive, perform) and the second one is to understand the trajectory of expertise development. These goals imply that the expertise paradigm is fundamentally optimistic, because it goes hand in hand with the premise that outstanding performance develops and can be influenced by environmental factors (e.g., feedback, support). Giftedness was not fully disputed, but got a less prominent role, although contentious nature-nurture debates about the impact of innate and environmental factors on outstanding performance are still being held (Ericsson, 2021; Hambrick et al., 2020). Motivation is an important driver to fully master the domain and bring it forward. It helps individuals to persevere and therefore gives rise to expertise. Accumulated amounts of deliberate practice were found to explain around 50% of the variance in performance, which is substantial but still leaves room for other impact (Ericsson et al., 1993; Macnamara & Maitra, 2019; Platz et al., 2014). From an educational perspective and the expertise paradigm point of view, however, it is promising that high levels of performance are not limited to unique innate talents that cannot be changed (Ericsson et al., 2007). On the contrary, understanding what characterises experts and how they attained their initial beginning levels of achievement is seen as a route "to develop practice environments and foster learning methods that help people to attain the fundamental

representations of the tasks and the self-regulatory skills that were necessary for the prospective experts to advance their learning to higher levels" (Ericsson, 2018, p. 18).

Another premise of expertise research is its focus on the individual and the key intraindividual changes in characteristic features. Expertise research rejects the assumption of stable cognitive knowledge structures. Instead, knowledge structures and memory organisation are ever-changing. Intraindividual changes occur throughout long-term practical experiences. Knowledge is restructured continuously in and through practice and the individual case-based knowledge is conceptualised as situated enabling pattern recognition (Boshuizen et al., 2020a,b).

This leads to another premise namely that expertise is domain specific. Features and requirements of the domain shape expertise and expertise development expressed in knowledge structures as well as cognitive and perceptual processing. This premise draws attention to the dynamic aspects of professional workplaces that can involve major changes during an expert's career (Boshuizen et al., 2020b). In addition, it also emphasises how domains that are characterised by comparable tasks still result in very different knowledge structures (e.g., highly specialised knowledge scripts). It also affects in how far findings from one expertise domain can be generalised to another domain (Boshuizen et al., 2020a).

Top performance or in other words superior reproducible performance on representative tasks is focussed and should be distinguished from satisfying performance. The domain specifies what the important tasks are and how performance can be assessed objectively to identify experts based on clearly identifiable superior performance or specific characteristics (cf. ELO rating in chess). Not general, but domain-specific performance and long-term practical experience are examined. Typically, the expert-(intermediate)-novice paradigm is used to investigate differences in processing representative domain-specific tasks (e.g., describing cognitive structures, examining cognitive and perceptual processes, analysing learning and professional development, or revealing influencing environmental factors). Usually, samples in empirical research are small, because individuals are often unique and have to be carefully chosen based on several criteria. Research studies in various fields have confirmed that experts outperform novices (e.g., concerning problem-solving, strategy use, perception) in their specific domain (see Reimann, 1998, for more detail and references to primary sources).

Attempts to replicate or transfer early expertise studies to professional domains like medicine started in the 1970s and steadily continued with a main focus on diagnostic decision-making and clinical reasoning (Claessen & Boshuizen, 1985; Elstein et al., 1978; Jaarsma et al., 2017; Lesgold et al., 1988; Schmidt et al., 1990) and expertise research in other professional domains such as business followed (Boshuizen et al., 2004). Challenges were identified when it was tried to transfer the expertise paradigm and the proposed theory of the "the mind's eye" to rather ill-defined domains like history or teaching (Voss & Wiley, 2006). In the following, we turn to the domain-specific characteristics of teaching and the challenge of identifying experts in this field.

3 Transferring expertise research to the teaching profession

3.1 The dynamic workplace of teachers and inherent challenges of the domain

Teaching is a versatile task, confined in space and time. It is performed in interaction and cooperation with students of various skills, backgrounds, and motivations. Many processes in the classroom have a life of their own. Teaching phases, unexpected events, and intended effects can be influenced by the teacher, but cannot be fully controlled or precisely predicted in terms of their course or the time of their occurrence (Nückles, 2021). Cognitive, perceptual, social as well as pedagogical and didactical processes need to go hand in hand.

Leinhardt and Greeno (1986) characterise teaching as a complex cognitive skill involving planning and decision-making (e.g., deciding to proceed or to repeat based on students' readiness for new material, asking questions or providing support). Planning and making appropriate decisions require teachers to simultaneously perceive, monitor, and interpret interactional events happening in class. This ongoing processing demands situational awareness and visual attention, which are influenced by teachers' knowledge (e.g., classroom management, lesson structure, subject matter) (Wolff et al., 2021). In addition to knowing what might be difficult for learners, the teacher needs to shape and manage the whole process within the constraints of formal education. Moreover, interpersonal processes and dynamics (e.g., human relations, emotional dimensions, participation patterns) need to be considered to deduce – or even implicitly sense – the most appropriate pedagogical and didactical strategies for individual learners as well as for the whole class (Wright, 2005).

A "good outcome" per session is often hard to determine as it is relative to context and students' prior learning. These characteristics make it difficult to define consistent superior performance, contrary to expertise research in domains like chess, which are relatively stable in time and across contexts allowing the measurement of unambiguous outcomes (Van Tartwijk et al., 2023).

