European Journal of Educational Research
Volume 11, Issue 2, 711 - 737.
ISSN: 2165-8714
http://www.eu-jer.com/

Effects of Generic and Subject-Didactic Teaching Characteristics on Student Performance in Mathematics in Secondary School: A Scoping Review

Carina Spreitzer* University of Klagenfurt, AUSTRIA

Samuel Hafner University of Klagenfurt, AUSTRIA

Konrad Krainer University of Klagenfurt, AUSTRIA

Andreas Vohns University of Klagenfurt, AUSTRIA

Received: November 12, 2021 ▪ Revised: December 25, 2021 ▪ Accepted: January 11, 2022

Abstract: Research on instructional quality has been of great interest for several decades, leading to an immense and diverse body of literature. However, due to different definitions and operationalisations, the picture of what characteristics are important for instructional quality is not entirely clear. Therefore, in this paper, a scoping review was performed to provide an overview of existing evidence of both generic and subject-didactic characteristics with regard to student performance. More precisely, this paper aims to (a) identify both generic and subject-didactic characteristics affecting student performance in mathematics in secondary school, (b) cluster these characteristics into categories to show areas for quality teaching, and (c) analyse and assess the effects of these characteristics on student performance to rate the scientific evidence in the context of the articles considered. The results reveal that teaching characteristics, and not just the instruments for recording the quality of teaching as described in previous research, can be placed on a continuum ranging from generic to subject-didactic. Moreover, on account of the inconsistent definition of subject-didactic characteristics, the category of 'subject-didactic specifics' needs further development to establish it as a separate category in empirical research. Finally, this study represents a further step toward understanding the effects of teaching characteristics on student performance by providing an overview of teaching characteristics and their effects and evidence.

Keywords: *Generic characteristics, instructional quality, mathematics achievement, mathematics instruction, subject-didactic characteristics.*

To cite this article: Spreitzer, C., Hafner, S., Krainer, K., & Vohns, A. (2022). Effects of generic and subject-didactic teaching characteristics on student performance in mathematics in secondary school: A scoping review. *European Journal of Educational Research*, *11*(2), 711-737. https://doi.org/10.12973/eu-jer.11.2.711

Introduction

Regarding quality teaching, Fischer (1998) and Krainer (2005) claimed that the precise meaning of "good" requires a continuous negotiation process linking experiences from school practice with scientific evidence. Similarly, Berliner (2005) stated that "Defining quality in teaching is unusually difficult. Were anyone serious about this issue, they would soon realize that quality is an ineffable concept […]. Defining *quality* always requires value judgments about which disagreements abound" (p. 206). Harvey and Green (2000) declared that a "transformative view of quality is rooted in the notion of 'qualitative chance', a fundamental change of form. […] A quality education is one that affects changes in the participants and, thereby, presumably enhances them" (pp. 24-25). Schratz (2009) and Helmke (2017) called for an emphasis on learner orientation and Fenstermacher and Richardson (2005) defined quality using both the concepts of transformation and learner orientation:

> Quality teaching, we argue here, consists of both good and successful teaching. By *good* teaching we mean that the content taught accords with disciplinary standards of adequacy and completeness, and that the methods employed are age appropriate, morally defensible, and undertaken with the intention of enhancing the learner's competence with respect to the content studied […]. By *successful* teaching we mean that the learner actually acquires, to some reasonable and acceptable level of proficiency, what the teacher is engaged in teaching (p. 191).

In our scoping review, we understand quality teaching in line with Fenstermacher and Richardson (2005). The output of a successful lesson can be the subject-related competence (Weinert et al., 1989) or the motivation of the students.

* **Corresponding author:**
Carina Spreitzer, Department of Instructional and School Development, University of Klagenfurt, Sterneckstraße 15, 9020 Klagenfurt, Austria.
✉ carina.spreitzer@aau.at

"Good" ("successful" in terms of Fenstermacher and Richardson (2005)) teaching should promote students' intrinsic motivation and increase their interest (Krapp, 1999). In the context of mathematics, Kunter (2005) concluded that teaching aims to produce both cognitive and non-cognitive outputs, such as interest in a subject.

Fenstermacher and Richardson (2005) also referred to principles of instructional practices fostering successful teaching. However, no established theory of teaching quality currently exist. The most recent *Handbook of Research on Teaching* (Gitomer & Bell, 2016), lacks a general theory of teaching quality (Praetorius, Klieme, et al., 2020). Praetorius, Klieme, et al. (2020) advocated "producing stronger theories of teaching quality" (p. 30). So far, empirical research on instruction was primary process-product research (Ditton, 2002). The core of this paradigm is the search for teaching characteristics related to good educational outcomes in terms of subject-related student output (Gage & Needels, 1989). In recent years, there has been a large amount of research on what constitutes good teaching, including meta-analyses (e.g., Hattie, 2008; Scheerens & Bosker, 1997) and collections (e.g., Nilsen & Gustafsson, 2016). To synthesise the plethora of findings, theoretical frameworks have been developed (e.g., Seidel & Shavelson, 2007). These range from generic models [constructs referring to all subjects, e.g., the Three Basic Dimensions framework (Klieme et al., 2006) or the Classroom Assessment Scoring System (CLASS) (Pianta & Hamre, 2009)] to subject-didactic models [constructs specifically for one subject, e.g., the Mathematics Quality of Instruction (MQI) (Learning Mathematics for Teaching Project, 2011) or Teacher Education and Development Study in Mathematics (TEDS-M) (Schlesinger & Jentsch, 2016)] (Charalambous & Praetorius, 2018). Considering only the generic models is limited in explaining student performance (Praetorius, Rogh, & Kleickmann, 2020). Klieme and Rakoczy (2008) also preferred considering subject-didactic principles for successful teaching. Bruder (2018) added domain-specific aspects from a mathematical-didactic perspective to reveal the effective development potential of mathematics teaching.

Focusing subject-didactic characteristics, there are increasingly many theoretical frameworks and instruments for lesson observation. For mathematics, in particular, there is a great deal of interest in identifying subject-didactic characteristics that contribute to increased student performance. Except for a few studies (e.g., Kane et al., 2011), analyses of theoretical frameworks focus on the elementary level (e.g., Walkowiak et al., 2014) or on the secondary level (e.g., König et al., 2021). Doig et al. (2005) observed differences in mathematics teaching theories and Perry et al. (2006) identified differences in the beliefs of mathematics teachers between primary and secondary school. Therefore, we advocate the separate synthesis of frameworks for these two domains.

Review Aims

This scoping review intends to provide an overview of existing evidence of both generic and subject-didactic characteristics with regard to student performance in mathematics in secondary school. Systematic reviews provided extensive meta-analyses investigating numerous generic teaching characteristics (e.g., Creemers & Kyriakides, 2008; Scheerens & Bosker, 1997; Seidel & Shavelson, 2007). Many of these focused on specific teaching characteristics, such as duration of teaching (e.g., Dagli, 2019; Schlesinger & Jentsch, 2016; Walshaw & Anthony, 2008; Wayne & Youngs, 2003), and did not limit the review to one subject or grade. The present scoping review examines both generic and subject-didactic characteristics in mathematics teaching in secondary education. It has three aims: (a) identify both generic and subject-didactic characteristics affecting student performance in mathematics in secondary school, (b) cluster these characteristics into categories or areas to show areas for quality teaching, and (c) analyse the effects of these characteristics on student performance to rate the scientific evidences.

The main objective of scoping reviews is to identify existing evidences and to cluster these with some narrative commentary (von Elm et al., 2019). The theoretical framework for the observation of TEDS-Instruct (Schlesinger et al., 2018) provided an orientation for the categories in the clustering process. We chose this approach, because it is based "on a framework widely accepted especially in German-speaking countries, namely the generic framework of three basic dimensions […] which we extended the generic dimensions of instructional quality by additional subject-specific dimensions" (Schlesinger et al., 2018, p. 475) and because it is specifically suitable for instructional quality in secondary schools (Busse, 2015).

Methodology

Literature Review - Research Design

The systematic search was conducted on February 8, 2021 using the literature database Web of Science, which includes only peer-reviewed articles, complemented by a hand search at a later time. The search was carried out using the "Basic Search" mode in the "Topic" field with the following three search terms: 1) mathematics AND (observation OR observations OR observational) AND (classroom OR instrument OR instruments OR teacher OR teaching); 2) mathematics AND quality AND (lesson OR instruction OR instructional OR teaching); 3) mathematics AND ("generic teaching practices" OR "content specific teaching practices" OR "teaching effectiveness"). The results were limited to the 2005–2021 time span, following existing reviews (see Hartmann et al., 2012; Scheerens et al., 2013). Limiting the time period is necessary because of the PRISMA statement (Moher et al., 2009).

The large number of results (2.311 articles) required a pre-selection of articles based on the Web of Science Categories, excluding research fields irrelevant to the topic (e.g., "computer science interdisciplinary applications"). Together with two articles found in a hand search, we disposed of 2.109 articles. Figure 1 provides an overview of the selection process following the PRISMA statement (Moher et al., 2009).

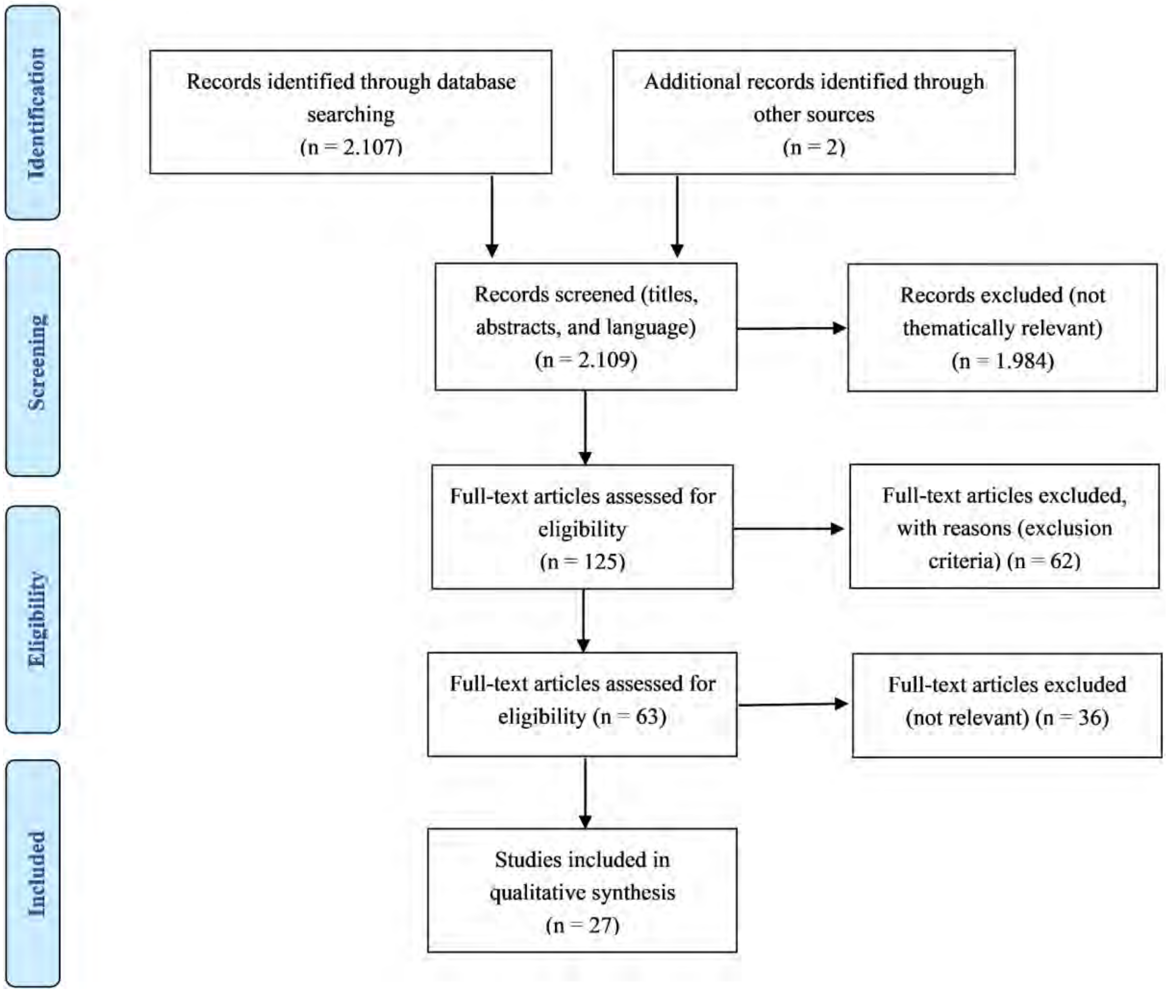

Figure 1. Flow Chart of the Study Selection Process According to PRISMA (Moher et al., 2009)

