ZDM – Mathematics Education (2024) 56:265–280
https://doi.org/10.1007/s11858-024-01548-0

ORIGINAL PAPER

Students' mathematics self-efficacy: a scoping review

Karin Elisabeth Sørlie Street[1] · Lars-Erik Malmberg[2] · Stanislaw Schukajlow[3]

Accepted: 13 January 2024 / Published online: 17 February 2024
© The Author(s) 2024, corrected publication 2024

Abstract
Students' mathematics self-efficacy (MSE) is strongly associated with learning behaviours and performance, and students' future career choices. In our scoping review, we screened what *substantive foci* (conceptualization, directionality and role of MSE, change in MSE, and situational specificity of MSE) have been posed and which *methodological approaches* (participants, analytical methods, data sources, and congruence of measures) have been used in recent (2018–2022) studies of MSE. Studies of MSE were clearly in the mathematics domain with 21 of 49 included studies exploring specific mathematics areas. The key focus was on strength of MSE. International databases (i.e., PISA) have enabled broad generalization, while in-depth qualitative studies enable minute situation-specificity. Cross-sectional and longitudinal studies using hierarchically nested designs (i.e., students in classrooms) enable us to draw conclusions at both individual, classroom, and school levels. The current state-of-the-art indicates methodological advancement is rapid and we are likely to see further methodological-substantive synergies in the field of MSE in future studies. We see the potential and need for future mixed-methods studies that continue the focus on MSE as a multidimensional and dynamic concept. Careful consideration of the theoretical background of the construct of MSE continues to be important to bring the field forward.

Keywords Motivation · Self-efficacy · Affect · Review · Beliefs

1 Introduction

The term self-efficacy (SE) refers to individual's confidence they can carry out future tasks and challenges (Bandura, 1997 p. 3). Previous meta-analyses have demonstrated the importance of academic SE in terms of its relationship with performance (e.g., Honicke & Broadbent, 2016) and demonstrated the contribution of SE as a predictor, and not only as an outcome, of performance (e.g., Talsma et al., 2018). Students' mathematics self-efficacy (MSE) is important because it is related to motivation for learning and studying mathematics, effort and perseverance, performance, and future career choices (see, e.g., Martin et al., 2015; Pajares & Graham, 1999).

✉ Karin Elisabeth Sørlie Street
karin.street@hvl.no

[1] Western Norway University for Applied Sciences, Sogndal, Norway

[2] University of Oxford, Oxford, UK

[3] Institute of Mathematics Education and Computer Science Education, University of Münster, Münster, Germany

Within education SE is studied in numerous domains, e.g., reading, language, music, physical activity, and, of course, mathematics. Investigations have targeted students', teachers', and parents' SE. More than a decade ago, Klassen and Usher (2010) provided a review of the status of research on SE in education, and a review by Pantziara (2016) on the state-of-the-art of research on SE within the field of mathematics education is now 7 years old. Furthermore, the field of mathematics education research is dynamic, with a large increase in the number of studies focusing on emotions in mathematics education research 2014–2021 (Schukajlow et al., 2023). Considering this backdrop, gaining oversight of the various approaches to investigating this important construct as it relates to the field of mathematics education is valuable.

In this review study we took a scoping approach (Munn et al., 2018) to provide a snapshot of research on students' MSE conducted within the last 5 years. Scoping reviews are appropriate tools to report on the types of available evidence of a given field as well as the way the research has been conducted (Munn et al., 2018, p. 2). Having knowledge about the kinds of research currently conducted is important to understand the direction research on students' SE has

taken recently particularly within the field of mathematics education, considering both substantive themes and methodological approaches to investigating these themes. This might inform the field of mathematics education research whether there are important gaps in terms of research we might have a need for in the future. There is a relationship between the substantive and the methodological, where methodological advances can promote new research designs and advance our knowledge, and unresolved issues in current research can promote the use of new methods for approaching these (Marsh & Hau, 2007). To achieve our aims, we posed the following research question: Which substantive foci and methodological approaches have been used in recent (2018–2022) studies of MSE?

2 Theoretical background

The construct of SE was coined by Albert Bandura, whose conceptualization of the construct has remained the most influential. According to Bandura, SE is a multidimensional construct related to an individual's confidence he/she can accomplish specific tasks or challenges in the future (Bandura, 1997 p. 42). Within mathematics one can consider SE at the domain-level, or for specific mathematical areas or tasks. This gives a structure of SE with specific areas, in learning domains, nested within fields. Another way of distinguishing SE beliefs is whether these are specific for an individual learner or of situations, which is also related to the stability of the construct.

Figure 1 provides a model of students' MSE, including its' relationship to predictor and outcome variables. In the following sections we will build on this model to provide a brief overview of the key theoretical and methodological considerations in relation to the construct of SE, as it relates to the field of mathematics education.

2.1 Conceptualization of SE

Bandura emphasized SE as a dynamic and multidimensional form of human agency (Bandura, 1997), which can change across time and situations, that is (1) in terms of the content or task in question, (2) the level of difficulty of that task, and (3) the strength of confidence students have in their