3.2 The challenge of choosing an expert in the domain of teaching

Clearly, the described domain-specific characteristics of teaching make it more challenging to identify an expert in this professional field where objective criteria are disputable or even lacking. Two distinct views emerged: performance-oriented ("expert teacher") (Berliner, 1986) versus knowledge-oriented ("teacher as expert") (Bromme, 1992/2014). Both views belong to what Chi (2006) calls the "relative" approaches meaning that expertise is a level of proficiency that novices in principle can achieve. This contrastive approach is used to compare experts to novices and intermediates. An important distinction between the two views is the operationalisation of expertise and thus the selection of the "expert".

3.2.1 The "expert teacher": focus on top performance

Driven by the search for the "expert teacher" – referring to persons who are identified based on their high performance – the research by Berliner (1986, 2001, 2004) is exemplary for expertise research in teaching. Different criterion-based performance indicators (e.g., level of education/training, professional success in competition-oriented domains, years of work experience, performance measures in the professional field, quality of work experience) and normative indicators (e.g., nominations or evaluations by colleagues, supervisors, and/or students) for expertise have been used in research (Palmer et al., 2005), but the question raised by Berliner (2001) is how to have surety that true "expert teachers" are selected. He distinguished between "good" and "successful" teaching and cautioned that these characteristics of expertise are often confused: "Good teaching is judged through reliance on standards applied to the tasks of teaching and related to norms for professional behaviour, including moral considerations. Successful teaching is about whether intended learnings were achieved" (p. 468). Both aspects are considered relevant when trying to define an expert teacher.

National educational reforms in the United States of America intending to improve student learning pushed the teaching profession and expertise research forward. In 1987, the National Board for Professional Teacher Standards was established aiming at (1) maintaining high and rigorous standards for what accomplished teachers should know and be able to do; (2) providing a national voluntary system certifying teachers who meet these standards; and (3) advocating related education reforms to integrate National Board Certification in American education and to capitalise on the expertise of National Board Certified Teachers (National Board for Professional Teacher Standards, n.d.). The validity study by Bond et al. (2000) tried to identify good and successful teachers and created objective criteria for designating expert teachers. In total, 13 prototypical features of expertise (e.g., better perception of classroom events, better monitoring of learning and providing feedback to students) were developed and assessed. In a recent meta-summary, Anderson and Taner (2023) even identified 73 features across six areas from cognition to practice that expert teachers frequently share and proposed an expert teacher prototype. Bond et al. (2000) capture successful teaching by measuring the outcomes of instruction including higher motivation to learn and higher feelings of self-efficacy, deeper understanding of the subject matter, and higher levels of achievement. The empirical work showed that Board Certified teachers in comparison to those that failed to meet the Board standards on the assessments, excelled on each prototypical feature. The data on student outcomes is less conclusive. For students' motivation, only a few differences were revealed, but the quality of students' work samples was found to be higher when taught by Board Certified teachers (Berliner, 2001; Bond et al., 2000).

If researchers are truly interested in top performance, then individuals in line with the "expert teacher" should be chosen. Convincingly in this respect are studies focussing on a group of teachers with very high expertise (National Board Certified teachers), in which also student achievement as a function of teacher expertise was investigated (Belson & Husted, 2015; Laczko-Kerr & Berliner, 2002; Vandevoort et al., 2004). Only longitudinal studies with careful measures that aggregate over students and time allow us to examine students' progress. Interestingly, these studies compared teachers who were certified within different programmes and standards rather than applying the novice-(intermediate)-expert paradigm. Comparing these certified teachers with novices or intermediates would have certainly resulted in even larger differences.

3.2.2 The "teacher as expert": focus on professional experience

When Bromme (1992/2014) refers to a "teacher as expert", he stresses the aspect of a professional – "a person who is specialised and masters a specific task; related to a professional context" (p. 37) – rather than a comparison between weak and strong performance. It concerns an experienced teacher, not necessarily a top performer. Teachers acquire a stock of professional knowledge that is needed to cope with the task. This knowledge is not stable but changes throughout practice and is shaped by the workplace. Knowledge becomes integrated, encapsulated, and script formation takes place allowing to process and act faster and more accurately (Boshuizen et al., 2020a; Moberg et al., 2025; Wolff et al., 2021). Primarily, the knowledge-oriented view of expertise is concerned with the content and structure of knowledge

in relation to cognitive and perceptual processes. Both the quality and the quantity of domain-specific expert knowledge are crucial. Knowledge underlies and accompanies perception, thinking and action; it is considered a prerequisite to cope with complex professional requirements. The patterns of beginners' and experts' eye movements and representations of the classroom suggest that beginners' attention is largely drawn to forceful but probably irrelevant features of the classroom whereas experts use their knowledge of classes and situations to direct their gaze to places where things might happen. Obviously, knowledge patterns include temporal information related to events and, thus, can be used to create predictions. Due to the time characteristics of classroom events, however, such knowledge structures need to be activated and run just-in-time, which requires a high degree of automaticity (Wolff et al., 2021).

In comparison to the performance-oriented view, the operationalisation of expertise and the selection of teachers for research studies appears to be easier, because the most important criterion-based indicator is professional teaching experience rather than the very precise delineation of different levels of expertise. Certainly, more indicators can be considered such as the level of training or professional successes.

4 The literature review: aim and research question

Revisiting the roots of expertise research and relating these to the domain of teaching has clarified several differences and pertinent challenges that need to be considered when designing empirical studies that intend to be in line with the expertise paradigm. The presented theoretical considerations motivated the literature review and guided the identification of themes for the analyses. We were particularly interested in the selection of participant groups, the role of the domain, the authenticity of the setting and the task, the development of tools and methods, and the dynamics of the research goals in order to answer the following research question: How did expertise studies in the domain of teaching evolve over time and what trends can be identified?