Article Selection

Two researchers [CS and SH] compiled the final selection of articles in three successive steps, with decisions made by consensus:

Step 1: Assessing titles and abstracts of the 2.109 articles and excluding articles in languages other than English or German, the number of articles was reduced to 125.

Step 2: Screening the 125 full texts applying the inclusion criteria (a) a focus on a sample of teachers, pupils, or other groups in secondary education due to the differences between secondary and elementary education (Doig et al., 2005; Perry et al., 2006); (b) clear inferences of the effects of teaching characteristics on student performance; and (c) a clear description of each characteristic to be able to cluster the characteristics adequately resulted in 63 articles.

Step 3: After carefully reading the remaining 63 articles, a total of 27 articles remained as relevant for the present study.

Analysis and Synthesis of the Results

Each of the 27 articles was coded by two researchers [CS and SH], with the following elements annotated: study questions/objectives, description of the teaching characteristics (definitions used, explanations), operationalisation of the characteristics, sample, units of analysis (students, classes, schools, countries), research design, results, and effects described in the discussion section. The annotation process was conducted individually and independently. The results of the coding procedure were summarised by consensus. Then, the instructional characteristics described in the articles were categorized based on deductive-inductive category formation following Kuckartz (2014).

Following Kuckartz's (2014) recommendation for deductive categorization, characteristics of instructional quality were clustered into an existing category system developed by Schlesinger et al. (2018), who proposed to include both generic and subject-didactic categories: classroom management, student support, cognitive activation, subject-related quality, and teaching-related quality. These categories served as a guide for categorizing the characteristics described in the analysed articles. Some characteristics, however, did not fit into this scheme, necessitating some revisions. Following Kuckartz (2014) guidelines, a new clustering was performed based on the theoretical descriptions of the characteristics. As the descriptions comprised mostly literature reviews, their references were also considered. Clustering was again performed independently by two researchers [CS and SH], and disagreement in the allocation of individual characteristics was solved by consensus.

Additionally, the effects of the characteristics on student performance in mathematics were assessed and described. For that purpose, we created a rating scheme to assess the reported effects of a teaching characteristic on student performance. The rating scheme comprises three categories:

1. + (plus) denotes that the described trait has been shown to improve performance,
2. - (minus) denotes that the described trait has been shown to have an adverse effect, and
3. ~ (tilde) denotes that no effect was found, or the effect was negligible.

For standardized values of the beta coefficients in the reported models, ±0.05 is usually considered a reasonable threshold for distinguishing positive, negative or negligible effects (Pigott & Polanin, 2020; Toste et al., 2020). A threshold value of ±0.05 was applied to studies reporting standardized values. For non-standardized values, we performed a subjective assessment, considering the authors' perspective, the context, the size of the values, and the relationship to the other values of the study. To counteract the effects of personal idiosyncrasies, the effects were first described independently by two researchers [CS and SH] and differences were again solved by consensus.

Based on the recommendations of Alexander (2020), the results of the reviews are presented as textual descriptions, highlighting the special aspects of the studies and describing the effects of the characteristics.

Results

Overview of the Selected Articles

Table 1 provides an overview of the articles analysed for the review. Each article is assigned a numerical study ID (in brackets), which is used for reference in the following tables. In addition to frequent use of international data sets, such as Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS), 4 of the 25 articles deal with data from the "Pythagoras-Study" (Leibniz Institute for Research and Information in Education, n.d.).

Table 1. Selected Studies

Study ID	Source	Country	Data set (study)	N	Category	Perspective		
						Teacher	Student	External
1	Baumert et al. (2010)	DEU	COACTIV – PISA 2003	4353 St, 181 T, 194 C	CM, SS, CA	x	x	x
2	Bardach et al. (2021)	AUT	Own survey	1743 St, 89 C	SS		x	
3	Bellens et al. (2019)	BEL, DEU, NOR	TIMSS 2015	5404 St, 295 C, 153 S	CM, SS, CA		x	
4	Cheema and Kitsantas (2014)	USA	PISA 2003	5456 St, 274 S	CM		x	
5	Creemers and Kyriakides (2008)	CYP	Own survey	2503 St, 108 C, 52 S	CM, SS, CA, misc.		x	x
6	Demir (2018)	TUR	PISA 2012	4848 St, 170 S	CM, SS, CA		x	
7	Dubberke et al. (2008)	DEU	COACTIV – PISA 2003	2483 St, 155 C	SS, CA	x	x	
8	Eriksson et al. (2019)	TIMSS countries, especially SWE	TIMSS 2003, 2007 (SWE also 2011, 2015)	SWE: 2003: 4255 St, 271 C; 2007: 5215 St, 307 C; 2011: 5816 St, 266 C; 2015: 4090 St, 206 C TIMSS countries: no information	SD, misc.	x	x	
9	Faber et al. (2018)	NLD	Own survey	953 St, 26 S, 51 T	SS			x
10	Gamlem (2019)	NOR	Own survey	234 St, 10 C	CM, SS, CA, SD, misc.			x
11	Kane et al. (2011)	USA	Own survey (Cincinnati Public Schools' Teacher Evaluation System (TES))	no information	CM, CA			x
12	Kelcey et al. (2019)	USA	Own survey	10987 St, 302 T, 52 S	CM, SD			x
13	Kistner et al. (2010)	DEU	"Pythagoras-Study" (Leibniz Institute for Research and Information in Education, n.d.)	538 St, 20 T, 40 S	SS			x
14	Kuger et al. (2017)	DEU	PISA 2012 (complementary survey)	2556 St	CM, SS, CA, SD, misc.		x	
15	Kunter et al. (2013)	DEU	COACTIV – PISA 2003	4353 St, 181 T, 194 C	CM, SS, CA	x	x	x
16	Lipowsky et al. (2009)	DEU, CHE	"Pythagoras-Study" (Leibniz Institute for Research and Information in Education, n.d.)	913 St, 38 C	CM, SS, CA			x
17	Ni et al. (2018)	CHN	Existing data set (Ni et al., 2011)	1779 St, 30 T, 10 S	CA			x

Table 1. Continued

Study ID	Source	Country	Data set (study)	N	Category	Perspective		
						Teacher	Student	External
18	Ozberk et al. (2018)	TUR	PISA 2012	317 St	CM, SS, CA, SD, misc.		x	
19	Ozdemir and Sahal (2018)	TUR	Own survey	69 St	CA		N/A [a]	
20	Pauli et al. (2008)	DEU, CHE	"Pythagoras-Study" (Leibniz Institute for Research and Information in Education, n.d.)	773 St, 38 C	CA, SD			x
21	Pinger et al. (2018)	DEU	Own survey (Conditions and Consequences of Classroom Assessment Co^{2}CA)	426 St, 17 T	SS		x	x
22	Reusser and Pauli (2013)	DEU, CHE	"Pythagoras-Study" (Leibniz Institute for Research and Information in Education, n.d.)	39 C	SD, misc.		x	x
23	Seidel and Shavelson (2007)	N/A [b]	N/A [b]	N/A [b]	CM, SS, CA, SD, misc.		N/A [b]	
24	Stronge et al. (2011)	USA	Own survey	Phase 1: 307 T, 110 S, ca. 4600 St Phase 2: 32 T	CM, SS, SD, misc.	x		x
25	Vanlaar et al. (2015)	BEL, CYP, DEU, GRC, IRL, SVN	Project "Establishing a knowledge base for quality in education: Testing a dynamic theory of educational effectiveness" (Creemers et al., 2013)	10742 St, 571 C, 334 S	CM, SS, CA, misc.	x	x	
26	Yi and Lee (2017)	KOR, SGP	PISA 2012	KOR: 5033 St, 156 S SGP: 5546 St, 172 S	CM, SS, CA, misc.		x	
27	Zepeda et al. (2019)	USA	Measures of Effective Teaching Longitudinal Database (METLDB)	no information	CA			x

Note. Country codes refer to the ISO 3166 (alpha-3) standard.
N = sample size; St = students; T = teachers; S = schools; C = classes; CM = classroom management; SS = student support; CA = cognitive activation; SD = subject-didactic specifics; misc. = miscellany; N/A = not applicable.
[a] In this study, the effect of a teaching method was measured with a pre- and post-test.
[b] Meta-analysis.

Description of the Clustered Categories

The clustering resulted in five categories: classroom management, student support, cognitive activation, subject-didactic specifics, and miscellaneous (Table 2; see also Table A1 in the Appendix for a more extensive version).

The first three categories represent the three basic dimensions of quality teaching of Klieme et al. (2006), which are frequently used both in German-speaking countries and internationally (Praetorius et al., 2018). Almost all articles (except for [2], [8], [11], [22]) link the basic dimensions to student performance. However, the interpretations of the dimensions vary considerably, leaving a heterogeneous picture of quality teaching. The numerous instruments for determining instructional quality through the three basic dimensions also lead to different operationalisations with varying focus. Owing to this situation, we formed various subcategories to give a more detailed and meaningful overview. However, not all aspects of the basic dimensions found in the articles could be assigned to a subcategory.

The three basic dimensions described primarily generic components of quality teaching. Schlesinger et al. (2018) added the categories "subject-related quality" and "teaching-related quality" to the three basic dimensions to include subject-related components. Since we could only identify a few characteristics in the selected articles, we created a single new category 'subject-didactic specifics' (with subcategories) to encompass Schlesinger's concepts.

The 'miscellaneous' category collected characteristics affecting student performance in mathematics fitting neither into an existing category nor allowing for the creation of a new category due to their heterogeneity. This category will not appear in Table 2.