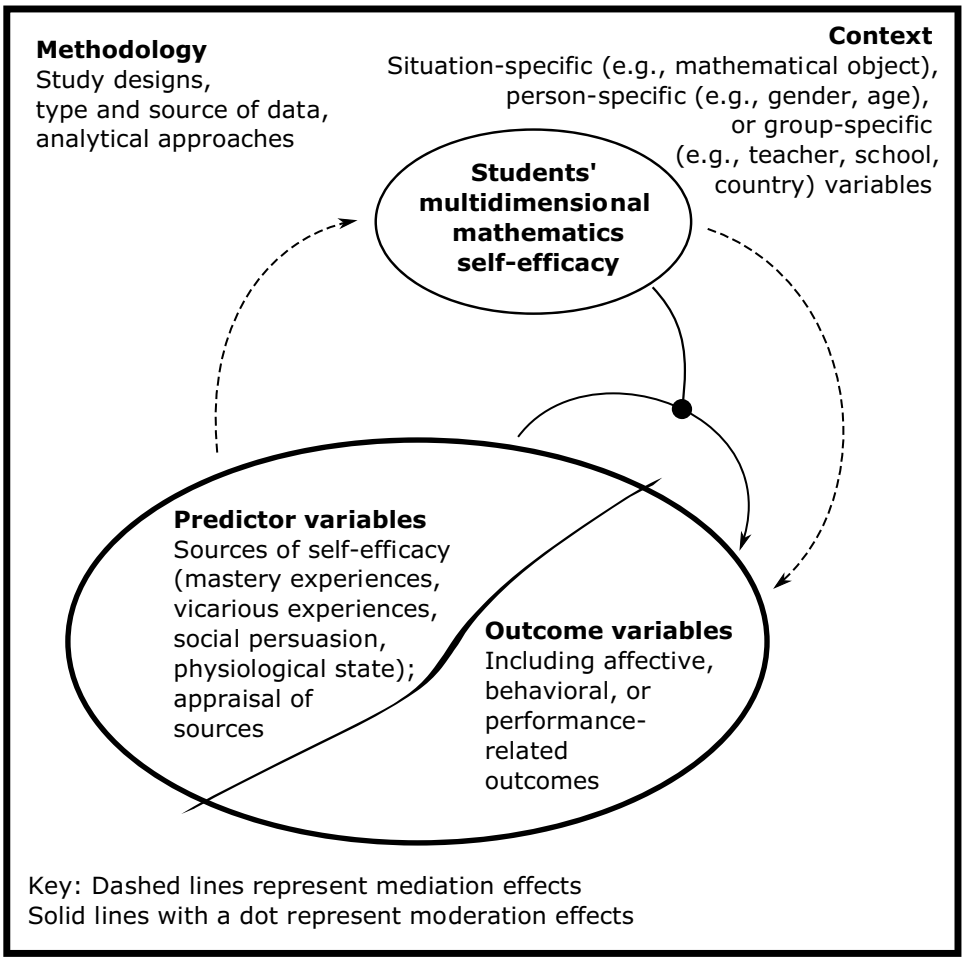

Fig. 1 Model of students' MSE

ability to carry the task out. Although Bandura highlighted multidimensionality as a key conceptual feature of SE and provided recommendations for how to design measures that incorporated all three dimensions, the single dimension of SE strength has largely dominated education research (Street et al., 2017) to the exclusion of other critical dimensions. To also investigate the role of the content specificity (e.g., what is the mathematical object or task in focus) and the level of difficulty of the content (e.g., is this task easy or hard for me) is valuable in terms of gaining detailed knowledge regarding students' MSE across different learning situations, and how best to support students' SE in classroom situations. Another way to describe MSE is through the characteristics listed by Schukajlow et al. (2023), where the construct of SE is of positive *valence*, and its *temporal stability* is likely to vary according to the assigned *object-specificity* (with higher stability for more general SE constructs).

As a type of self-belief, SE is related to but also different from other self-beliefs such as self-esteem, self-worth, outcome expectancies, and self-concept (see, e.g., discussions by Bong & Skaalvik, 2003 and Klee et al., 2022, for important distinctions between various self-belief constructs). The key distinctive features of SE are its specificity (it should be conceptualized at the domain level or more specific), its future orientation (it should be related to specific tasks or challenges in the future, e.g., a test or learning situation), its mastery orientation (including a goal focus rather than comparative focus), and its relation to individuals' confidence they can carry out tasks or accomplish challenges ("can" rather than "will" statements) (e.g., Bong & Skaalvik, 2003). In a previous review as much as 50% of the assessed research on SE was a-theoretical, that is, the studies did not follow key recommendations to measuring SE (Klassen & Usher, 2010). A more recent review by Diego-Mantecón et al. (2019) on conceptualization issues in mathematics education research indicated this problem is relevant within this field of research, and still prevails. One of the most common mix-ups is likely with the related construct of self-concept. Despite the similarities, repeated studies have concluded there are important distinctions between these two constructs (see, e.g., Klee et al., 2022; Marsh et al., 2019) and advocated for the careful consideration of theory when constructing their respective scales. While past studies have demonstrated both constructs are relevant and important to consider for the field of education, it is important to observe to what extent clarity of conceptualization and operationalization of SE prevails in the field of mathematics education in relation to other constructs which are often conflated with SE.

2.2 The relationship between SE and other constructs

SE can be studied as an independent or dependent variable, or both (see Fig. 1). Initially, many studies on SE have been concerned with establishing the importance of the construct, i.e., what is the effect of SE on other variables, e.g., Pajares and Graham (1999). This could include the effects of SE on outcomes (SE as a *predictor*), whether the effects of one variable on another is transmitted through SE (SE as *mediator*), and whether the effects of one variable on another differ according to students' SE (SE as *moderator*). As the construct has become established, more studies have focused on how other variables might affect students' SE (SE as an *outcome*, e.g., Pampaka et al., 2011). Given the theoretically proposed reciprocal determinism of Bandura's social cognitive theory (1997), which postulates a continuous reciprocal interaction between behavior, cognition, and environment, studies have also sought to determine what is the dominant direction of the relationship between SE and performance. Previous research has investigated SE in relation to, e.g., gender, school grade (/age), or cognitive constructs such as grades or test performance (Klassen & Usher, 2010) showing that SE both orients students to engagement, and is itself modified by the outcomes of engagement.

2.2.1 SE as an outcome

SE is proposed to change according to the most recent influential experiences – that is, individuals' cognitive appraisals of their learning and performance experiences (Bandura, 1997). According to Bandura, there are four key sources to students' SE beliefs. These are mastery experiences (i.e., individuals' own experiences of success or failure), vicarious experiences (i.e., the experiences of success or failure or relevant others, e.g., friends or classmates), social persuasion (i.e., verbal or non-verbal encouragement), and physiological states (i.e., bodily experiences of pulse, sweat, arousal, serenity, etc.). Importantly, the sources do not influence SE directly but are cognitively appraised, or interpreted, by the individual, such that the same experience might lead to highly different outcomes in different students. This highlights the role of students' wider self-beliefs, e.g., their ability beliefs (e.g., Dweck, 2008) and emotions (Mega et al., 2014), as potential sources of SE or moderators of the relationship between sources of SE and SE. A meta-analysis by Byars-Winston et al. (2017) confirmed mastery experiences as the most influential SE changing source, while the relative effect of the other sources appears to differ across participants and contexts (e.g., Usher & Pajares, 2008).