5 Methodology of the study

5.1 Literature search strategy

As the research led by Berliner and Leinhardt was seen as important starting point for expertise research in teaching, a manual search process was the first strategy to find empirical studies from their teams through forward and backchecks of citations. In a second step, we performed a systematic search in the database *Web of Science*. This interdisciplinary repository was selected as it has high-quality indexing standards, a decent international reputation, and is one of the largest curated scientific, publisher-neutral search systems. In view of our goal, this large database was considered substantial enough for our search to identify trends. Our search process was based on the PRISMA 2020 Checklist as guideline for reporting reviews with a focus on eligibility criteria, information sources, search strategy, selection and data collection process (Page et al., 2021). In three search rounds performed in September 2024, February 2025, and September 2025, we explored several search strings including the keywords expert*, novic*, teach*, class*, empirical, and math*. The broadest search syntax (All fields: expert* AND novic* AND teach*) resulted in 1349 hits. Finally, to capture studies that follow the expertise paradigm, we used a high-precision search syntax for the present analysis: Abstract: expert* AND novic* AND teach* AND class* NOT competence. This search string resulted in 174 hits in September 2025. The third step was another manual search based on the snowballing technique to identify further studies, in which we checked reference lists of studies we had assessed as relevant for addressing our research question.

5.2 Eligibility criteria and selection process

Table 1 presents the inclusion and exclusion criteria that we specified for the review and used for selecting studies that address our research question.

As we were interested in providing a historical overview and identifying trends, no restriction on publication year was applied or any additional refinement. Studies had to be of adequate quality (e.g., providing sufficient information about the study design), which we ensured by selecting publications in peer-reviewed journals as a major criterion. All articles had to be open access or accessible via the university licence or the internet. Figure 1 depicts the flow chart of the article selection process. In total, 40 studies were included in our review.

5.3 Critical appraisal

The first and second author screened the 174 hits from the database *Web of Science* individually and then compared their selection based on the abstracts. No automation tools were used. Only in one case, the authors came to a conflicting result, but this disagreement was resolved by rereading the abstract. When in doubt, the abstract was discussed considering the eligibility criteria until agreement was reached. Seven studies that were found manually did not contain the keyword class* in the abstract, but the study design and context was identified as eligible. The oldest study in our review (Leinhardt & Smith, 1985) did not mention novic* in the abstract although an expert-novice comparison was performed. We included it as we consider it as one of the pioneering expertise studies in the domain of teaching.

Table 1 Eligibility Criteria

Inclusion Criteria (IC)	Exclusion Criteria (EC)
IC1: Language of the papers is English	EC1: Language of the papers is not English
IC2: Peer-reviewed empirical research articles	EC2: Editorials, books, book chapters, conference papers, theoretical articles
IC3: Studies following the expertise paradigm and focus on differences between expert and novice teachers	EC3: Studies without expertise differences as we were interested in the contrastive expert-novice paradigm and selection of participant groups
IC4: Studies about teaching, instruction and/or the organisation of it	EC4: Studies with focus on teacher professional development (e.g., co-teaching) or innovation (e.g., use of artificial intelligence)
IC5: Studies with focus on cognitive processing and domain-specific skills in schools of general education as setting	EC5: Studies in higher education (e.g., medical or nursing domain) or vocational education or special education or sports, arts and music
IC6: Full-text articles accessible at no charge (open access, university licence, internet)	EC6: Full-text articles with costs

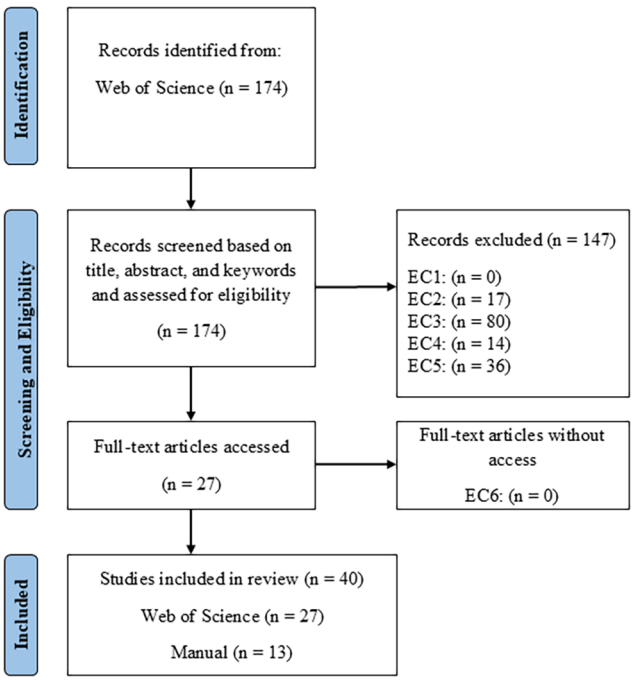

Fig. 1 Flow Chart of the Article Selection Process

The inspection of the outcomes showed that certain groups of researchers, i.e. those who are more reluctant in describing their samples as novices and experts, although they use similar selection criteria, are not included. Missing in the selection are, for instance, the studies by Van den Bogert et al. (2014) and Van Driel et al. (2021, 2023) who prefer terms like student-teachers, pre-service teachers, beginning and experienced teachers. These studies are, however, clearly related to the studies with the mentioned keywords. It is important to note that the high-precision search syntax might have caused some biases in the included studies.