Table 2. Categories and Subcategories of the Studies Surveyed

Categories and subcategories	Study ID (see Table 1)	Number of articles in respective (sub)category
Classroom management		16
Dealing with disruptions	1, 3, 4, 6, 10, 12, 14, 15, 16, 18, 23, 25, 26	13
Supporting learning processes	5, 10, 11, 12, 18, 25, 26	7
Managing learning time	1, 15, 16, 18, 25, 26	6
Student support		18
Feedback (teacher's feedback to students)	3, 10, 14, 16, 18, 21, 23, 24	8
Constructive approach to errors	1, 3, 7, 15	4
Appreciation of students	1, 3, 5, 7, 10, 15, 16, 23	8
Teacher support	1, 2, 3, 7, 9, 11, 15, 18, 24, 26	10
Student orientation	1, 6, 7, 14, 15, 18, 24, 26	8
Cognitive activation		18
Challenging tasks	1, 3, 5, 6, 7, 14, 15, 16, 17, 18, 20, 26, 27	13
Challenging questions/class discussions	3, 5, 6, 7, 10, 11, 14, 16, 18, 20, 25, 26	12
Activation of prior knowledge	16	1
Support through metacognition	7, 10, 14, 16, 23, 25, 27	7
Subject-didactic specifics		9
Relation to students' daily lives	8, 14, 18	3
Mathematical connections/mathematical depth	20, 22, 24	3
Teacher's mathematical accuracy	12	1

Classroom management

Classroom management and its aspects have been proposed for many years as the basis for successful teaching (e.g., Klieme et al., 2006; [10], [23]). Klieme et al. (2006) described classroom management as a teacher's task, who, as the organizer and facilitator of institutionalized learning, should ensure that the lesson takes place largely without disruption and in a structured form to gain the students' attention. Therefore, classroom management is characterized by disciplinary and organizing elements. We split classroom management into the three components 'dealing with disruptions', 'managing learning time', and 'supporting learning processes'.

'Dealing with disruptions' represents the disciplinary component that entails the teacher to react to conflicts or interruptions that occur in the classroom and, in the best case, to use preventive disruption management to avoid such disruptions in the future [1, 4, 15, 23].

The organizing component includes 'management of learning time' and 'support of learning processes'. 'Supporting learning processes' aims at helping students to understand the matter, e.g., by explaining the significance of a task at a specific phase of the lesson or providing clear lesson expectations [5, 16]. This should ensure that the teaching-learning processes run as smoothly as possible, which, in turn, fosters a deeper engagement with the material [11, 14]. Both the disciplinary and organizing components facilitate efficient time management, which, in turn, also translates to more effective learning time for students [26].

Student support

Klieme et al. (2006) stated that the teacher, who as an educator is also responsible for social leadership in the classroom, should ensure that students integrate socially, learn independently, and receive positive feedback on the skills achieved. This category focuses on the quality of social interactions and classroom relationships to support students in their personal learning processes, (e.g., Schlesinger et al. (2018); Praetorius et al. (2018)). Klieme's et al. (2001) first description sparked interest in the research community, but the interpretations of this category are very heterogeneous. Based on the multitude of extensive descriptions and operationalisations given of previous literature, the following subcategories containing the supporting and relationship components were formed: feedback (teacher's feedback to students), constructive approach to errors, appreciation of students, teacher support, and student orientation.

'Support processes' in students' learning include adaptive explanations and tasks at different stages of the learning process, respectful and patient handling of students' mistakes, and constructive feedback or support for students with difficulties in the learning process [1, 3, 7, 14, 15, 16, 26].

The 'relationship level' includes components such as respectful and appreciative teacher-student interaction, caring teacher behaviour, social orientation of teaching, and support of the students' desire for autonomy [5, 7, 10]. These aspects promote active engagement of learners and increase their motivation. According to motivational psychological theories (e.g., Ryan & Deci, 2000), autonomy, competence, and social relatedness are necessary to maintain motivation for learning and increase student interest in the subject.

Cognitive activation

In the process of knowledge acquisition, context and situational factors are important for the design of learning environments that focus on the construction of knowledge by the learners themselves (De Corte, 2004). Learners select information, arrange it, and integrate it with their previous knowledge. Thus, knowledge schemata that enable learners to act detached from the original learning situation and gain independence emerge (Nückles & Wittwer, 2014). This dimension is based on the cognitive-constructivist learning theory. Lipowsky et al. (2009) [16] stated that "Cognitive activation is an instructional practice that encourages students to engage in higher-level thinking and thus to develop an elaborated knowledge base" (p. 529). In other words, cognitive activation is teaching that promotes active intellectual engagement with the learning material and networking of old and new knowledge.

From the 18 articles dealing with cognitive activation, we formed four subcategories: challenging tasks, challenging questions/class discussions, activation of prior knowledge, and support through metacognition.

The teacher has an essential role, arranging the learning environment fostering insightful learning and cognitive activation of students in their respective subject (Klieme et al., 2006). Teachers must explore students' prior knowledge and subsequently build on it (Nückles & Wittwer, 2014). 'Challenging tasks' that are classified as cognitively activating learning opportunities are one way to help students building on previous knowledge (Klieme et al., 2001). The potential for cognitive activation can only be exploited if students are confronted with "challenging exercises" (Kuger et al., 2017, p. 66) [14]. To achieve such potential, 'challenging questions' and class discussions are required, meaning the focus is on an interaction between the social context and cognition. "Situative metacognition involves the interaction of multiple entities (e.g., students, teachers, computer tutors) engaging in the processes of planning, monitoring, and/or evaluating performance on a task" (Zepeda et al., 2019, p. 524) [27]. Praetorius et al. (2018) argued that participation in content discourses in class discussions supports the development of conceptual understanding. 'Metacognition' is used to describe a wide range of concepts (Akturk & Sahin, 2011). Essentially, metacognition describes an individual's awareness of their cognitive abilities and the ability to organize them (Flavell, 1979). Supporting students' metacognition is a task of the teacher, which Zepeda et al. (2019) [27] referred to as "metacognitive delivery."

Subject-didactic specifics

Hiebert and Grouws (2007) emphasized that considering subject-didactic characteristics is crucial for good mathematics teaching, stating that "the nature of classroom mathematics teaching significantly affects the nature and level of students' learning" (p. 371). Numerous other authors shared this opinion (e.g., Bruder, 2018; Drollinger-Vetter, 2011; Klieme & Rakoczy, 2008; Lindmeier & Heinze, 2020; Praetorius, Rogh, & Kleickmann, 2020; Seidel & Shavelson, 2007).

The review process showed that the categories proposed by Schlesinger et al. (2018) (subject-related quality and teaching-related quality) could not be applied to the present research, as subject-didactic aspects are so far not sufficiently established as an independent characteristic. Some studies (e.g., [26]) included subject-didactic aspects in the operationalisation of the three basic dimensions and did not define a separate category. Nine articles [8, 10, 12, 14, 18, 20, 22, 23, 24] dealing with subject-specific elements addressed specific didactic issues that could not be assigned to the three basic dimensions. Based on the data, the category of subject-didactic specifics was split into three subcategories: connection to students' daily lives, mathematical connections/mathematical depth, and teacher's mathematical accuracy.

Eriksson et al. (2019) [8] discussed the relevance of mathematics material for students' daily lives, which van den Heuvel-Panhuizen and Drijvers (2014) considered one of the most significant ideas in mathematics education in recent years. Similarly, Kuger et al. (2017) [14] and Ozberk et al. (2018) [18] addressed tasks with strong practical orientation. Mathematical connections require special consideration by the teacher through "meaningful conceptualization of knowledge rather than on isolated facts" (Stronge et al., 2011, p. 341) [24]. However, all subject-didactic specifics are irrelevant if the teacher is imprecise, unclear, or incorrect. Kelcey et al. (2019) [12] related the teacher's mathematical accuracy and their mistakes to student performance.

Effects and their Relevance for the Categories Formed

Table 3, 4, 5, 6 and 7 depict the findings on the effect of teaching characteristics on student performance in mathematics for each category. To avoid publication bias, the tables provide an overview of both significant and non-significant results, including calculations at individual and class levels. Because most articles report aggregated variables, a closer assessment of our subcategories was not possible. Table A2 [Appendix] provides a comprehensive overview of all variables and their effects.

The narrative description addressed conspicuous features for each category and discusses them in more detail. For this purpose, additional literature was used in the descriptions, including articles that did not appear in the primary search but were found in additional literature search. These articles were on par with the reviewed articles and were intended to provide further descriptions and explanations, including the (reported) expected effects on student performance. For the three basic dimensions, the expectations of the authors first describing them were reported. There were no concrete expectations for subject-didactic specifics, as this is a newly formed.

Classroom management

For the category classroom management, clear, positive effects on student performance were expected (Lipowsky et al., 2009). The results were consistent with this expectation. Eleven out of 16 articles [1, 3, 4, 6, 10, 11, 14, 15, 16, 24, 25] examining classroom management reported a positive effect on student performance in mathematics (Table 3).

There was a difference between class and individual level. While results at the class level were mostly positive, two studies [14, 18] showed no effect at the individual level. This difference was particularly evident in Kuger et al. (2017) [14]. At the individual level, no effect was established for the variable "clarity of rules and structure," whereas a positive effect was found at the class level. Ozberk et al. (2018) [18] only performed calculations on an individual level, and no effects were found for "classroom management" and "discipline." These findings suggest that the class as a whole benefited more from good leadership and management than from individual students' participation.

Another interesting finding was found in Yi and Lee (2017) [26], who reported a negative effect of classroom management for Singapore. The authors suspected that weaker students required more individual support, and therefore, structuring at the class level did not have a positive effect on learning success. Moreover, they found indications of a non-linear relationship between classroom management and student performance. Helmke (2017) assumed an inverse U-shaped relationship for classroom management. He stated that in educational research, it is often not a question of a maximum but rather of an optimum, and both a complete lack and excessive, exaggerated classroom management are equally unfavourable for learning success.

Table 3. The Effects of Individual Characteristics on Student Performance in Mathematics: Classroom Management

Variable names	**Study ID (see Table 1)**		
	[+]	**[~]**	**[-]**
Classroom management	[1]; [3 - DEU]; [3 - NOR]; [15]; [16]; [24]	[3 - BEL (Flanders)]; [18*]; [26 - KOR]	[26 - SGP]
Disciplinary climate	[4]; [6]	[18*]	
Behavior management	[10]		
Classroom environment relative to instructional practices	[11]		
Dealing with misbehavior	[25]		
Regelklarheit und Struktur [carity of rules and structure]	[14]	[14*]	
Orientation		[5]	
Structuring		[5]	
Productivity		[10]	
Classroom organization		[12]	
Organization of learning		[23]	
Instructional focus on learning		[24]	
Quantitative structuring	[25]		
Qualitative structuring		[25]	
Time management	[25]	[5]	

Note. (*) Calculated at the individual level; Variable names were taken directly from the articles and translated [in brackets]. The country codes refer to the ISO 3166 (alpha-3) standard.

Student support

It was assumed that student support had a direct, positive effect on student motivation and an indirect effect on student performance (Klieme et al., 2006; Lipowsky et al., 2009). Lipowsky et al. (2009) stated that there were no clear results for effects on student performance. The review is in line with previous literature as only eight out of 18 articles [2, 3, 10, 13, 18, 21, 25, 26] reported a positive effect of certain aspects of student support. Ozberk et al. (2018) [18] and Yi and Lee (2017) [26] examined the effect of the 'teacher support' subcategory on student performance. Ozberk et al. (2018) [18] investigated resilient students and stated "Findings of the study reveal that teacher support has an important role in the success of resilient students. .... It is also based on the fact that the teacher provides students with extra help when needed" (p. 123). Yi and Lee (2017) [26] found a positive effect on student performance and interest for the "learning support" variable. They referred to Anderman et al. (2001), who suggested that for learning to be effective, the teacher must succeed in creating a learning environment for students in which they feel or have the intention to feel successful.