2.2.2 SE as a predictor

Numerous studies have demonstrated the importance of SE, in terms of positive effects on effort, perseverance, motivation, self-regulation, (reduced) anxiety, and performance experiences (Zimmerman, 2000). SE can influence performance directly through increased effort or

perseverance, or indirectly through affecting motivations or learning behaviours that in turn, lead to higher competence. SE is generally considered an affective or motivational construct and is also associated with other such constructs. Stronger SE predicts higher intrinsic motivation, effort, and perseverance, as individuals have stronger confidence in their own ability to bring about desired outcomes. In other words, it is worth trying, when you believe you have the potential to change an outcome. In contrast, a weak sense of SE is associated with low sense of agency or control. For students with experiences of failure in mathematics, a low sense of SE might mean they do not perceive that they can bring about a more positive outcome in the future, which might in turn be associated with mathematics anxiety (see, e.g., Dowker et al., 2016).

2.2.3 Reciprocal relationships

There is a theoretically proposed mutual influence between people's behaviors and cognition and their environment (reciprocal determinism, see, e.g., Bandura, 1997). The nature of this reciprocal relationship has often been considered in terms of the relationship between students' SE and their performance, where on the one hand SE can influence performance experiences, while on the other appraised performance (mastery) experiences are considered a key source of SE. Investigations have focused both on (a) to what degree is SE a reflection of, versus a contributor to, performance, (b) what is the dominant direction of effect, and (c) is the reciprocal relationship constant or does it change over time (e.g., Weidinger et al., 2018). A meta-analysis (Talsma et al., 2018) found support for the theoretically proposed reciprocal relationship between SE and performance, where the dominant direction of effect was from performance to SE. The relationships varied according to participant age, but only two of the studies included school-age children. Furthermore, the studies focused on a variety of domains (including mathematics).

2.3 Change in SE over time

Capturing the minute mechanisms of SE change (i.e., a degree of temporal stability) is challenging, considering the multiple potential influencing sources as well as the role of individual students' appraisals of these. Studies have documented domain-specific declines in, e.g., students' motivation and achievement (Gottfried et al., 2007) and perceived competence and task values (Jacobs et al., 2002) with increasing age, and investigating changes in students' MSE is important. From a quantitative perspective, at least two measurement points are required to study change (e.g., increased strength of SE between time points T_1 and T_2), while a minimum of three are needed to detect non-linear change (e.g., a steep rise in SE between T_1 and T_2, followed by a flattening trend between T_2 and T_3). Appropriate time-lags (the time between occasions) should be considered depending on whether one investigates developmental change or transitions (in years), sequences of learning in class (days or weeks), or processes during a task (minutes or hours) (McNeish & Hamaker, 2020; for an overview see Fig. 1 in Street et al., 2022). Longitudinal designs are more costly, which is probably why crossectional designs dominate the field (Pantziara, 2016).

2.4 Situational specificity of SE

While change in SE can be investigated at the situational versus the long-term perspective, differential effects of situational, personal, or group factors on students' MSE can also be investigated. Situational factors might change from one lesson to the next (e.g., the task at hand, quality of feedback), individual factors (e.g., motivation), while group factors could be considered at the classroom, school, district, or country level (e.g., the teacher of a class, the socio-economic make-up of a school). Multi-level analyses are important when students are nested within classrooms or schools, to partition variance into within and between-level variance, adjusting standard errors appropriately. Importantly, if a study includes nested data (e.g., situations nested within students or students nested within classrooms/schools), failing to account for the hierarchical structure of the data could lead to Type I errors which could give rise to atomistic fallacies (i.e., wrongly attributing the "cause" to individuals) or ecological fallacies (i.e., wrongly attributing the "cause'' to group-level characteristics) (Hox et al., 2017).

Methodological advances mean there is more data available and there are also a range of approaches for investigating this large-scale data. Within psychology there have been repeated calls that researchers should consider the principle of ergodicity (Molenaar, 2004), when investigating learning and development processes. In short, we cannot make inferences about development (change over time) based on cross-sectional findings. Theories that have been developed on the basis of crossectional designs might not fit individuals, i.e., they may be inadequate for describing or understanding the development of individual students' learning. As one important focus of the field of mathematics education is to provide knowledge for the field of practice, e.g., teachers in the classroom, this is an important point to consider. In short, it is important to consider individual variation within a group, and nestedness in terms of, e.g., time-points within individuals within classrooms within schools.

2.5 Other methodological considerations

Depending on the aim of a study many study designs and analytical approaches are possible and meaningful. Study

designs can be conceptualized in terms of the timeline of the study (e.g., cross-sectional/longitudinal), whether the aim is to describe students' SE or to change it (e.g., observational/experimental), or to what degree the study considers nestedness (e.g., time-points in students in classrooms in schools in countries). Further methodological choices related to the design of a study are the *participants* included (how many? how old? from where?). Previous studies (e.g., Ahn et al., 2016; Klassen, 2004) have demonstrated cultural differences in students' SE beliefs, in terms of both strength of SE and its relationship with performance outcomes. There is a need for research from different cultures to represent a variety of perspectives and participants, as different cultural contexts may bring different assumptions and social mechanisms, relevant to the study of MSE. Other methodological choices are related to the *analytical approaches* applied (quantitative/qualitative/mixed methods), and the *data sources* included (e.g., questionnaires/performance indicators/interviews/observations). Furthermore, when performance indicators are considered in relation to SE, the measures can be more or less *congruent* in specificity (e.g., task-specific SE and task-specific mathematics test performance/task specific SE and domain-specific mathematics test performance) as well as in relation to the timing of measurement (e.g., SE as a *predictor* of performance should be measured prior to, not after, the performance test) (Pajares & Miller, 1995).

Considering the large number of studies including a focus on emotions within mathematics education, together with the highly interrelated relationships between SE and student outcomes such as engagement and performance and a large variety of available methodological approaches to investigate these relationships, there is a need to gain oversight over the types of available evidence on students' SE in mathematics, how research is conducted in this field, and identify potential knowledge gaps. Taking a systematic approach, we conducted a scoping review (Munn et al., 2018) to investigate the kinds of substantive foci and methodological approaches that have been used to investigate students' MSE in the recent 5-year period.