5.4 Analysis of the studies

Based on our theoretical considerations, the selection of participant groups, the role of the domain, the authenticity of the setting and the task, the development of tools and methods, and the dynamics of the research goals were the themes that guided our analysis process. To organise all 40 studies, we arranged the studies in a historical order. The following categories and subcategories were used to examine all studies in more detail: country of data collection; number per expertise level; expertise (levels; criteria for expert selection); effects on/samples of; tasks and settings (action in authentic or simulated context; vignettes, paper and pencil); methods of data collection (real-time or video-based; eye-tracking, brain-based methods; other methods); dependent variables/foci/scoring categories (e.g., knowledge; perception; recall); goals and domain taught (see Appendix for a detailed overview). An abbreviated version of the categorisation is presented in Table 2.

6 Results

In our analysis of the 40 articles that were selected we used several foci derived from the expertise paradigm to cluster studies and see what trends could be discerned. Several commonalities, systematic differences and trends are apparent. The findings are organised according to the identified themes: participant selection; role of the domain; authenticity of setting and task; development of tools and methods; and dynamics in research goals. Based on Table 2, we refer to a study by reporting the assigned number (1-40). Thus, a lower number represents an older study than a higher one. Numbers that are separated by a "/" are based on the same data set.

Table 2 Overview of the Expertise Studies in Teaching

Year	Publication	Country of data collection	Teaching domain	Type of goal
1980s	1. Leinhardt & Smith (1985)	United States of America	Math	filling gap
	2. Carter et al. (1987)	United States of America	Math & Science	filling gap, extending
	3. Carter et al. (1988)	United States of America	Math & Science	further deepening
	4. Leinhardt (1989)	United States of America	Math	specifying
1990s	5. Brown Clarridge & Berliner (1991)	United States of America	Math & Science	further deepening
	6. Sabers et al. (1991)	United States of America	Science	connecting theory
	7. Tochon & Munby (1993)	Switzerland	Language	further deepening
	8. Allen & Casbergue (1997)	United States of America	Science	addressing limitations of previous studies
2000s	9. Torff (2003)	United States of America	Language	updating
	10. Meyer (2004)	United States of America	not stated	further deepening, understanding differences
	11. Krull et al. (2007)	Estonia	Language	identifying indicators for professional development
	12. Hogan & Rabinowitz (2009)	United States of America	not stated	filling gap
2010s	13. Huang & Li (2012)	China	Math	enhance understanding
	14. Cortina et al. (2015)	United States of America	Math	validating
	15. Wolff et al. (2015)	The Netherlands	CM	updating and enhancing understanding
	16. Wolff et al. (2016)	The Netherlands	CM	filling gap
	17. McIntyre et al. (2017)	United Kingdom & Hong Kong	Math, Science, Language, Humanities, other	filling gap, ascertaining, exploring
	18. Wolff et al. (2017)	The Netherlands	see Wolff et al. (2016)	further deepening
	19. McIntyre & Foulsham (2018)	United Kingdom & Hong Kong	see McIntyre et al. (2017)	expanding
	20. McIntyre et al. (2019)	United Kingdom & Hong Kong	see McIntyre et al. (2017)	further deepening
2020s	21. Sun et al. (2020)	China	Math	further deepening, elucidating
	22. Seidel et al. (2021)	Germany	Math	expanding
	23. Stahnke & Blömeke (2021a)	Germany	Math & Biology	further deepening
	24. Stahnke & Blömeke (2021b)	Germany	see Stahnke & Blömeke (2021a)	expanding
	25. Yang et al. (2021)	China	Math	further deepening understanding
	26. Bastian et al. (2022)	Germany	Math	filling gap
	27. Herbst et al. (2022)	United States of America	Math	developing methods, validating, exploring
	28. Keller et al. (2022)	Germany	Math, English, Social Studies & History	further deepening understanding
	29. Liu et al. (2022)	China	Math & Biology	testing and validating
	30. McIntyre et al. (2022)	The Netherlands & United Kingdom	see McIntyre et al. (2017) & Wolff et al. (2016)	validating
	31. Zhao et al. (2022)	China	Math	updating
	32. Huang et al. (2023)	United States of America	Math, Science, Literacy, History & Social Studies	further deepening
	33. Pilous et al. (2023)	United States of America	Math	further deepening

Table 2 (*Continued*)

Year	Publication	Country of data collection	Teaching domain	Type of goal
	34. Siller et al. (2023)	Israel	Math	filling gap
	35. Zhang et al. (2023)	China	Math	developing methods
	36. Kosko et al. (2024)	United States of America	Math	further deepening
	37. Lin et al. (2024)	China	CM	expanding
	38. Tong et al. (2024)	China	Language	filling gap
	39. McIntyre et al. (2025)	United Kingdom & Hong Kong	see McIntyre et al. (2017)	further deepening
	40. Stahnke & Gegenfurtner (2025)	Germany	CM	developing methods

Note: Full references are listed in the electronic appendix. CM = Classroom Management.

Selection of participant groups. Most studies used a combination of a minimal number of years of experience, recognition (according to national criteria), and identification by superiors to select participating experts. Researchers of the group around Berliner tried to identify excellent teachers and did extra effort to improve selection by doing classroom observations (2/3, 6), while Leinhart used students' growth and achievement as an (extra) criterion (1, 4). Most research conducted in China was able to use more objective criteria in expert selection (21, 29, 31, 32, 35, 37, 38). Especially the Chinese system of teacher assessment seems to add much rigour to studies done in this country (see 13 and 31 for a description). The system applies to all teachers, which makes participant selection quite efficient.