Bellens et al. (2019) [3] used the aggregated variable "supportive climate" to address not only teacher support but also the aspects of 'feedback' (teacher's feedback to students), constructive approach to errors, and appreciation of students. This variable showed a positive effect for Flanders (Belgium) but no effect for Germany and Norway [3]. Similar to Bellens et al. (2019) [3], Gamlem (2019) [10] provided evidence of a positive effect on student performance for "positive climate." This variable also encompassed a relationship component; however, no positive effect was identified for the whole category or for individual aspects. This characteristic affects socio-emotional output in addition to cognitive output. The lack of evidence of an effect is often justified by the impact of student support on student performance via socio-emotional output [14, 16, 23, 26].

Two studies [14, 21] reported a negative effect of the 'feedback' subcategory (teacher's feedback to students) on student performance. Kuger et al. (2017) [14] assumed that feedback is provided to a greater extent to classes with lower-achieving students. Pinger et al. (2018) [21] suspected that the reason for the negative effect of feedback on post-test performance was the negative correlation between student pre-test performance and feedback. The authors assumed a negative correlation between student pre-test performance and feedback. Their multi-level analysis revealed a negative effect of feedback on the post-test performance show that weaker students receive more feedback. However, this does not affect their performance positively.

Another noteworthy observation is that 'student orientation' had a negative impact on student performance (Table 4). Demir (2018) [6] believed that one of the main reasons for this may be inexperience in implementing student orientation strategies, such as teaching students to work in small groups. This inexperience could cause the teacher to feel insecure and create an adverse effect on the classroom atmosphere, inhibiting cognitive performance gain (Randler & Bogner, 2002).

Dubberke et al. (2008) [7] reported a negative effect of "student support" on student performance. They saw a negative suppression effect as a reason for this. Student support seems to suppress variance components of cognitive activation that are irrelevant to mathematics performance.

Table 4. The Effects of Individual Characteristics on Student Performance in Mathematics: Student Support

Variable names	**Study ID (see Table 1)**		
	[+]	**[~]**	**[-]**
(Individual) learning support	[26 - KOR]; [26 - SGP]	[1]; [15]	
Teacher support	[18*]		
Konstruktive Unterstützung [student support]		[14]; [14*]	[7]
Supportive climate	[3 - BEL (Flanders)]	[3 - DEU]; [3 - NOR]; [16]	
Positive climate	[10]		
Negative climate		[10]	
Classroom-level environment (teacher-student relationships and relationships among students)		[5]	
Teacher-student interaction	[25]		
Teachers' student orientation			[6]
Differentiated instruction		[9]	
Teacher sensitivity		[10]	
Teacher behavior – student orientation		[18*]	
Teacher behavior – formative assessment		[18*]	
Regard for adolescent perspectives		[10]	
Total number of observed strategies - instructions		[13]	
Explicit		[13]	
Implicit		[13]	
Indirect (constructivism, transfer)	[13]		
Indirect (cooperation, self-direction)		[13]	
Steuerung durch die Lehrkraft [management by the teacher]		[14]; [14*]	
Regulation/monitoring		[23]	
Social/direct experiences		[23]	
Time for learning		[23]	
Instructional clarity	[2]	[24]	
Assessment for understanding		[24]	
Quality of feedback		[10]	
Direct (organization; monitoring and evaluation; resource management)		[13]	
Direct (elaboration; planning and systematic activity; feedback)		[13]	
Lehrerfeedback [teacher feedback]		[14*]	[14]
Number of comments			[21*]
Specificity			[21*]
Feedback utilization I	[21]		
Feedback utilization II	[21]		
Feedback		[24]	

Note. (*) Calculated at the individual level; Variable names were taken directly from the articles and translated [in brackets]. The country codes refer to the ISO 3166 (alpha-3) standard.

Cognitive activation

"Empirical evidence for the benefits of cognitive activation from classroom studies is still weak" (Lipowsky et al., 2009, p. 529). Nevertheless, a positive effect of cognitive activation on student performance was expected (Klieme et al., 2001; Klieme et al., 2006). Eleven of the 18 articles addressing cognitive activation [1, 6, 7, 15, 16, 17, 19, 20, 25, 26, 27] showed a positive effect on student performance (Table 5). These effects related to the whole construct of cognitive activation as well as to individual aspects. A comprehensive operationalisation covering all aspects of cognitive activation was used in the study by Lipowsky et al. (2009) [16], who found positive evidence for the entire construct. Two aspects that were often used to operationalise cognitive activation and for which clear positive results were found were 'challenging tasks' and 'challenging questions and classroom discussions'. Ten of the 12 showed positive evidence on student performance, with both aspects reflected in the operationalisation.

However, despite the fairly clear tendency for a positive influence on student performance, there were noticeable deviations in the results. First, Bellens et al. (2019) [3] found a negative effect of one aspect of cognitive activation, which was operationalised by the item, "Our mathematics teacher asks me what I have understood and what I haven't" (p. 25). The authors stated, "As we did not take into account prior achievement, reversed causality might explain this negative relation. Considering this assumption, our results might point to the fact that high achieving students in Flanders do not receive sufficient support for learning, resulting in reduced chances for the highest possible achievement" (p. 20).

Second, ten of 18 studies [4, 10, 11, 14, 17, 18, 20, 23, 25] found no effect of cognitive activation or its individual aspects. Kuger et al. (2017) [14] reported that in questionnaires, students tended to express their individual preferences to the question of whether certain aspects (e.g., mathematics task types) are used "frequently" or "rarely" rather than reflecting an objective frequency. Ni et al. (2018) [17] explained the lack of effectiveness of various problem-solving strategies for building cognitive performance by the incorrect implementation of such strategies by the teacher. They stated as follows: "Teachers may pursue a task with multiple solution methods primarily to encourage students to talk more, rather than making strategic uses of students' generated methods to advance a greater understanding of mathematics" (p. 715).

Table 5. The Effects of Individual Characteristics on Student Performance in Mathematics: Cognitive Activation

Variable names	**Study ID (see Table 1)**		
	[+]	**[~]**	**[-]**
Cognitive activation	[7]; [15]; [16]; [26 - KOR]; [26 - SGP]	[18*]	[3 - BEL (Flanders) Component 2]
Potential zur kognitiven Aktivierung [potential for cognitive activation]		[14]; [14*]	
Cognitive activation strategies	[6]		
Questioning	[25]	[5]	
Teaching-modeling		[5]; [25]	
Analysis and inquiry		[10]	
Instructional dialogue		[10]	
Problem-posing approach	[19]		
Länge der Schülerbeiträge im Klassengespräch [length of student contribution in class discussions]	[20]		
Questions & discussion approach relative to standards & content focus		[11]	
Anteil anspruchvolles Üben [proportion of challenging practice]		[20]	
Basic processing		[23]	
Cognitive level of tasks	[1]		
Curricular level of tasks	[1]		
Effect of "Cognitive demands of mathematics instructional tasks" to calculation, routine problem solving and complex problem-solving		[17]	
Effect of "External (multiple) representations of mathematics instructional tasks" to solving complex problems	[17]		
Effect of "External (multiple) representations of mathematics instructional tasks" to calculation and routine problem-solving		[17]	
Multiple solution methods of calculation, routine problem-solving, and complex problem-solving		[17]	
Metacognitive support	[27]		

Note. (*) Calculated at the individual level; Variable names were taken directly from the articles and translated [in brackets]. The country codes refer to the ISO 3166 (alpha-3) standard.

Subject-didactic specifics

The findings regarding the effectiveness of subject-specific didactics on student performance varied (Table 6). Some aspects – particularly "structural clarity" – clearly showed positive relationships [20, 22]. Other aspects, however, had mixed results, such as the characteristic "experience with school algebra". While Kuger et al. (2017) [14] observed a positive effect for this characteristic, no effect was found by Ozberk et al. (2018) [18].

Studies also found mixed results for the characteristic 'connection to students' daily lives'. While Eriksson et al. (2019) [8] and Kuger et al. (2017) [14] reported a negative effect at the class level, Ozberk et al. (2018) [18] reported a positive relationship at the individual level.

The divergent effects of this characteristic regarding the allocation levels (individual or class level) may result from differences in its operationalisation resulting from different class compositions [14]. In contrast, Eriksson et al. (2019) [8] theorized that student performance and 'connection to students' daily lives' are driven in opposite directions by a third variable (e.g., less time on everyday aspects). Additionally, the authors hypothesized that such an effect could be explained by a causal relationship between performance and instruction.

Table 6. The Effects of Individual Characteristics on Student Performance in Mathematics: Subject-Didactic Specifics

Variable names	**Study ID (see Table 1)**		
	[+]	**[~]**	**[-]**
Strukturelle Klarheit [structural clarity]	[20]; [22]		
We relate what we are learning in mathematics to our daily lives (Daily)		[8 - 45 countries in TIMSS, 2003]; [8 - 50 countries in TIMSS, 2007]	[8 - SWE TIMSS, 2003-2015]
Experience with applied mathematical tasks	[18*]	[14*]	[14]
We memorize formulas and procedures (Memo)	[8 - SWE TIMSS 2003-2015]	[8 - 50 countries in TIMSS, 2007]	
Experience with pure mathematical tasks	[14*]; [14]	[18]	
Familiarity with mathematical concepts	[18]		
Familiarity with mathematics concepts (corrected for overclaiming)		[18]	
Content understanding		[10]	
Domain-specific processing	[23]		
Instructional complexity		[24]	
Errors		[12]	

Note. (*) Calculated at the individual level; Variable names were taken directly from the articles and translated [in brackets]. The country codes refer to the ISO 3166 (alpha-3) standard.

Miscellany

Here, we observed heterogeneous results in terms of content (Table 7). However, we could combine characteristics addressing forms of teaching, such as the variable “lecture-style presentation” in Eriksson et al. (2019) [8]. No study investigating such aspects [8, 10, 18, 22, 24] found a significant effect on student performance.

Noteworthy is the negative effect of the “student-oriented instruction” variable on student performance for Singapore and Korea (Yi & Lee, 2017) [26]. The authors stated that “student-oriented instruction may not be effective if it fails to engage students in learning but only makes them work on math-related activities in a superficial way” (pp. 36-37).

Table 7. The Effects of Individual Characteristics on Student Performance in Mathematics: Miscellaneous

Variable names	Study ID (see Table 1)		
	[+]	[~]	[-]
We listen to the teacher give a lecture-style presentation (Listen)	[8 - SWE TIMSS, 2003-2015]	[8 - 45 countries in TIMSS, 2003]; [8 - 50 countries in TIMSS, 2007]	
Instructional learning formats		[10]	
Schülerorientiertes Unterrichten [student-oriented teaching]		[14*]; [14]	
Teacher behavior, teacher-directed instruction		[18*]	
Unterrichtsinszenierung [lesson design]		[22]	
Instructional differentiation		[24]	
Teacher-student orientation (student-oriented instruction)			[26 - KOR]; [26 - SGP]
Klassengespräche (ko-konstruktives Gespräch) [class discussions (co-constructive)]	[22]		
Social context		[23]	
Evaluation of learning		[23]	
Goal setting and orientation		[23]	
Expectations for student learning		[24]	
Application	[5]; [25]		
Assessment	[25]	[5]	

Note. (*) Calculated at the individual level; Variable names were taken directly from the articles and translated [in brackets]. The country codes refer to the ISO 3166 (alpha-3) standard.