3 Method

To conduct our scoping review, we developed a coding scheme and coded a selection of publications in which SE was included in the title to get a snapshot of studies in the last 5years (2018–2022). As a precursor of a systematic review or meta-analysis we followed the PRISMA guidelines for systematic (scoping) reviews of a given field (Tricco et al., 2018) and guidelines of Munn et al. (2018): to identify types of available evidence, clarify key concepts and dimensions, examine how research is conducted, identify key characteristics or factors related to a concept, and identify and analyze knowledge gaps.

3.1 Search process and eligibility criteria

We conducted a systematic search (search #1) on March 7th 2023, including the search terms self-efficac* AND math* in the title, using the Web of Science Core collection of research titles. Furthermore, we limited our search to English language, peer-reviewed, empirical reports, published from Jan 2018 to March 2023. This gave 182 hits. In addition, on June 2nd 2023, we included the search term self-efficac* in the title in six key mathematics education journals (Journal for Research in Mathematics Education, Educational Studies in Mathematics, International Journal of Science and Mathematics Education, Journal of Mathematics Teacher Education, Mathematical thinking and learning, and ZDM Mathematics Education), the other search terms and limitations being equal (search #2). Authors are perhaps less likely to include "mathematics" in the title in such journals, and we did not want to risk research in the field of mathematics education to be underrepresented. Search #2 resulted in 29 hits, of which 13 were duplicates from search #1, rendering 16 unique hits. The abstracts of the 198 hits from search #1 and #2 were screened for inclusion and exclusion criteria (see Table 1), where 108 studies were excluded. 90 hits were kept for full text screening, considering full text eligibility and (the same) inclusion and exclusion criteria. A further 41 studies were excluded, giving 49 studies that were included in our review and coded (see Fig. 2).

In terms of inclusion/exclusion criteria number 2 (conceptualization of SE), we followed previous reviews (e.g., Klassen & Usher, 2010), considering the SE measure as a whole. That is, we considered whether most items (e.g., three out of four) of a measure were congruent with theory and inclusion criteria 2a–2d (see Table 1). The response options (scale and anchors) of the measure were also included in this consideration, where the response option and the statement together should reflect the participant's degree of confidence of their ability to carry out a certain task in the future (versus, e.g., participants' intentions to do things; their expectancies of outcomes not contingent on their own actions; how often they believe things will occur; or their assessment of how good they are/how much better than others).

A large number of studies focusing on undergraduates' (36) and teachers' (42) SE were excluded from the 198 initially identified studies (see Fig. 2). Of the 90 records included for full text screening, 26 were excluded due to inadequate/inconsistent conceptualization (inclusion/exclusion criteria 2a-2d; see Fig. 2). That is, among the studies we examined more carefully 29% were excluded because their definition or operationalization of SE did not align with theoretical tenets.

Table 1 Inclusion and exclusion criteria

Inclusion criteria (IC)		Exclusion criteria (EC)	
1	Studies focusing on mathematics self-efficacy	1	Studies on self-efficacy that are not related to mathematics (e.g., general academic self-efficacy or self-efficacy for learning English; studies that focus on STEM rather than just mathematics; studies that focus on emotional self-efficacy; studies that focus on self-regulation self-efficacy)
2	Studies meeting our definition of self-efficacy	2	Studies that define and/or operationalise self-efficacy in a way that does not meet our definition (e.g., studies on self-concept or outcome expectancies)
2a	Future-oriented statements	2a	Past-oriented or general statements
2b	Statements reflecting a particular context	2b	Statements not reflecting a particular context
2c	Evaluation of confidence	2c	Evaluation of competence/ability/intention/outcome
2d	Goal-referenced	2d	Normative (comparative)
3	Studies at primary and secondary levels of mathematics education	3	Studies including a sample other than mathematics learners at primary/elementary, secondary /high school level of learning (e.g., early childhood education, graduate student education, teacher education, teachers, adults, parents)
4	Empirical, peer reviewed, studies	4	Theoretical papers, editorials, books, book reviews, commentaries, notes, meeting abstracts, literature reviews, meta analyses, review papers
5	Studies published in English	5	Studies published in a language other than English
6	Studies published between 2018 and 2023	6	Studies published before 2018
7	Studies indexed in Web of Science Core Collection	7	Studies indexed in databases other than the Web of Science Core Collection

3.2 Coding process

To code the included studies, we used a coding manual including 13 categories (see supplementary materials). Following our model (see Fig. 1), the first author developed a draft for the code, which was then discussed, adjusted, and agreed upon by all authors. Two coders (the first and second author) trial coded five papers, followed by a discussion and clarifications to the coding manual. The two coders then proceeded to independently code 27 and 22 (respectively) of the remaining 44 studies each, including an overlap on 5 articles (= 10% of the studies). Inter-rater reliability for these five articles indicated substantial agreement, with Cohen's Kappa ranging from 0.46 to 0.96, and an average of 0.71, while exact match ranged from 0.77–0.97).

4 Results

49 studies were included and coded for our review. An overview of the results (number and percentage of studies assigned to each code) is provided in Table 2. In the following sections we describe the coding results and provide examples.

4.1 Substantive foci of studies

4.1.1 Conceptualization of SE

Considering the multidimensional definition of SE, all studies included a focus on SE strength (students' degree of confidence), while only a few focused on variation of SE either as a function of the mathematical content (five studies = 10%) or the level of difficulty of the tasks (five studies = 10%). Only two studies included a focus on all three dimensions of SE (i.e., level of difficulty, strength of confidence, and specificity of the mathematical object or content area). Liu et al. (2020) investigated students' domain- and task-specific problem-posing SE and performance for different levels of difficulty problems, while Street et al. (2022) investigated starting point and change in students' SE across a series of lessons according to perceived task difficulty and mathematical content area.