Most participating novices are pre-service teachers, though at different stages of their training. This group is more diverse than might be expected as teacher training programmes in different countries vary widely in duration and in amount of time spent on didactic and pedagogical topics as well as practical placements. Most studies compared beginners and experts, but some also included laypersons (postulant teachers; 2/3, 5/6), or intermediates (in-service teachers with between 1 and 5 years of experience; 8, 21, 25, 26, 38).

Role of the domain. Most studies reported the teaching domain involved mathematics, as topic taught by the participants or in the vignettes presented. Other studies concerned science teachers (6, 8), EFL (38), history (9), but also mixed groups were used (2/3/5, 17/19/20/30/39, 23/24, 28, 32) or domain was not even stated (10, 12). A separate group of studies explicitly focussed on classroom management (15/16/18, 37, 40) and also used a mix of teaching domains. The role of the domain varied in these studies. The domain could be the central theme of research into teachers' macro- and micro-activities, such as in some noticing studies (13, 25, 26) or the studies that focussed the development of teacher content and pedagogical content knowledge or professional knowledge structure (1, 33). In the other studies, the teaching domain rather provided the context, such as in classroom management studies. But also, in studies that used a mix of domains more generic tasks could be focussed, such as recall of information as a prerequisite for reflection (3), perception of causes of student behaviour (5), representation of classroom behaviour (12), gaze type (17, 32), observing students' behavioural engagement (22), noticing of classroom management (24). Together the studies show an interesting spectrum of domain-specific and more generic tasks investigated.

Authenticity of setting and task. Twenty of the 40 studies used authentic tasks of which 18 were conducted in authentic settings. Most authentic tasks concerned teacher actions or macro-activities such as 'teach a class' or 'prepare a lesson'. Yet data sets varied in obtrusiveness of the data acquisition conditions. For instance, study 29 used video recordings that had been collected routinely for observational and communicational purposes; at the other extreme of the continuum teachers taught a lesson wearing a mobile eye-tracker, which could have interfered with their own and their students' behaviour. These tasks could be supplemented by measures focussing micro-activities elicited through methods such as stimulated think-aloud or free-recall protocols (the basis for the measurement of different aspects and components of underlying information processing, knowledge application, representation and thinking processes). For instance, the studies 30/39 that combined the task of teaching a class while wearing a mobile eye-tracker and provide stimulated think-aloud based on the eye-tracker footage with eye-movements made visible. Analysis of these verbalisations provided insights into perceptual and cognitive processing of the classroom events. Other, less frequently used methods included card sorting (1), document analysis (10), similarity judgment (12), profile matching (22), which give more insight into the structure of knowledge applied but are more remote from daily teaching practice. The other 20 studies applying less authentic tasks and settings used (more

or less immersive) visual or text materials in combination with mostly micro-activities evoked through specific open, forced choice or rating questions to measure perceptual and cognitive processing, and knowledge structure and application. These studies provided less rich data sets but are able to test more specific hypotheses and can accommodate large samples. An example of such a study (26) was conducted in the context of the TEDS-projects. It used scripted video vignettes, aimed to get deeper insights into teacher noticing and the factors affecting that, and included 457 participants.

Development of tools and methods. The development of new tools such as eye trackers gave a boost to expertise research. Older studies had few options to investigate the perceptual processes involved in expert and novice performance, although study 6 tried to investigate classroom monitoring behaviour by playing classroom videos over three displays positioned next to each other. The host of European studies (e.g., 15, 16, 17) started when eye-tracking apparatus became more feasible and prices dropped, in combination with the more powerful computer systems that allowed more complex analyses. This development opened opportunities to investigate 'old' research questions at a more micro-level. Interestingly, the use of eye-tracking led to unexpected insights into how experts and novices differed in what they attended to and how they systematically monitored the classroom. Mobile eye-tracking has a similar potential as it allows teachers freely moving around in their classes (14, 19/20/30/39, 28, 32). Yet, analysis is still an issue; problems associated with the coordination of teacher location, head orientation and tilt, and eye movements need to be solved to enable a similar breakthrough.

The most recent technical development is seen in the studies that use brain-based measures (to measure brain synchronisation in dyads (21); the use of EEG-waves and brain locations (37)), which were used in two Chinese studies. As the researchers of study (37) state, this type of method can again expand the expertise development paradigm as it can differentiate between aspects of event perception and detection at an even more detailed level.

Dynamics in research goals. Where the projects in the 1980s used extensive combinations with various tasks, methods and measures to investigate effects of expertise, more recent studies apply more streamlined and dedicated methods. And where researchers in the early years reported 'gaps' in research done thus far, a 'paradigm that is still under development', to apply a paradigm, approach or theory 'borrowed from another domain', research following later reported 'previous findings and their limitations' and 'detailed hypotheses' as the starting points or aims of the research. We also see studies investigating the measurement qualities of instruments and categorisation systems coming up, suggesting that the field approached a situation of shared tools and concerns. And in the last decennium we see studies coming up that seem to merge the theoretical and methodological approach on noticing which focussed traditionally on teacher students with the expertise paradigm (13, 24, 25, 26, 28, 36).