Discussion

To achieve the three objectives formulated in the introduction, 27 articles were selected, analysed, and clustered in a scoping review process. This clustering helped to identify both generic and subject-didactic characteristics and assess their effects. The results yielded some additional insights into the characteristics of quality in mathematics instruction over the last two decades. The results of this review can be seen as a supplement to a very comprehensive and well-researched field.

Separability of Generic and Subject-Didactic Characteristics

Charalambous and Praetorius (2018) made an important contribution to the field of instructional quality through "setting the ground for understanding instructional quality more comprehensively" as they stated in ~~the title of~~ their article. For obtaining a comprehensive picture of teaching they gave an overview of existing frameworks by "organizing them along a continuum form more generic to more content-specific" (p. 355). This contribution has initiated a discussion for developing a theory of teaching quality (e.g., Praetorius, Klieme, et al., 2020). So far, the idea of a continuum refers to frameworks of teaching quality (Charalambous & Praetorius, 2018) and teaching characteristics seem more dichotomous, i.e., they are either generic or subject-didactic. However, the analysis of the 27 articles demonstrated that the teaching characteristics should be placed on a continuum, depending on the operationalisation. The first three categories (classroom management, student support, and cognitive activation) reflect the three basic dimensions and are generally considered generic, although cognitive activation is regarded subject-dependent in its design (Klieme et al., 2006; Lipowsky et al., 2009; Praetorius et al., 2014). Schlesinger and Jentsch (2016) pointed out that characteristics of instructional quality related to mathematics instruction have, to date, not been considered in subject-didactic research. Therefore, the three basic dimensions of quality teaching were often supplemented by subject-didactic characteristics. Jentsch et al. (2021) proposed adding a fourth dimension, including aspects of subject-related, diagnostic, evaluative, or structuring measures in mathematics lessons. Drollinger-Vetter (2011) also investigated whether each characteristic of instructional quality has both a generic and a subject-didactic component. They provided an example showing that classroom management has a subject component, where under- and overcharged students show increased disquietude in otherwise disciplined classes. However, Bruder (2018) argued for explicitly considering technical and didactical aspects as basic dimensions relevant to the quality of teaching in mathematics. This call for adding a subject-didactic category is in line with the demands of other authors (Brunner, 2018; Drollinger-Vetter, 2011; Lindmeier & Heinze, 2020; Praetorius, Klieme, et al., 2020; Praetorius, Rogh, & Kleickmann, 2020; Schlesinger et al., 2018). Also, Praetorius, Herrmann, et al. (2020) synthesized the teaching quality in different subject matters. This discussion, combined with the findings of the present scoping review, suggests the following two requirements. First, a uniform, content-related, and theoretically well-founded definition of subject-didactic characteristics along with uniform operationalisations is required to establish these characteristics in addition to the three basic dimensions. This would also help better study and understand the effects on student performance. Second, as the review shows, there is a large body of literature that has contributed to the development of teaching quality. Due to the fact that there are many individual pieces of the puzzle, it is now necessary to create an overall picture. For this Praetorius, Rogh, and Kleickmann (2020) already published a first synthesis framework to organize the heterogeneous field and show limits of the three basic dimensions model. This is in line with the scoping review. In summary, both generic and subject-didactic characteristics turned out to influence the performance of mathematics students. Hence, the entire range of characteristics is necessary to assess the quality of mathematics teaching.

Discussion of Descriptions and Effects

Although cognitive activation is a "generic" dimension (like classroom management and student support), some articles had subject-didactic elements in their operationalisation [1, 14, 16]. Through the idea with the continuum from generic to subject-didactic characteristics, it is possible to position the category cognitive activation and their subcategories in between the two ends of the continuum. The continuum makes it possible to arrange the characteristics in relation to other characteristics between generic and subject-didactic. The description of the category student support shows that this category focuses on social and emotional components, which, in addition to students' subject-related performance, also affect their social-emotional performance (Lipowsky et al., 2009). This result is in line with Kunter (2005).

The effects on student performance and the associated features in Tables 3, 4, 5, 6 and 7 still provide an incomplete picture. Depending on the study, several characteristics were reported to have positive, negative, or negligible effects. Helmke (2017) provided a possible explanation by arguing that "good teaching" is not identical with optimal and certainly not with maximum expression of all characteristics. Considerably different patterns of successful teaching and deficits in one area can be compensated to a certain extent by strengths in other areas. The different effects could be attributed to differences at the national level (e.g., school system, educational policy), which influences variables at lower levels (e.g., teachers and teaching), as Nilsen et al. (2016) suggested. For example, Bellens et al. (2019) [3] found that the variable "classroom management" had no effect on performance for Flanders. This might be offset by the positive effect of the variable "supportive climate" (which is part of student support) on student performance, as suggested by the significant correlation between classroom management and supportive climate ($r = .36$, $p < .001$). Another example of differences at the national level especially for educational policy is given by Balagtas (2021). She underlines the need for continuous updating of the teacher education program in the Philippines to meet the international standards for quality teaching. Without a curriculum producing the desired quality of mathematics teachers, it is hardly possible to guarantee instructional quality. For comparability, more emphasis was placed on international large-scale assessments (ILSAs), as Klieme (2013) described: "ILSAs establish a monitoring structure that provides reliable comparative information on education systems, describing system structures as well as the functioning and the productivity (i.e., the gross outcome or 'yield') of education systems" (p. 115).

Conclusion

Working within the field of instructional quality means focusing on mechanisms and teaching characteristics leading to change in teacher practice and student performance. This research interest has been the focus of many researchers for several decades, leading to an immense and diverse body of literature. However, due to different definitions and operationalisation, the picture of which characteristics are important for instructional quality is not entirely clear. To provide an overview of the current state of research, this review aimed to (a) identify both generic and subject-didactic characteristics affecting student performance in mathematics in secondary school, (b) cluster these characteristics into categories or areas to show key areas for quality teaching, and (c) analyse and assess the effects of these characteristics on student performance to rate the scientific evidences in the context of the analysed articles. The analysis showed that, on the one hand, there is a consensus within the scientific community on the three general basic dimensions. On the other hand, the definition of subject-didactic characteristics is inconsistent. Our conclusion is to place teaching characteristics on a continuum, ranging from generic to subject-didactic. Moreover, our study revealed the need for further empirical investigation of the category we termed 'subject-didactic specifics' which should be further developed to establish it as a separate category in empirical research.

Recommendations

To achieve the goal of identifying both generic and subject-didactic characteristics affecting student performance in mathematics in secondary school, it was necessary to deal with a broad range of characteristics from different categories. This fact, in combination with the complex and untidy nature of the body of literature, impeded an in-depth analysis of each category. However, this review provides a starting point for further in-depth analyses of teaching quality. A promising task could be analysing classroom management and its effects on student performance. Another interesting direction for research is to examine the effects of teaching characteristics not on student performance but on student motivation.

Limitations

The present review focused solely on the subject of mathematics. In recent years, there have been efforts to investigate the stability of certain characteristics of instructional quality on student performance across subjects (Praetorius et al., 2016) and whether the instructional quality of, for instance, tutoring lessons can be described by characteristics (Guill et al., 2020).

This review did not consider teacher personality and characteristics, which are also related to the quality of teaching and partly considered as indicators for quality teaching (Nilsen et al., 2016; Praetorius et al., 2016). Future research could investigate the effect of teacher personality and characteristics on student performance.

Limitations in the analyses resulted from different operationalisation in the included studies. To standardize the understanding of teaching quality in the future, there are two essential requirements. First, uniform definitions and operationalisation of teaching characteristics are needed. Second, methodological approaches should be unified and certain guidelines should be established to better compare results, such as reporting both standardized and non-standardized values.

Acknowledgements

The authors thank Rainer W. Alexandrowicz for his valuable contribution to this article. The authors acknowledge the financial support from the University of Klagenfurt.

Declaration of interest

We have no known conflict of interest to disclose.

Authorship Contribution Statement

Spreitzer: Conceptualization, design, data analysis/ interpretation, writing. Hafner: data analysis/ interpretation, critical revision of manuscript. Krainer: Edition/reviewing, supervision, final approval. Vohns: Editing/reviewing, supervision.

References

(* Studies included in the review are denoted with an asterisk)

Akturk, A. O., & Sahin, I. (2011). Literature review on metacognition and its measurement. *Procedia - Social and Behavioral Sciences*, *15*, 3731–3736. https://doi.org/10.1016/j.sbspro.2011.04.364

Alexander, P. A. (2020). Methodological guidance paper: The art and science of quality systematic reviews. *Review of Educational Research*, *90*(1), 6–23. https://doi.org/10.3102/0034654319854352

Anderman, E. M., Eccles, J. S., Yoon, K. S., Roeser, R., Wigfield, A., & Blumenfeld, P. (2001). Learning to value mathematics and reading: Relations to mastery and performance-oriented instructional practices. *Contemporary Educational Psychology*, *26*(1), 76–95. https://doi.org/10.1006/ceps.1999.1043

Balagtas, M. U. (2021). Alignment of the Philippine mathematics teacher education curriculum with the programme for international student assessment. *European Journal of Mathematics and Science Education, 2*(2), 145–161. https://doi.org/10.12973/ejmse.2.2.145

* Bardach, L., Yanagida, T., Morin, A. J. S., & Lueftenegger, M. (2021). Is everyone in class in agreement and why (not)? Using student and teacher reports to predict within-class consensus on goal structures. *Learning and Instruction*, *71*, 1–14. https://doi.org/10.1016/j.learninstruc.2020.101400

* Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*(1), 133–180. https://doi.org/10.3102/0002831209345157

* Bellens, K., van Damme, J., van den Noortgate, W., Wendt, H., & Nilsen, T. (2019). Instructional quality: catalyst or pitfall in educational systems' aim for high achievement and equity? An answer based on multilevel SEM analyses of TIMSS 2015 data in Flanders (Belgium), Germany, and Norway. *Large-Scale Assessments in Education*, *7*, 1–27. https://doi.org/10.1186/s40536-019-0069-2

Berliner, D. C. (2005). The near impossibility of testing for teacher quality. *Journal of Teacher Education*, *56*(3), 205–213. https://doi.org/10.1177/0022487105275904

Bruder, R. (2018). Fachliche Unterrichtsqualität im Kontext der Basisdimensionen guten Unterrichts aus mathematikdidaktischer Perspektive [Subject-related teaching quality in the context of the basic dimensions of good teaching from a mathematical-didactic perspective]. In M. Martens, K. Rabenstein, K. Bräu, M. Fetzer, H. Gresch, I. Hardy, & C. Schelle (Eds.), *Konstruktionen von Fachlichkeit: Ansätze, Erträge und Diskussionen in der* *empirischen Unterrichtsforschung* [Constructions of professionalism: approaches, outputs and discussions in empirical classroom research] (pp. 203–218). Klinkhardt, Julius.