Most commonly, the studies focused on students' MSE at the domain- level (55%), for example the relationship between students' perceived responsibility for learning, MSE and sources of MSE (Lau et al., 2018). A relatively large percentage of studies included students' SE for a mixture of mathematical topics (29%)—several of these studies used available PISA data (e.g., Soland, 2019), which includes different topic areas of mathematics in the SE measure. A smaller number of studies included a focus on students' SE for specific topic areas, including geometry (five studies), functions (four studies), arithmetic (three studies), algebra (two studies), or combinatorics/probability (two studies). In addition, five studies focused on students' SE for other objects of mathematics, including programming/coding (Jiang et al., 2022), mathematical modelling (Krawitz & Schukajlow, 2018), and problem posing (Q. Liu et al., 2020).

4.1.2 The relationship between SE and other variables

Across the studies, SE was investigated in terms of its role in predicting and affecting other outcomes (41%), such as mathematics achievement (Hiller et al., 2022), mathematics

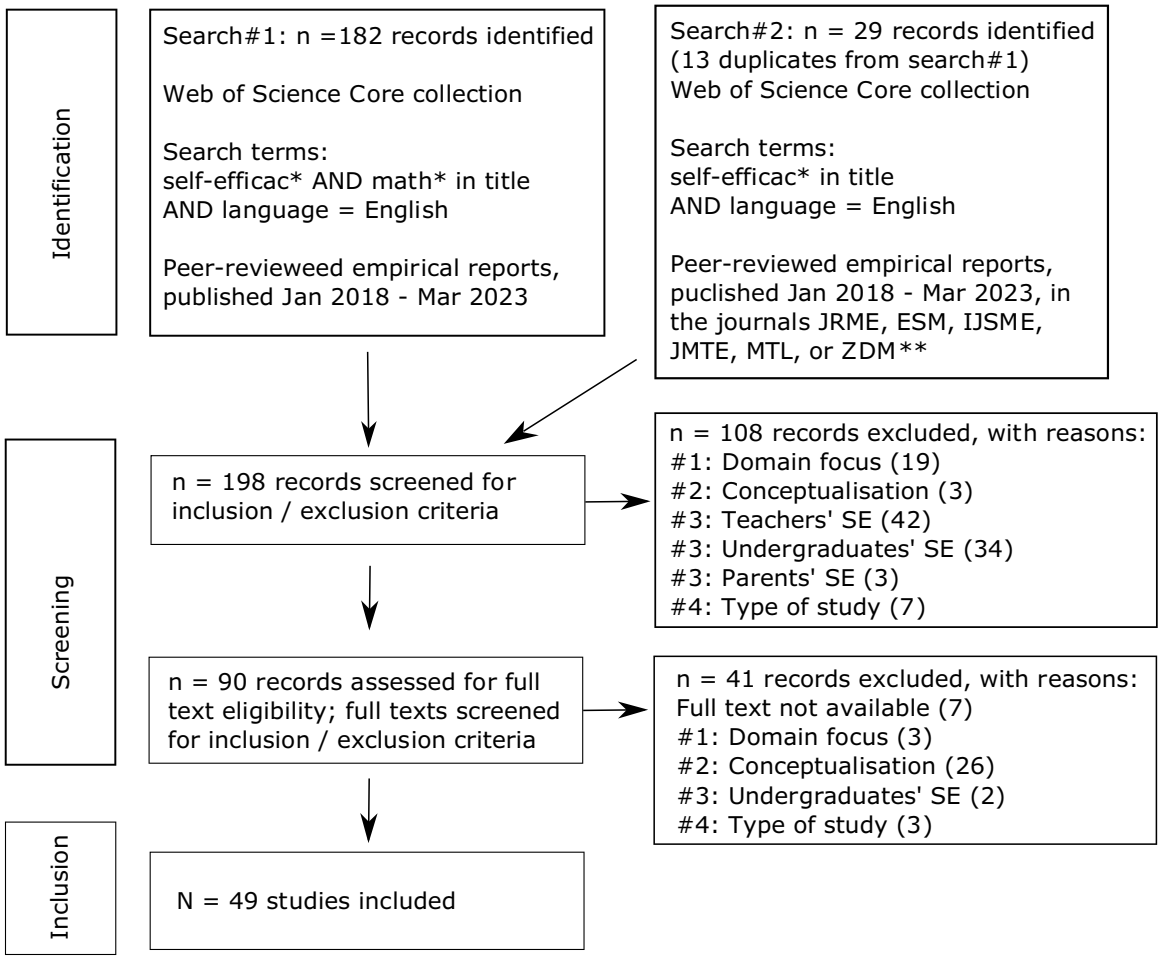

Fig. 2 Article selection process

anxiety (Du et al., 2021), or mathematical creative thinking (Rahyuningsih et al., 2022). Studies also investigated SE as a mediator of the relationship between other variables (31%), such as the relationship between teacher-student relationships and students' problem solving ability (Zhou et al., 2020), the relationship between implicit theories of intelligence and mathematics career interest (Huang et al., 2019), or the effects of perceived parental support on students' engagement in mathematics (Sağkal & Sönmez, 2022). Only two studies investigated the role of SE as a moderator on the relationship between other variables. Schukajlow et al. (2019) investigated whether prior SE moderated the effect of developing multiple solutions on students' MSE and perceived competence, while Cai et al. (2019) investigated whether there were differential effects of the use of tablets on students' conceptions and approaches to learning mathematics according to their SE.

Furthermore, studies investigated SE as an outcome variable (35%), such as the role of cognitive activation on students' MSE (Liu et al., 2022), how sources of SE raised or lowered students' MSE in a rural, high poverty, area (Usher et al., 2019), or the role of perfectionism on how students perceive information from sources of SE (Ford et al., 2023).

Six studies (12%) investigated reciprocal effects between SE and other variables. Five of these studies included a focus on performance, where, e.g., Arens et al. (2022) compared the constructs of SE and self-concept, in terms of their reciprocal relationships with performance measures (achievement test and school grades), Grigg et al. (2018) investigated reciprocal relationships between math interest, intentions, SE, and performance, and Du et al. (2021) investigated the reciprocal relationships between students' mathematics interest, anxiety, SE, and achievement.