A last trend that deserves broader attention in the expertise community is the cultural and workplace component. Expertise is shaped by the environment in which it develops. The studies comparing UK and Hong Kong teachers showed different gaze patterns (19/20/30/39); the studies from the TEDS research programme (25, 26) done in separate countries but using parallel instruments and definitions suggest that teacher workplace requirements and qualification systems affect continuous expertise development. These types of macro effects have long remained out of sight in the mind's eye metaphor of expertise but are of great importance from a societal perspective.

7 Discussion

The goal of this study was to identify trends by exploring how expertise studies have evolved over time. To do so, we focussed on the selection of participant groups, the role of the domain, the authenticity of the setting and the task, the development of tools and methods, and the dynamics of the research goals.

Identifying the "expert teacher", who is good and successful (Berliner, 2001) is a difficult endeavour and seems to be dependent in part on the policy of a country. In the United States of America and more recently especially in China certification levels exist that are assessed either on a voluntary basis or inherent in the system. In European countries, for instance in Germany, certification levels do not exist, making it more difficult to select teachers based on objective criteria. We see a tendency to select homogeneous groups using the best available combination of criteria, in the absence of hard indicators. It is striking that the effects on students' learning are almost absent in these studies. Only longitudinal studies with careful measures that aggregate over students and time allow to examine students' progress. In line with capturing superior reproducible performance (Ericsson, 2018) of teachers, a longitudinal study design also has added value as it can rule out measurement errors due to cross-sectional one-shot data collection. Ultimately, however, it is essential to focus on advancing the profession and supporting learners.

As teacher training programmes in different countries vary widely in duration and amount of time spent on didactic and pedagogical topics as well as practical placements, the selection and description of the sample needs to be documented clearly. In fact, however, Caspari-Sadeghi and König (2018) report how inconsistently teachers are described and selected. Based on their review, they recommend that the

educational community should substantiate operational definitions and identification of an "expert teacher" and other expertise levels so that research becomes more informative, reliable, and useful. It is paramount to include more expertise levels in future research, to develop a better understanding of the developmental path of knowledge structures, and how they relate with cognitive and perceptual performance. Deliberate practice was not studied directly in the selected studies, but different professional practices described in the studies (e.g., in China and Germany) seem to result in different performance.

The role of the domain differed in the included studies. The domain can be the central focus or provides the context. The majority of our included studies had to do with mathematics. Yet, research conclusions are drawn at a more generic or abstract level.

The emergence of new tools and methods provide a boost to the research field as they may enable investigation of central questions at a level of detail or from an angle that was not possible without this development. The most noticeable change is the arrival of the eye-tracking methodology, which resulted in more detailed studies of perception and understanding of classroom processes compared to earlier studies using video analysis and verbal reports only or head movement as indication for gaze patterns. The latest shift bears the traces of brain-based methods. However, these methods are still too recent to indicate new patterns emerging. Newer methods and technologies in expertise research such as (mobile) eye-tracking, electroencephalography (EEG) or functional magnetic resonance imaging (fMRI) certainly have benefits (e.g., visualising expertise in brain activity) and caveats (e.g., restricting participants in time and space) (Gegenfurtner et al., 2017). Thus, validation is required before new research questions and relations should be introduced. Some of the included studies focussed on validation indicating that expertise research matures. When charting new territories, it would be good to take stock occasionally and ask as a research community where we are heading and what we overlook.

The dynamics of the goals show a trend of growing maturity of expertise research. The early studies in our review originated from very broad projects that aimed to get a grasp of the field and in recent studies more precise questions and methods are used. However, there might also be a risk that study designs become very specific or complex threatening authenticity in tasks and settings or making analyses less feasible. Moreover, we see interactions with other research paradigms, which can be a way to expand the scope of the studies. In this line are also studies that included factors like culture or that allowed comparisons between diverging workplaces, which stresses the important role of context. These studies draw attention to a phenomenon that is central to expertise development but is hardly focussed on in expertise research namely that the workplace shapes expertise development.

Concluding, teacher education programmes, workplaces, and opportunities of professional development for teachers differ worldwide. The research interest in expert teachers, however, is shared by many countries (Guo, 2024; Yang, 2014). The mind's eye in expertise research can help explain surprising findings in large-scale studies. For instance, the conundrum that pedagogical content knowledge does not increase with years of professional experience (Krauss et al., 2017). One possible explanation could be the lack of deliberate practice in teacher professional development. Another reason might be that the instruments are not sensitive enough to detect knowledge restructuring processes that occur during a professional career (Blömeke et al., 2015). Longitudinal investigations with an intraindividual focus increase our understanding of how individuals develop. The mind's eye in teaching with a strong cognitive focus on the development of individual expertise also needs to be carefully scrutinised, because of the risk of ignoring the social embeddedness of thinking and knowledge in situations, practices, and cultures (Berliner, 2001; Engeström, 2018; Nückles, 2021; Stigler & Miller, 2018; Van Tartwijk et al., 2023).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11858-025-01745-5.

Funding Information Open Access funding enabled and organized by Projekt DEAL.

Declarations

Competing Interests We have no known conflict of interest to disclose.

References

Anderson, J., & Taner, G. (2023). Building the expert prototype: A metasummary of teacher expertise studies in primary and secondary education. *Educational Research and Reviews*, *38*, 100485. https://doi.org/10.1016/j.edurev.2022.100485.

Belson, S. I., & Husted, T. A. (2015). Impact of national board for the professional teaching standards certification on student achievement. *Education Policy Analysis Archives*, *23*(91). https://doi.org/10.14507/epaa.v23.2074.