Brunner, E. (2018). Quality of mathematics teaching: A question of perspective. *Journal für Mathematik-Didaktik*, *39*(2), 257–284. https://doi.org/10.1007/s13138-017-0122-z

Busse, A. (2015). *Teds-Unterricht: Ziele und Forschungsfragen* [Teds-instruct: Targets and research questions]. Universität Hamburg. https://bit.ly/3HEkSpT

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

* Cheema, J. R., & Kitsantas, A. (2014). Influences of disciplinary classroom climate on high school student self-efficacy and mathematics achievement: A look at gender and racial-ethnic differences. *International Journal of Science and Mathematics Education*, *12*(5), 1261–1279. https://doi.org/10.1007/s10763-013-9454-4

* Creemers, B. P. M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools. Contexts of learning*. Routledge. https://doi.org/10.4324/9780203939185

Creemers, B. P. M., Kyriakides, L., & Antoniou, P. (2013). *Teacher professional development for improving quality of teaching*. Springer. https://doi.org/10.1007/978-94-007-5207-8

Dagli, U. Y. (2019). Effect of increased instructional time on student achievement. *Educational Review*, *71*(4), 501–517. https://doi.org/10.1080/00131911.2018.1441808

De Corte, E. (2004). Mainstreams and perspectives in research on learning (mathematics) from instruction. *Applied Psychology*, *53*(2), 279–310. https://doi.org/10.1111/j.1464-0597.2004.00172.x

* Demir, S. B. (2018). The effect of teaching quality and teaching practices on PISA 2012 mathematics achievement of Turkish students. *International Journal of Assessment Tools in Education*, *5*(4), 645–658. https://doi.org/10.21449/ijate.463409

Ditton, H. (2002). Unterrichtsqualität - Konzeptionen, methodische Überlegungen und Perspektiven [Teaching quality - concepts, methodological considerations and perspectives]. *Unterrichtswissenschaft*, *30*(3), 197–212. https://doi.org/10.25656/01:7685

Doig, B., Groves, S., Tytler, R., & Gough, A. (2005). Primary and secondary mathematics practice: How different is it? In P. Clarkson (Ed.), *Building connections: Research, Theory and Practice. Proceedings of the annual conference held at RMIT, Melbourne* (pp. 305–312). Mathematics Education Research Group of Australasia.

Drollinger-Vetter, B. (2011). *Verstehenselemente und strukturelle Klarheit: Fachdidaktische Qualität der Anleitung von mathematischen Verstehensprozessen im Unterricht* [Elements of understanding and structural clarity: Didactic quality of the instruction of mathematical processes of understanding in the classroom]. Waxmann.

* Dubberke, T., Kunter, M., McElvany, N., Brunner, M., & Baumert, J. (2008). Mathematics teachers' beliefs and their impact on instructional quality and student achievement. *Zeitschrift für Pädagogische Psychologie*, *22*(3-4), 193–206. https://doi.org/10.1024/1010-0652.22.34.193

* Eriksson, K., Helenius, O., & Ryve, A. (2019). Using TIMSS items to evaluate the effectiveness of different instructional practices. *Instructional Science*, *47*(1), 1–18. https://doi.org/10.1007/s11251-018-9473-1

* Faber, J. M., Glas, C. A. W., & Visscher, A. J. (2018). Differentiated instruction in a data-based decision-making context. *School Effectiveness and School Improvement*, *29*(1), 43–63. https://doi.org/10.1080/09243453.2017.1366342

Fenstermacher, G. D., & Richardson, V. (2005). On making determinations of quality in teaching. *Teachers College Record*, *107*(1), 186–213. https://doi.org/10.1111/j.1467-9620.2005.00462.x

Fischer, R. (1998). Wissenschaft und Bewusstsein der Gesellschaft [Science and awareness in society]. *Zeitschrift für Hochschuldidaktik*, *3*, 106–120.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Gage, N. L., & Needels, M. C. (1989). Process-product research on teaching: A review of criticisms. *The Elementary School Journal*, *89*(9), 253–300. https://doi.org/10.1086/461577

* Gamlem, S. M. (2019). Mapping teaching through interactions and pupils' learning in mathematics. *SAGE Open*, *9*(3), 1–13. https://doi.org/10.1177/2158244019861485

Gitomer, D. H., & Bell, C. A. (Eds.). (2016). *Handbook of research on teaching* (5th ed.). American Educational Research Association. https://doi.org/10.3102/978-0-935302-48-6

Guill, K., Lüdtke, O., & Köller, O. (2020). Assessing the instructional quality of private tutoring and its effects on student outcomes: Analyses from the German National Educational Panel Study. *The British Journal of Educational Psychology*, *90*(2), 282–300. https://doi.org/10.1111/bjep.12281

Hartmann, J., Schauer, S., Krauth, C., & Amelung, V. (2012). Methoden zur Prädiktion von Hochnutzern: ein systematischer Literatur-Review [Methods for predicting heavy users: a systematic literature review]. *GMS Medizinische Informatik, Biometrie und Epidemiologie*, *8*(1), 1–11. https://doi.org/10.3205/mibe000126

Harvey, L., & Green, D. (2000). Qualität definieren: fünf unterschiedliche Ansätze [Defining quality: Five different approaches]. *Zeitschrift für Pädagogik, (Suppl. 41)*, 17–40. https://doi.org/10.25656/01:8483

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* Routledge. https://doi.org/10.4324/9780203887332

Helmke, A. (2017). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts* [Teaching quality and teacher professionalism: diagnosis, evaluation and improvement of teaching] (7th ed.). Klett Kallmeyer.

Hiebert, J., & Grouws, D. A. (2007). The effects of classoom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning: A project of the national council of teachers of mathematics* (pp. 371–404). Information Age Pub.

Jentsch, A., Schlesinger, L., Heinrichs, H., Kaiser, G., König, J., & Blömeke, S. (2021). Erfassung der fachspezifischen Qualität von Mathematikunterricht: Faktorenstruktur und Zusammenhänge zur professionellen Kompetenz von Mathematiklehrpersonen [Measuring the subject-specific quality of mathematics teaching: Factor structure and links to the professional competence of mathematics teachers]. *Journal für Mathematik-Didaktik*, *42*, 97–121. https://doi.org/10.1007/s13138-020-00168-x

* Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. (2011). Identifying effective classroom practices using student achievement data. *Journal of Human Resources*, *46*(3), 587–613. https://doi.org/10.1353/jhr.2011.0010

* Kelcey, B., Hill, H. C., & Chin, M. J. (2019). Teacher mathematical knowledge, instructional quality, and student outcomes: A multilevel quantile mediation analysis. *School Effectiveness and School Improvement*, *30*(4), 398–431. https://doi.org/10.1080/09243453.2019.1570944

* Kistner, S., Rakoczy, K., Otto, B., van Dignath Ewijk, C., Buettner, G., & Klieme, E. (2010). Promotion of self-regulated learning in classrooms: investigating frequency, quality, and consequences for student performance. *Metacognition and Learning*, *5*(2), 157–171. https://doi.org/10.1007/s11409-010-9055-3

Klieme, E. (2013). The role of large-scale assessments in research on educational effectiveness and school development. In M. von Davier, E. Gonzalez, I. Kirsch, & K. Yamamoto (Eds.), *The role of international large-scale assessments: Perspectives from technology, economy, and educational research* (pp. 115–148). Springer. https://doi.org/10.1007/978-94-007-4629-9

Klieme, E., Lipowsky, F., Rakoczy, K., & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht: theoretische Grundlagen und ausgewählte Ergebnisse des Projektes "Pythagoras" [Quality dimensions and effectiveness of mathematics teaching: theoretical foundations and selected results of the "Pythagoras" project]. In M. Prenzel & L. Allolio-Näcke (Eds.), *Untersuchungen zur Bildungsqualität von Schule: Abschlussbericht des DFG-Schwerpunktprogramms* [Studies on the educational quality of schools: Final report of the DFG priority programme] (pp. 127–146). Waxmann.

Klieme, E., & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik: Outcome-orientierte Messung und Prozessqualität des Unterrichts [Empirical teaching research and subject didactics: Outcome-oriented measurement and process quality of teaching]. *Zeitschrift für Pädagogik*, *54*(2), 222–237. https://doi.org/10.25656/01:4348

Klieme, E., Schümer, G., & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur" und Unterrichtsgestaltung [Mathematics teaching in lower secondary education: "task culture" and lesson design]. In Bundesministerium für Bildung und Forschung (Ed.), *TIMSS - Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente* [TIMSS - Impulses for schools and teaching. Research findings, reform initiatives, practice reports and video documents] (pp. 43–58). BMBF.

König, J., Blömeke, S., Jentsch, A., Schlesinger, L., Nehls, C. F. N., 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. https://doi.org/10.1007/s10649-020-10021-0

Krainer, K. (2005). Was guter Mathematikunterricht ist, müssen Lehrende ständig selbst erarbeiten!: Spannungsfelder als Orientierung zur Gestaltung von Unterricht [Teachers must constantly work out for themselves what good mathematics teaching is!: Areas of conflict as orientation for the design of lessons]. In C. Kaune, I. Schwank, & J.

Sjuts (Eds.), *Schriftenreihe des Forschungsinstituts für Mathematikdidaktik: Vol. 40,1. Mathematikdidaktik im Wissenschaftsgefüge: Zum Verstehen und Unterrichten mathematischen Denkens; Festschrift für Elmar Cohors-Fresenborg* [Mathematics Didactics in the Structure of Science: On Understanding and Teaching Mathematical Thinking] (1st ed., pp. 165–178). Forschungsinstitut für Mathematikdidaktik Osnabrück.

Krapp, A. (1999). Intrinsische Lernmotivation und Interesse: Forschungsansätze und konzeptuelle Überlegungen [Intrinsic learning motivation and interest: Research approaches and conceptual considerations]. *Zeitschrift für Pädagogik*, *45*(3), 387–406. https://doi.org/10.25656/01:5958

Kuckartz, U. (2014). *Qualitative Text Analysis: A Guide to Methods, Practice & Using Software*. SAGE Publications Ltd. https://doi.org/10.4135/9781446288719

* Kuger, S., Klieme, E., Luedtke, O., Schiepe-Tiska, A., & Reiss, K. (2017). Student learning in secondary school mathematics classrooms: On the validity of student reports in international large-scale studies. *Zeitschrift für Erziehungswissenschaft*, *20*(2), 61–98. https://doi.org/10.1007/s11618-017-0750-6

Kunter, M. (2005). *Multiple Ziele im Mathematikunterricht* [Multiple objectives in mathematics teaching]. Waxmann.

* Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional Competence of Teachers: Effects on Instructional Quality and Student Development. *Journal of Educational Psychology*, *105*(3), 805–820. https://doi.org/10.1037/a0032583

Learning Mathematics for Teaching Project. (2011). Measuring the mathematical quality of instruction. *Journal of Mathematics Teacher Education*, *14*(1), 25–47. https://doi.org/10.1007/s10857-010-9140-1

Leibniz Institute for Research and Information in Education. (n.d.). *Unterrichtsqualität und mathematisches Verständnis in verschiedenen Unterrichtskulturen* [Teaching quality and mathematical understanding in different teaching cultures]. https://www.dipf.de/de/forschung/projektarchiv/Pythagoras

Lindmeier, A., & Heinze, A. (2020). Die fachdidaktische Perspektive in der Unterrichtsqualitätsforschung: (bisher) ignoriert, implizit enthalten oder nicht relevant? [The subject didactic perspective in teaching quality research: (so far) ignored, implicitly included or not relevant?]. *Zeitschrift für Pädagogik, (Suppl. 66),* 255–268.

* Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, *19*(6), 527–537. https://doi.org/10.1016/j.learninstruc.2008.11.001

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Ni, Y., Li, Q., Li, X., & Zhang, Z.-H. (2011). Influence of curriculum reform: An analysis of student mathematics achievement in Mainland China. *International Journal of Educational Research*, *50*(2), 100–116. https://doi.org/10.1016/j.ijer.2011.06.005

* Ni, Y., Zhou, D. -H. R., Cai, J., Li, X., Li, Q., & Sun, I. X. (2018). Improving cognitive and affective learning outcomes of students through mathematics instructional tasks of high cognitive demand. *Journal of Educational Research*, *111*(6), 704–719. https://doi.org/10.1080/00220671.2017.1402748

Nilsen, T., & Gustafsson, J. -E. (Eds.). (2016). *Teacher quality, instructional quality and student outcomes* (Vol. 2). Springer International Publishing. https://doi.org/10.1007/978-3-319-4122-8

Nilsen, T., Gustafsson, J.-E., & Blömeke, S. (2016). Conceptual framework and methodology of this report. In T. Nilsen & J.-E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes* (pp. 1–20). Springer International Publishing. https://doi.org/10.1007/978-3-319-41252-8

Nückles, M., & Wittwer, J. (2014). Lernen und Wissenserwerb [Learning and knowledge acquisition]. In T. Seidel & A. Krapp (Eds.), *Psychologie 2014. Pädagogische Psychologie: Mit Online-Materialien* (6th ed., pp. 225–252). Julius Beltz.

* Ozberk, E. B. U., Findik, L. Y., & Ozberk, E. H. (2018). Investigation of the variables affecting the math achievement of resilient students at school and student level. *Education and Science/ Egitim ve Bilim, 43*(194), 111–129. https://doi.org/10.15390/EB.2018.7153

* Ozdemir, A. S., & Sahal, M. (2018). The effect of teaching integers through the problem posing approach on students' academic achievement and mathematics attitudes. *Eurasian Journal of Educational Research,* (78), 117–137. https://doi.org/10.14689/ejer.2018.78.6

* Pauli, C., Drollinger-Vetter, B., Hugener, I., & Lipowsky, F. (2008). Cognitive activation and mathematics instruction. *Zeitschrift für Pädagogische Psychologie*, 22(2), 127–133. https://doi.org/10.1024/1010-0652.22.2.127

Perry, B., Wong, N. -Y., & Howard, P. (2006). Comparing primary and secondary mathematics teachers' beliefs about mathematics, mathematics learning and mathematics teaching in Hong Kong and Australia. In F. K. S. Leung, K.-D. Graf, & F. J. Lopez-Real (Eds.), *New ICMI study series: Vol. 9. Mathematics education in different cultural traditions - a comparative study of East Asia and the West: The 13th ICMI study* (Vol. 9, pp. 435–448). Springer. https://doi.org/10.1007/0-387-29723-5_26

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*(2), 109–119. https://doi.org/10.3102/0013189X09332374

Pigott, T. D., & Polanin, J. R. (2020). Methodological guidance paper: High-quality meta-analysis in a systematic review. *Review of Educational Research*, *90*(1), 24–46. https://doi.org/10.3102/0034654319877153

* Pinger, P., Rakoczy, K., Besser, M., & Klieme, E. (2018). Implementation of formative assessment - effects of quality of programme delivery on students' mathematics achievement and interest. *Assessment in Education-Principles Policy & Practice*, *25*(2), 160–182. https://doi.org/10.1080/0969594X.2016.1170665

Praetorius, A.-K., Herrmann, C., Gerlach, E., Zülsdorf-Kersting, M., Heinitz, B., & Nehring, A. (2020). Unterrichtsqualität in den Fachdidaktiken im deutschsprachigen Raum – zwischen Generik und Fachspezifik [Teaching quality in subject didactics in German-speaking countries - between generics and subject specifics]. *Unterrichtswissenschaft*, *48*(3), 409–446. https://doi.org/10.1007/s42010-020-00082-8

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

Praetorius, A.-K., Klieme, E., Kleickmann, T., Brunner, E., Lindmeier, A., Taut, S., & Charalambous, C. Y. (2020). Towards developing a theory of generic teaching quality: origin, current status, and necessary next steps regarding the three basic dimensions model. *Zeitschrift für Pädagogik, (Suppl. 66),* 15–36.

Praetorius, A.-K., Pauli, C., Reusser, K., Rakoczy, K., & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction*, *31*, 2–12. https://doi.org/10.1016/j.learninstruc.2013.12.002

Praetorius, A.-K., Rogh, W., & Kleickmann, T. (2020). Blinde Flecken des Modells der drei Basisdimensionen von Unterrichtsqualität? Das Modell im Spiegel einer internationalen Synthese von Merkmalen der Unterrichtsqualität [Blind spots of the model of the three basic dimensions of quality of teaching? The model in light of an international synthesis of characteristics of quality of teaching]. *Unterrichtswissenschaft*, *48*, 303–318. https://doi.org/10.1007/s42010-020-00072-w

Praetorius, A.-K., Vieluf, S., Saß, S., Bernholt, A., & Klieme, E. (2016). The same in German as in English? Investigating the subject-specificity of teaching quality. *Zeitschrift für Erziehungswissenschaft*, *19*(1), 191–209. https://doi.org/10.1007/s11618-015-0660-4

Randler, C., & Bogner, F. (2002). Comparing methods of instruction using bird species identification skills as indicators. *Journal of Biological Education*, *36*(4), 181–188. https://doi.org/10.1080/00219266.2002.9655830

* Reusser, K., & Pauli, C. (2013). Verständnisorientierung in Mathematikstunden erfassen: Ergebnisse eines methodenintegrativen Ansatzes [Measuring comprehension orientation in mathematics lessons: Results of a method-integrative approach]. *Zeitschrift für Pädagogik*, *59*(3), 308–335. https://doi.org/10.25656/01:11940

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Scheerens, J., & Bosker, R. J. (1997). *The foundations of educational effectiveness* (1st ed.). Pergamon.

Scheerens, J., Witziers, B., & Steen, R. (2013). A meta-analysis of school effectiveness studies. *Revista De Educación,* (361), 619–645. https://doi.org/10.4438/1988-592X-RE-2013-361-235

Schlesinger, L., & Jentsch, A. (2016). Theoretical and methodological challenges in measuring instructional quality in mathematics education using classroom observations. *ZDM Mathematics Education*, *48*(1-2), 29–40. https://doi.org/10.1007/s11858-016-0765-0

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. https://doi.org/10.1007/s11858-018-0917-5

Schratz, M. (2009). „Lernseits" von Unterricht: Alte Muster, neue Lebenswelten - was für Schulen? [Paradigm shift in learning culture]. *Lernende Schule*, *12*(46-47), 16–21.

* Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and

research design in disentangling meta-analysis results. *Review of Educational Research*, *77*(4), 454–499. https://doi.org/10.3102/0034654307310317

* Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education*, *62*(4), 339–355. https://doi.org/10.1177/0022487111404241

Toste, J. R., Didion, L., Peng, P., Filderman, M. J., & McClelland, A. M. (2020). A meta-analytic review of the relations between motivation and reading achievement for k–12 students. *Review of Educational Research*, *90*(3), 420–456. https://doi.org/10.3102/0034654320919352

van den Heuvel-Panhuizen, M., & Drijvers, P. (2014). Realistic mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (Vol. 1, pp. 521–525). Springer. https://doi.org/10.1007/978-94-007-4978-8_170

* Vanlaar, G., Kyriakides, L., Panayiotou, A., Vandecandelaere, M., McMahon, L., de Fraine, B., & van Damme, J. (2015). Do the teacher and school factors of the dynamic model affect high- and low-achieving student groups to the same extent? a cross-country study. *Research Papers in Education*, *31*(2), 183–211. https://doi.org/10.1080/02671522.2015.1027724

von Elm, E., Schreiber, G., & Haupt, C. C. (2019). Methodische Anleitung für Scoping Reviews (JBI-Methodologie) [Methodological manual for scoping reviews]. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, *143*, 1–7. https://doi.org/10.1016/j.zefq.2019.05.004

Walkowiak, T. A., Berry, R. Q., Meyer, J. P., Rimm-Kaufman, S. E., & Ottmar, E. R. (2014). Introducing an observational measure of standards-based mathematics teaching practices: Evidence of validity and score reliability. *Educational Studies in Mathematics*, *85*(1), 109–128. https://doi.org/10.1007/s10649-013-9499-x

Walshaw, M., & Anthony, G. (2008). The teacher's role in classroom discourse: A review of recent research into mathematics classrooms. *Review of Educational Research*, *78*(3), 516–551. https://doi.org/10.3102/0034654308320292

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, *73*(1), 89–122. https://doi.org/10.3102/00346543073001089

Weinert, F. E., Schrader, F. -W., & Helmke, A. (1989). Quality of instruction and achievement outcomes. *International Journal of Educational Research*, *13*(8), 895–914. https://doi.org/10.1016/0883-0355(89)90072-4

* Yi, H. S., & Lee, Y. (2017). A latent profile analysis and structural equation modeling of the instructional quality of mathematics classrooms based on the PISA 2012 results of Korea and Singapore. *Asia Pacific Education Review*, *18*(1), 23–39. https://doi.org/10.1007/s12564-016-9455-4

* Zepeda, C. D., Hlutkowsky, C. O., Partika, A. C., & Nokes-Malach, T. J. (2019). Identifying teachers' supports of metacognition through classroom talk and its relation to growth in conceptual learning. *Journal of Educational Psychology*, *111*(3), 522–541. https://doi.org/10.1037/edu0000300

Appendices

Table A1. Formed Categories and Subcategories with Example Indicators

Categories and subcategories	Example indicators	Study ID (see Table 1)	# num
Classroom management			16
Dealing with disruptions	The teacher makes clear what happens if students break the rules (Praetorius et al., 2018, p. 412) The teacher successfully prevents emerging disturbances immediately (Schlesinger et al., 2018, p. 487) Die Lehrerin / der Lehrer muss lange warten, bis die Schülerinnen und Schüler ruhig werden (Kuger et al., 2017, p. 91) [The teacher has to wait a long time until the pupils are quiet]	1, 3, 4, 6, 10, 12, 14, 15, 16, 18, 23, 25, 26	13
Supporting learning processes	At the beginning of the lesson, the teacher starts with what we covered in the previous lessons (Vanlaar et al., 2015, p. 205) The lesson is separated clearly into sections (Schlesinger et al., 2018, p. 487)	5, 10, 11, 12, 18, 25, 26	7
Managing learning time	The lesson starts and ends on time (Schlesinger et al., 2018, p. 487) Transitions between lesson phases run smoothly (Schlesinger et al., 2018, p. 487) The teacher clearly communicates the targeted time for tasks (e.g., '5 min left') (Praetorius et al., 2018, p. 412)	1, 15, 16, 18, 25, 26	6
Student support			18
Feedback (teacher's feedback to students)	Die Schülerinnen und Schüler erhalten differenzierte Leistungsrückmeldung (Ministerium für Bildung, Jugend und Sport, 2012, p. 35) [Pupils receive differentiated feedback] The teacher's feedback is constructive (Schlesinger et al., 2018, p. 487)	3, 10, 14, 16, 18, 21, 23, 24	8
Constructive approach to errors	The teacher uses mistakes to make students learn from them (Praetorius et al., 2018, p. 413) Mit Schülerfehlern wird konstruktiv umgegangen (Ministerium für Bildung, Jugend und Sport, 2012, p. 34) [Students' mistakes are handled constructively]	1, 3, 7, 15	4
Appreciation of students	Der Umgangston zwischen der Lehrkraft und den Schülerinnen und Schülern ist wertschätzend und respektvoll (Ministerium für Bildung, Jugend und Sport, 2012, p. 34) [Teachers and students appreciate and respect each other] The teacher is interested in the students' perspectives and opinions (Praetorius et al., 2018, p. 413)	1, 3, 5, 7, 10, 15, 16, 23	8
Teacher support	The teacher asks about students' individual difficulties/individual progress (Schlesinger et al., 2018, p. 487) The teacher provides exercises with different difficulty levels (Praetorius et al., 2018, p. 412) Unsere Lehrerin/unser Lehrer erklärt etwas so lange, bis wir es verstehen (Kuger et al., 2017, p. 90) [Our teacher explains until we understand]	1, 2, 3, 7, 9, 11, 15, 18, 24, 26	10
Student orientation	Students can choose between different solution strategies (Praetorius et al., 2018, p. 413) The teacher initiates collaborative learning processes between students (Schlesinger et al., 2018, p. 487)	1, 6, 7, 14, 15, 18, 24, 26	8