Table 2 Results

Category	Code	Count	Percentage
Country (participants)			
(tick all that apply)	China	11	22.45
	USA	7	14.29
	Germany	6	12.24
	Turkey	4	8.16
	OECD	3	6.12
	Greece	2	4.08
	The Netherlands	2	4.08
	Norway	2	4.08
	Australia	1	2.04
	Costa Rica	1	2.04
	Estonia	1	2.04
	Hong Kong	1	2.04
	Indonesia	1	2.04
	Iran	1	2.04
	Israel	1	2.04
	New Zealand	1	2.04
	Saudi Arabia	1	2.04
	Singapore	1	2.04
	South Korea	1	2.04
	Spain	1	2.04
	Switzerland	1	2.04
	UK	1	2.04
Research methods			
(tick one)	Quantitative	43	87.76
	Qualitative	2	4.08
	Mixed methods	4	8.16
Participants			
(tick all that apply)	Elementary students	9	18.37
	Secondary students	42	85.71
Sample size			
(tick one)	0−50	4	8.16
	51−100	4	8.16
	101−200	4	8.16
	201−500	10	20.41
	501−1000	8	16.33
	> 1000	19	38.78
Design			
(tick all that apply)	Crossectional	33	67.35
	Longitudinal (= 2)	10	20.41
	Longitudinal (> 2)	3	6.12
	Microlongitudinal	3	6.12
	Experiment	5	10.20
Level of data			
(tick one)	Case	3	6.12
	Individ/pooled	37	75.51
	Multilevel	9	18.37
Dimensions of self-efficacy			
(tick all that apply)	Strength	49	100.00
	Specificity	5	10.20
	Difficulty	5	10.20

Table 2 (continued)

Category	Code	Count	Percentage
Treatment of self-efficacy			
(tick all that apply)	Correlate	12	24.49
	Predictor	20	40.82
	Outcome	17	34.69
	Reciprocal effects	6	12.24
	Mediator	15	30.61
	Moderator	2	4.08
Congruence[a]			
(applicable in 31 studies)	Congruence (of measures)	18	58.06
	Incongruence (of measures)	13	41.94
	not applicable	18	
Order[a]			
(applicable in 31 studies)	Congruence (of order)	23	74.19
	Incongruence (of order)	4	12.90
	Not reported	4	12.90
	Not applicable	18	
Other main variables			
(tick all that apply)	Student gender	16	32.65
	Collective self-efficacy	0	0.00
	Sources of self-efficacy	8	16.33
	Grades	6	12.24
	Test performance	31	63.27
	Anxiety	11	22.45
	(Other) affective/motivational variables	33	67.35
	Student behaviour	3	6.12
	Teacher variables	1	2.04
	(Other) contextual variables	9	18.37
Data sources			
(tick all that apply)	Performance test	27	55.10
	Questionnaire	46	93.88
	Reports/worksheets	0	0.00
	Interviews	3	6.12
	Observations	1	2.04
	Secondary data	14	28.57
Self-efficacy topic			
(tick all that apply)	Geometry	5	10.20
	Algebra	2	4.08
	Arithmetic	3	6.12
	Combinatorics/probability	2	4.08
	Fractions	0	0.00
	Functions	4	8.16
	Mixture of topics	14	28.57
	Other math object	5	10.20
	Domain self-efficacy (mathematics)	27	55.10

[a]These categories are only coded if relevant

24% of the studies treated SE as a correlate, including, e.g., the association between mathematical literacy and visual math literacy SE (Katranci & Şengül, 2019), and the relationship between parents' and children's MSE and emotional arousal to mathematics (Bartley & Ingram, 2018).

A range of variables were included as key foci in the studies assessed. Many studies (67%) focused on some type of

affective variable, including interest (Grigg et al., 2018), self-regulation (Tian et al., 2018), and emotional arousal (Bartley & Ingram, 2018), not counting anxiety. As the number of studies including a focus on anxiety turned out to be large, we coded this affective variable separately, and found that 22% of studies included a focus on this (e.g., Du et al., 2021). As could be expected, many studies included a focus on performance, through either test performance (63%), assigned grades (12%), or both, (e.g., Arens et al., 2022). Furthermore, many studies included a focus on gender (33%, e.g., Huang et al., 2019), while a smaller number of studies included a focus on sources of SE (16%, e.g., Usher et al., 2019). Other foci in the studies included student behaviour variables (three studies = 6%), e.g., student response times (Soland, 2019), teacher variables (one study), i.e., teacher behaviour, SE, and knowledge (Kaskens et al., 2020), and other contextual variables (18%), such as the role of the educational center (Zamora-Araya et al., 2022). No studies emerged that focused on collective SE, although the search procedure as well as inclusion–exclusion criteria would have enabled this.

4.1.3 Change in SE over time

A majority of the studies (67%) were crossectional in design, where data came from a single time of measurement, e.g., using data from PISA (Borgonovi & Pokropek, 2019). A total of 33% of studies had some type of longitudinal design: 20% of the studies included two measurement occasions (e.g., Rakoczy et al., 2019), while six studies (12%) included three or more measurement occasions (e.g., Street et al., 2022). 13 studies (27%) included long gaps in between each measurement occasion (i.e., weeks, months, or years apart), focusing, e.g., on the effect of a three-week intervention (Rakoczy et al., 2019) or the reciprocal relationship between SE and performance across the space of three years (Arens et al., 2022). Only three studies (6%) were micro-longitudinal in design, that is, measuring SE across learning situations (i.e., minutes or days apart), investigating e.g., the starting point and change of students' SE across a sequence of lessons in mathematics (Street et al., 2022). This micro-longitudinal code includes one case study where the measurement points were not specified in time, i.e., Leong (2021), who investigated how the SE of a single year-nine student in Singapore changed from moment to moment and across different tasks during a lesson in mathematics.

Five of the longitudinal studies applied an experimental design (10% of all studies). Examples are the impact of instruction based on dynamic versus static visualization on students' SE and problem solving in real-time as well as three months later (Kohen et al., 2022), and the effect of prompting students to construct multiple solutions on their perceived competence and SE (Schukajlow et al., 2019). All other studies were observational (90%).