Berliner, D. C. (1986). In pursuit of the expert pedagogue. *Educational Researcher*, *15*, 5–13. https://doi.org/10.3102/0013189X015007007.

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

Berliner, D. C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology & Society*, *24*, 200–212. https://doi.org/10.1177/o270467604265535.

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

Blömeke, S., Jentsch, A., Ross, N., Kaiser, G., & König, J. (2022). Opening up the black box: Teacher competence, instructional quality, and students' learning progress. *Learning and Instruction*, *79*, 101600. https://doi.org/10.1016/j.learninstruc.2022.101600.

Blömeke, S., Kaiser, G., König, J., & Jentsch, A. (2020). Profiles of mathematics teachers' competence and their relation to instructional quality. *ZDM – Mathematics Education*, *52*, 329–342. https://doi.org/10.1007/s11858-020-01128-y.

Bond, L., Smith, T., Baker, W. K., & Hattie, J. (2000). *The certification system of the National Board for Professional Teaching Standards: A construct and consequential validity study*. Center for Educational Research and Evaluation, University of North Carolina at Greensboro.

Boshuizen, H. P. A., Bromme, R., & Gruber, H. (2004). *Professional learning: Gaps and transitions on the way from novice to expert. Innovation and change in professional education: Vol. 2*. Kluwer. https://doi.org/10.1007/1-4020-2094-5_10.

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

Boshuizen, H. P. A., Vosniadou, S., & Lehtinen, E. (2020b). Conceptual changes for and during working life. *International Journal of Educational Research*, *104*, 101682. https://doi.org/10.1016/j.ijer.2020.101682.

Bromme, R. (2014). *Der Lehrer als Experte: Zur Psychologie des professionellen Wissens [The teacher as expert: The psychology of professional knowledge]*. D. H. Rost (Ed.). Verlag Hans Huber. (Original work published 1992).

Cambridge Dictionary (2024). *Expert*. Retrieved September 18, 2024, from https://dictionary.cambridge.org/.

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

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In *Visual information processing* (pp. 215–281). https://doi.org/10.1016/B978-0-12-170150-5.50011-1.

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

Claessen, H. F. A., & Boshuizen, H. P. A. (1985). Recall of medical information by students and doctors. *Medical Education*, *19*, 61–67. https://doi.org/10.1111/j.1365-2923.1985.tb01140.x.

De Groot, A. D. (1978). *Thought and choice in chess* (2nd ed.). Mouton. (Original work "Het denken van den schaker" published in 1946).

Elo, A. E. (1986). *The rating of chessplayers, past and present* (2nd ed.). Arco.

Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). *Medical problem solving: An analysis of clinical reasoning*. Harvard University Press. https://doi.org/10.4159/harvard.9780674189089.

Engeström, Y. (2018). *Expertise in transition: Expansive learning in medical work*. Cambridge University Press. https://doi.org/10.1017/9781139023009.

Ericsson, K. A. (2018). An introduction to the second edition of the Cambridge handbook of expertise and expert performance: Its development, organization, and content. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (2nd ed., pp. 3–20). Cambridge University Press. https://doi.org/10.1017/9781316480748.

Ericsson, K. A. (2021). Given that the detailed original criteria for deliberate practice have not changed, could the understanding of this complex concept have improved over time? A response to Macnamara and Hambrick (2020). *Psychological Review*, *85*, 1114–1120. https://doi.org/10.1007/s00426-020-01368-3.

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. https://doi.org/10.1037/0033-295X.100.3.363.

Ericsson, K. A., Roring, R. W., & Nandagopal, K. (2007). Giftedness and evidence for reproducibly superior performance: An account based on the expert performance framework. *High Ability Studies*, *18*, 3–56. https://doi.org/10.1080/13598130701350593.

Gegenfurtner, A., Kok, E. M., van Geel, K., de Bruijn, A. B. H., & Sorger, B. (2017). Neural correlates of visual perceptual expertise: Evidence from cognitive neuroscience using functional neuroimaging. *Frontline Learning Research*, *5*, 95–111. https://doi.org/10.14786/flr.v5i3.259.

Grabner, R. H., Stern, E., & Neubauer, A. C. (2007). Individual differences in chess expertise: A psychometric investigation. *Acta Psychologica*, *124*, 398–420. https://doi.org/10.1016/j.actpsy.2006.07.008.

Gruber, H. (1994). *Expertise. Modelle und empirische Untersuchungen [Expertise. Models and empirical studies]*. Westdeutscher Verlag.

Guo, Y. (2024). Improving teacher performance: Evolution and enlightenment of primary and middle school teacher certification institution in China (1978–2023). *European Journal of Education: Research, Development & Policy*, *59*, e12701. https://doi.org/10.1111/ejed.12701.

Hambrick, D. Z., Macnamara, B. N., & Oswald, F. L. (2020). Is the deliberate practice view defensible? A review of evidence and discussion of issues. *Frontiers in Psychology*, *11*, 1134. https://doi.org/10.3389/fpsyg.2020.01134.

Hattie, J. A. C. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Jaarsma, T., Boshuizen, H. P. A., Jarodzka, H., Nap, M., Verboon, P., & van Merriënboer, J. J. G. (2017). Tracks to a medical diagnosis: Expertise differences in visual problem solving. *Applied Cognitive Psychology*, *30*, 314–322. https://doi.org/10.1002/acp.3201.

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*, 161–182, 183–184. https://doi.org/10.1007/s10649-016-9713-8.