Table A1. Continued

Categories and subcategories	Example indicators	Study ID (see Table 1)	# num
Cognitive activation			18
Challenging tasks	Die Lehrerin/der Lehrer stellt Aufgaben, bei denen wir das Gelernte in neuen Zusammenhängen anwenden müssen (Kuger et al., 2017, p. 90) [The teacher sets tasks in which we have to apply what we have learned in new contexts] The teacher presents problems for which there is no immediately obvious method of solution (OECD, 2014, p. 331)	1, 3, 5, 6, 7, 14, 15, 16, 17, 18, 20, 26, 27	13
Challenging questions/ class discussions	Im Mathematikunterricht stellen öfters verschiedene Schüler*innen ihre Lösungswege für eine Aufgabe vor (Dubberke et al., 2008, p. 197) [In mathematics lessons, different pupils often present their solutions to a problem] The teacher spontaneously encourages cognitive challenges (Schlesinger et al., 2018, p. 487) The teacher asks questions that make us reflect on the problem (OECD, 2014, p. 331)	3, 5, 6, 7, 10, 11, 14, 16, 18, 20, 25, 26	12
Activation of prior knowledge	The teacher activates and explores students' prior knowledge (Schlesinger et al., 2018, p. 488) Students are asked to brainstorm about a topic (Praetorius et al., 2018, p. 414) When asking about prior knowledge, the teacher is not only interested in a single, specific answer (Praetorius et al., 2018, p. 414)	16	1
Support through metacognition	Unser Mathematiklehrer/unsere Mathematiklehrerin verlangt häufig, dass wir unsere Arbeitsschritte ausführlich begründen (Dubberke et al., 2008, p. 197) [Our mathematics teacher often demands that we explain the steps we took to arrive at the solution in detail] The teacher provides time for metacognitive processes (Schlesinger et al., 2018, p. 487) Students reflect on their learning processes (Schlesinger et al., 2018, p. 487)	7, 10, 14, 16, 23, 25, 27	7
Subject-didactic specifics			9
Relation to students' daily lives	The teacher addresses the relevance of content (Schlesinger et al., 2018, p. 488) Students may bring in their own experience and interests to class (Schlesinger et al., 2018, p. 488) Die Lehrkraft stellt einen klaren Bezug zur Alltagswelt und/oder Berufswelt her (Ministerium für Bildung, Jugend und Sport, 2012, p. 33)) [The teacher makes a clear reference to the everyday and/or professional world]	8, 14, 18	3
Mathematical connections/mathematical depth	Mathematical generalizations are developed (Schlesinger et al., 2018, p. 488) The topic is related to other mathematical topics (Schlesinger et al., 2018, p. 488) The content is embedded in a broader mathematical structure (Schlesinger et al., 2018, p. 488) Die Verknüpfungen der Repräsentationsformen tragen zum Aufbau von Verständnis bei (Drollinger-Vetter, 2011, p. 257) [The links between the forms of representation contribute to building understanding]	20, 22, 24	3
Teacher's mathematical accuracy	The teacher does not make any content-related or formal mistakes (Schlesinger et al., 2018, p. 488) The teacher is precise concerning mathematical language and notation (Schlesinger et al., 2018, p. 488)	12	1

Note. # num Number of articles in respective category. The example indicators are printed in their original language, but translations are provided for languages other than English.

Table A2. The Effects of Individual Characteristics on Student Performance in Mathematics by Study ID

Study ID	Classroom management	Student support	Cognitive activation	Subject-didactic specifics	Miscellany
1	+ Effective classroom management	~ Individual learning support	+ Cognitive level of tasks + Curricular level of tasks		
2		+ Instructional clarity			
3	~ Classroom management (BEL - Flanders) + Classroom management (DEU) + Classroom management (NOR)	+ Supportive climate (BEL - Flanders) ~ Supportive climate (DEU) ~ Supportive climate (NOR)	- Cognitive activation (Component 2) (BEL - Flanders)		
4	+ Disciplinary climate in math lessons				
5	~ Orientation ~ Structuring ~ Time management	~ Classroom-level environment (Teacher-student relationships and relationships among students)	~ Questioning ~ Teaching-modeling		+ Application ~ Assessment
6	+ Disciplinary climate	- Teachers' student orientation	+ Cognitive activation strategies		
7		- Konstruktive Unterstützung [Student support]	+ Kognitive Aktivierung [Cognitive activation]		
8				- We relate what we are learning in mathematics to our daily lives (SWE) + We memorize formulas and procedures (SWE) ~ We relate what we are learning in mathematics to our daily lives (45 countries in TIMSS, 2003) ~ We relate what we are learning in mathematics to our daily lives (50 countries in TIMSS, 2007) ~ We memorize formulas and procedures (50 countries in TIMSS, 2007)	+ We listen to the teacher give a lecture-style presentation (SWE) ~ We listen to the teacher give a lecture-style presentation (45 countries in TIMSS, 2003) ~ We listen to the teacher give a lecture-style presentation (50 countries in TIMSS, 2007)
9		~ Differentiated instruction			
10	+ Behavior management ~ Productivity	+ Positive climate ~ Teacher sensitivity ~ Regard for adolescent perspectives ~ Negative climate ~ Quality of feedback	~ Analysis and inquiry ~ Instructional dialogue	~ Content understanding	~ Instructional learning formats

Table A2. Continued

Study ID	Classroom management	Student support	Cognitive activation	Subject-didactic specifics	Miscellany
11	+ Classroom environment relative to instructional practices		~ Questions & discussion approach relative to standards & content focus		
12	~ Classroom organization			~ Errors	
13		~ Total number of observed strategies - instructions ~ Direct (organization; monitoring and evaluation; resource management) ~ Direct (elaboration; planning and systematic activity; feedback) ~ Explicit ~ Implicit + Indirect (constructivism, transfer) ~ Indirect (cooperation, self-direction)			
14	~ Regelklarheit und Struktur [Clarity of rules and structure] (*) + Regelklarheit und Struktur [Clarity of rules and structure]	~ Konstruktive Unterstützung [Student support] (*) ~ Konstruktive Unterstützung [Student support] ~ Steuerung durch die Lehrkraft [Management by the teacher] (*) ~ Steuerung durch die Lehrkraft [Management by the teacher] ~ Lehrerfeedback [Teacher feedback] (*) - Lehrerfeedback [Teacher feedback]	~ Potential zur kognitiven Aktivierung [Potential for cognitive activation] (*) ~ Potential zur kognitiven Aktivierung [Potential for cognitive activation]	+ Erfahrungen Schulalgebra [Experience with pure mathematical tasks] (*) + Erfahrungen Schulalgebra [Experience with pure mathematical tasks] ~ Erfahrungen angewandte Aufgaben [Experience with applied mathematical tasks] (*) - Erfahrungen angewandte Aufgaben [Experience with applied mathematical tasks]	~ Schülerorientiertes Unterrichten [Student-oriented teaching] (*) ~ Schülerorientiertes Unterrichten [Student-oriented teaching]
15	+ Classroom management	~ Learning support	+ Cognitive activation		
16	+ Classroom management	~ Supportive climate	+ Cognitive activation		

Table A2. Continued

Study ID	Classroom management	Student support	Cognitive activation	Subject-didactic specifics	Miscellany
17			~ Effect of "Cognitive demands of mathematics instructional tasks" to calculation, routine problem solving and complex problem-solving + Effect of "External (multiple) representations of mathematics instructional tasks" to solving complex problems ~ Effect of "External (multiple) representations of mathematics instructional tasks" to calculation and routine problem-solving ~ Multiple solution methods of calculation, routine problem-solving, and complex problem-solving		
18	~ Classroom management (*) ~ Disciplinary climate (*)	+ Teacher support (*) ~ Teacher behavior – student orientation (*) ~ Teacher behavior – formative assessment (*)	~ Cognitive activation (*)	+ Experience with applied mathematical tasks (*) ~ Experience with pure mathematical tasks + Familiarity with mathematical concepts ~ Familiarity with mathematics concepts (corrected for overclaiming)	~ Teacher behavior, teacher-directed instruction (*)
19			+ Problem-posing approach		
20			+ Länge der Schülerbeiträge im Klassengespräch [Length of student contribution in class discussions] ~ Anteil anspruchvolles Üben [Proportion of challenging practice]	+ Strukturelle Klarheit der Theoriephasen [Strucural clarity]	
21		- Number of comments (*) - Specificity (*) + Feedback utilization I + Feedback utilization II			

Table A2. Continued

Study ID	Classroom management	Student support	Cognitive activation	Subject-didactic specifics	Miscellany
22				+ Strukturelle Klarheit [Strucural clarity]	~ Unterrichtsinszenierung [Lesson design] + Klassengespräche (ko-konstruktives Gespräch) [Class discussions (co-constructive)]
23	~ Organization of learning	~ Time of learning ~ Social/direct experiences ~ Regulation/monitoring	~ Basic processing	+ Domain-specific processing	~ Social context ~ Evaluation of learning ~ Goal setting and orientation
24	+ Classroom management ~ Instructional focus on learning	~ Assessment of understanding ~ Feedback ~ Instructional clarity		~ Instructional complexity	~ Instructional differentiation ~ Expectations for student learning
25	+ Quantitative structuring ~ Qualitative structuring + Time management + Dealing with misbehavior	+ Teacher-student interaction	+ Questioning ~ Teaching-modeling		+ Application + Assessment
26	~ Class management (KOR) - Class management (SGP)	+ Learning support (KOR) + Learning support (SGP)	+ Cognitive activation (KOR) + Cognitive activation (SGP)		- Teacher-student orientation (student-oriented instruction) (KOR) - Teacher-student orientation (student-oriented instruction) (SGP)
27			+ Metacognitive support		

Note. (*) Calculated at the individual level; Variable names were taken directly from the articles and translated [in brackets]. The country codes refer to the ISO 3166 (alpha-3) standard.