4.1.4 Situational specificity of SE

A majority of the studies (82%) focused on a single level of analysis (including case studies), while 18% of the studies investigated also effects of nestedness (e.g., the effects of variations across time-points, teachers, schools, or districts). Examples are the effects of teacher- (i.e., teacher behaviour, SE, and knowledge) and student-level (i.e., MSE, self-concept, and anxiety) variables on students' arithmetic fluency and problem-solving (Kaskens et al., 2020), and the moderating effect of the educational center on the relationship between students' MSE and performance in Costa Rica (Zamora-Araya et al., 2022). Importantly, we focused in our coding manual on whether studies included a *substantial* focus of, e.g., school effects, on students' MSE. Several additional studies accounted for the nestedness of their data by applying multilevel analytical techniques but did not report the estimates for the group-level variation (as this was not a focus or aim of the study). We did not code whether studies *should* have accounted for nestedness (but failed to do so), e.g., if students were grouped within different classrooms but the shared variance of the classroom was not accounted for.

4.2 Methodological approaches of studies

4.2.1 Participants

Many different countries (21) were represented in the studies, in terms of the country of the participants. The largest number of studies (22%) included Chinese students, followed by USA (14%), Germany (12%), and Turkey (8%). One study included participants from four countries (Pepper et al., 2018), and three studies included data across (unspecified) OECD countries (e.g., Borgonovi & Pokropek, 2019).

Most of the studies included secondary school participants (86%), while 18% of studies included participants from elementary levels of education (two studies included both). Of the studies that included elementary students, one study included participants from 3rd grade (Lau et al., 2018), while the other studies included participants from grade four and up.

Many studies (39%) included a large number of participants, i.e., more than 1,000 students. Among these, four studies included between 158,000 and 605,000 students. These numbers were largely enabled through large-scale assessment data such as PISA or national repositories. One of these large-scale studies included data from Chinese students (Zhang & Wang, 2020) while the other three studies

included data from students in the OECD (e.g., Borgonovi & Pokropek, 2019).

4.2.2 Analytical methods

A large majority of the studies (88%) took a quantitative approach to investigating MSE, including quantitative (mainly questionnaire) data and using statistical methods for analyses (including Maximum Likelihood and Bayesian approaches). Four studies (8%) collected both quantitative and qualitative data (i.e., mixed methods), where, e.g., Gao (2020) used a survey followed by Q-sorting procedure to untangle the differential effects of exposure and importance of different sources of MSE. Two studies included qualitative data alone (4%). These included using cognitive interviews to validate the PISA measure of SE (Pepper et al., 2018) and investigating switches in SE for a single student across a lesson in mathematics (Leong, 2021).

4.2.3 Data sources

All but three studies included questionnaires (94%) as a source of data, and more than half of the studies (55%) included a type of performance measure. In addition, many studies (29%) included secondary data—included here were all the very-large-scale studies mentioned above. In addition to these methods, three studies included interviews (6%) and one single study included observation data (Leong, 2021). Usher et al. (2019) drew on multiple data sources to demonstrate both general trends (structured questionnaire data to investigate sources of SE) and details of these trends (open-ended questions to investigate the importance of the sources).

4.2.4 Congruence of measures

For the studies that investigated the relationship between SE and a domain of functioning or performance variable (31 studies = 63%), we investigated the congruence or alignment between the SE measure and the task/performance measure, focusing on whether the measures corresponded in terms of content and degree of specificity. Slightly more than half of the studies were deemed congruent in this regard (58%). In Borgonovi and Pokropek (2019) students were asked about their exposure to specific types of tasks, e.g., "Solving an equation like $3x + 5 = 17$.", and for SE "How confident do you feel about having to do the following mathematics tasks … $3x + 5 = 17$". This illustrates an exact match between the exposure to a certain type of task and the SE measure of that type of task.

Similarly, we looked at the order of measurement, and whether this was aligned with the treatment of SE in the study (whether SE was treated as a predictor or an outcome). 74% of the relevant studies were aligned with theoretical recommendations in this regard (e.g., performance measured prior to SE in a study focusing on SE as an outcome), while 13% were mis-aligned. Four studies (13%) did not provide sufficient detail to determine the order of measurement.

5 Discussion

In order to review the state of SE research in mathematics education, we screened what *substantive foci* have been posed, and which *methodological approaches* have been used in recent (2018–2022) studies of MSE. Herein we identify and analyze knowledge gaps, which could serve as basis for systematic reviews or meta-analyses, and propose future directions on research in MSE.

5.1 Substantive foci

In terms of substantive foci, the reviewed studies included a clear focus on the domain of mathematics. Some studies also addressed a specific topic of mathematics (e.g., geometry or arithmetic), a specific competency (e.g., modelling), or other objects (e.g., programming). Because of the small number of studies on SE regarding specific objects, our knowledge of, for example, geometry SE or problem-solving SE remains limited. Very few studies compared SE for different objects (e.g., SE for solving modelling problems on linear functions and on the Pythagorean theorem in Krawitz & Schukajlow, 2018). Thus, we call for more research that focus on such comparisons in the future, to get deeper insights into the mathematics-specific sources of SE.

As most studies investigated strength of SE, few studies investigated the role of perceived difficulty on students' SE (e.g., differences in the starting point and change of students' SE as a function of the perceived level of difficulty of the object in Street et al., 2022). Future studies should expand our understanding of the role of MSE as a multidimensional construct, through investigating the role of both objective and subjective task difficulty on students' SE change and development.

We found no studies investigating students' collective MSE, which may indicate this type of research is conducted with a focus on teachers, rather than on students. Future studies should contribute to theory development by collecting evidence about students' collective MSE.

Students' (individual) MSE was defined as predictor, outcome, both predictor and outcome in reciprocal effects, *and* as mediator and moderator between other variables. While there has been a long-standing debate regarding the relative effects of SE and performance experiences in their theoretically proposed reciprocal relationship (see, e.g.,

Arens et al., 2022), SE is now often included as a learning outcome, indicating that the importance of students' SE is widely accepted, beyond its' influence on performance. Although Bandura (1997) defined physiological states as a source of SE we found no studies which incorporated biophysiological states as predictors. Future studies could incorporate measures of, e.g., heartrate or electrodermal activity as indices of stress or arousal, in additional to self-reported anxiety and emotions (Martin et al., 2023). Many studies included SE as a mediator, in order to analyze how effects of predictors (e.g., emotions or strategies) on learning outcomes (e.g., performance) can be transmitted via students' SE. In the future, we call for more studies in which SE is included as a moderator, where the effects of predictors on learning outcomes can differ according to students' SE, as shown by Cai et al. (2019). Thus, including SE as a moderator can bring additional evidence about the role of SE in mathematics learning.