Krauss, S. (2020). Expertise-Pradigma in der Lehrerinnen- und Lehrerbildung. In C. Cramer, J. König, M. Rothland, & S. Blömeke (Eds.), *Handbuch Lehrerinnen- und Lehrerbildung [Handbook on teacher training]* (pp. 154–162). Klinkhardt. https://doi.org/10.35468/hblb2020-018.

Krauss, S., Lindl, A., Schilcher, A., Fricke, M., Göhring, A., Hofmann, B., Kirchhoff, P., & Mulder, R. (2017). *Falko – Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik*. Waxmann.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (Eds.) (2013). *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project*. Springer.

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

Laczko-Kerr, I., & Berliner, D. C. (2002). The effectiveness of "Teach for America" and other under-certified teachers on student academic achievement: A case of harmful public policy. *Educational Policy Analysis Archives*, *10*(37). https://doi.org/10.14507/epaa.v10n37.2002.

Leinhardt, G., & Greeno, J. G. (1986). The cognitive skill of teaching. *Journal of Educational Psychology*, *78*, 75–95. https://doi.org/10.1037/0022-0663.78.2.75.

Leinhardt, G., & Smith, D. A. (1985). Expertise in mathematics instruction: Subject matter knowledge. *Journal of Educational Psychology*, *77*, 247–271.

Lesgold, A., Rubinson, H., Feltovitch, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing x-ray pictures. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (1st ed., pp. 311–342). Psychology Press. https://doi.org/10.4324/9781315799681.

Macnamara, B. N., & Maitra, M. (2019). The role of deliberate practice in expert performance: Revisiting Ericsson, Krampe, & Tesch-Römer (1993). *Royal Society Open Science*, *6*, 190327. https://doi.org/10.1098/rsos.190327.

Moberg, J., Hannula-Sormunen, M., Hähkiöniemi, M., & Lehtinen, E. (2025). Analyzing teachers' scripts from teachers' reflections after they tried to encourage students' flexible mathematical thinking. *Frontline Learning Research*, *13*, 45–75. https://doi.org/10.14786/flr.v13i1.1589.

National Board for Professional Teacher Standards (n.d.). *Mission and History*. Retrieved October, 24, 2024 from https://www.nbpts.org/about/mission-history/.

Nückles, M. (2021). Investigating visual perception in teaching and learning with advanced eye-tracking methodologies: Rewards and challenges of an innovative research paradigm. *Educational Psychology Review*, *33*, 149–167. https://doi.org/10.1007/s10648-020-09567-5les.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal*, *372*, n71. https://doi.org/10.1136/bmj.n71.

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*, 13–25. https://doi.org/10.1207/s15326985ep4001_2.

Platz, F., Kopiez, R., Lehmann, A. C., & Wolf, A. (2014). The influence of deliberate practice on musical achievement: A meta-analysis. *Frontiers in Psychology*, *5*, 151–163. https://doi.org/10.3389/fpsyg.201400646.

Reimann, P. (1998). Novizen- und Expertenwissen [Novice and expert knowledge]. In F. Klix & H. Spada (Eds.), *Wissen. Enzyklopädie der Psychologie, Themenbereich C: Theorie & Forschung, Serie II: Kognition* (pp. 325–367). Hogrefe.

Schmidt, H. G., Norman, G. R., & Boshuizen, H. P. A. (1990). A cognitive perspective on medical expertise: Theory and implications. *Academic Medicine*, *65*, 611–621.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist*, *61*, 394–403.

Stigler, J. W., & Miller, K. F. (2018). Expertise and expert performance in teaching. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (2nd ed., pp. 431–452). Cambridge University Press. https://doi.org/10.1017/9781316480748.

Van den Bogert, N., van Bruggen, J., Kostons, D., & Jochems, W. (2014). First steps into understanding teachers' visual perception of classroom events. *Teaching and Teacher Education*, *37*, 208–216. https://doi.org/10.1016/j.tate.2013.09.001.

Vandevoort, L. G., Amrein-Beardsley, A., & Berliner, D. C. (2004). National board certified teachers and their students' achievement. *Education Policy Analysis Archives*, *12*(46). https://doi.org/10.14507/epaa.v12n46.2004.

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

Van Driel, S., Crasborn, F., Wolff, C. E., Brand-Gruwel, S., & Jarodzka, H. (2023). Teachers' interactive cognitions in noticed classroom management events: Does experience matter? *Teaching and Teacher Education*, *126*, 104076. https://doi.org/10.1016/j.tate.2023.104076.

Van Tartwijk, J., van Dijk, E. E., Geertsema, J., Kluijtmans, M., & van der Schaaf, M. (2023). Teacher expertise and how it develops during teachers' professional lives. *International Encyclopedia of Education (Fourth Edition)*, *5*, 170–179. https://doi.org/10.1016/B978-0-12-818630-5.04057-4.

Voss, J. F., & Wiley, J. (2006). Expertise in history. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 569–584). Cambridge University Press. https://doi.org/10.1017/CBO9780511816796.033.

Wolff, C. E., Jarodzka, H., & Boshuizen, H. P. A. (2021). Classroom management scripts: A theoretical model contrasting expert and novice teachers' knowledge and awareness of classroom events. *Educational Psychology Review*, *33*, 131–148. https://doi.org/10.1007/s10648-020-09542-0.

Wright, T. (2005). *Classroom management in language education*. Palgrave MacMillan.

Yang, X. (2014). *Conception and characteristics of expert mathematics teachers in China*. Springer. https://doi.org/10.1007/978-3-658-03097-1.