While many studies included a focus on students' mathematics performance, a larger proportion of studies included a focus on affective/motivational variables, as per the affective turn in mathematics education research described by Schukajlow et al. (2023). Consistent with Bandura's theory on sources of SE, emotions being a key potential source, the associations between SE, affective, and motivational variables could be suitable for future meta-analyses. Given the prevalence of anxiety among the emotional correlates of SE, the *type* of emotional construct could be considered as a moderator of MSE in future studies. Other suitable moderators for meta-analyses, and constructs that should be addressed in future studies, are gender, sources of SE, and students' appraisals or interpretations of SE sources.

Longer term longitudinal studies enabled investigations into broader changes in SE over time, while a smaller number of micro-longitudinal studies focused on the nature and mechanisms of changes of SE in mathematics classrooms. Qualitative data and analytical approaches enabled investigations into substantive foci such as different students' interpretations and weighting of different sources of information, and shifts in perspective across events. Expanding our knowledge on intraindividual changes in students' MSE within and across lessons in mathematics, e.g., through applying mixed methods and within (micro) longitudinal designs, would be valuable in future studies. Given the importance of development of SE over time for theoretical and practical implications, we also call for more intervention studies, which allow collecting indications on theoretically proposed mechanisms of change and sources of SE within ecological classroom settings.

Multilevel studies enabled a focus on the impact of contextual factors (e.g., migrant background or school characteristics) on students' SE, while only few studies focused on student behaviour or teacher variables. Furthermore, while there is a growing body of research demonstrating the crucial role of teacher-student interactions on students' development (Vandenbroucke et al., 2018), only one study in our review included classroom observation data enabling such a focus. Investigating the role of teacher-student interactions on changes in students' SE beliefs is a promising avenue for future research. Future studies on MSE could focus on the role of teachers' involvement with students, their emotional and instructional support to students, and the quality of teacher-student interactions.

5.2 Methodological approaches

In terms of methodological approaches, studies using quantitative data and analytical approaches dominated in our review, along with surveys and questionnaires as data collection tools. This mirrors findings of a previous review of research on SE in education more broadly (Klassen & Usher, 2010). As qualitative approaches tend to dominate within mathematics education research (Schukajlow et al., 2018), this might indicate that the bulk of research on MSE is being conducted by researchers outside the field of mathematics education. The relatively few mixed method and qualitative inquiries included demonstrated the value of methodological diversity in terms of expanding our understanding of the complex interplay between students' SE and their appraisals of their experiences and wider contextual factors. Large-scale surveys are not well suited to capture, e.g., the process of SE change including situational, personal, temporal, and social conditions together with students' individual cognitive appraisals (Usher & Pajares, 2008, p. 784). The study by Usher et al. (2019) is one example of how mixed methods and qualitative in-depth recording of students' experiences can be used to fruitfully investigate the complex and unstable construct of MSE. Thus, we encourage more mixed methods and qualitative research, and we encourage mathematics education researchers to (keep) contributing to research on SE as it pertains to mathematics, to ensure the inclusion of mathematics education perspectives when developing knowledge on this important construct.

The inclusion of MSE in international databases (i.e., PISA) enables broad generalization, while in-depth qualitative studies enable minute situation-specificity. Sample sizes in our review ranged from 1 to more than 600,000, with many relatively larger (> 1000 participants) studies. This demonstrates a distinct increase in participants from a previous review (Klassen & Usher, 2010). Large-scale cross-sectional and longitudinal studies in hierarchically nested designs (i.e., students in classrooms) in a range of countries (e.g., Spain, China, Germany) enable us to draw conclusions at both individual, classroom and school levels. The largest single contributing country was China, followed

by the US, reflecting a shift from a previous "western focus" in education research (Klassen & Usher, 2010). Maintaining diversity is important in terms of testing cross-cultural applicability of theories and deriving knowledge pertaining to mathematics classroom practice across the world.

More studies in our review included secondary than elementary students, and no participants younger than grade three were included, possibly mirroring the dominance of questionnaires as data source among the studies in our review (as well as the methodological and ethical challenges associated with research on younger children). We concur with Schukajlow et al. (2023) who suggested there is a need for studies that include children, in particular elementary age. One example is the need to untangle the effects of age and domain-specificity on the SE- performance relationship.

A prevailing concern in SE research, is the issue of operationalization of the construct. While Klassen and Usher (2010) concluded that more than half of the studies in their review suffered from shortcomings in measurement, we excluded slightly fewer of the identified studies from our search due to issues with conceptualization and operationalization of SE. Furthermore, many of the included studies in our review were coded as lacking in congruence of measures, meaning there was a mismatch between the measure of SE and the domain of functioning. Also, while there is a convergence of measurement (e.g., item-selection) particularly in large-scale international comparisons, we need to continue to be aware of cultural and sub-group differences in respondents' understanding of items (Pepper et al., 2018). We reiterate recommendations by, e.g., Klassen and Usher (2010) and Diego-Mantecón et al. (2019), that researchers pay close attention to available guidelines for theoretically based approaches to conceptualization and measurement of the construct of interest.

Through our work with this review, we identified many studies of potential interest to researchers within mathematics education. We present a selection of these within our references (marked **), as illustrations of what we consider the state-of-the-art in current research on MSE. The special consideration articles highlight in different ways either best-practice contributions or innovative approaches to bring the field forward, in substantive and/or methodological terms. We summarise briefly: Many studies have been concerned with investigating reciprocal effects between SE and performance, and Arens et al. (2022) illustrated a best-practice approach to inform this long-standing issue. While most studies investigated MSE at the domain level or for a mixture of topics, Krawitz and Schukajlow (2018) demonstrated the value of focusing on students' MSE for specific topics and types of problems and comparing these in the same study. One trend in our review was the many studies with a large number of participants, enabled through large-scale assessment data such as PISA. Soland (2019) found that item response times on very difficult items in PISA was strongly correlated with students' SE, demonstrating a promising avenue for measuring students' SE more directly. A key issue for the future is to develop new methodologies to capture change in a dynamic construct that interacts with situational and contextual factors. Two studies (Gao, 2020; Usher et al., 2019) illustrated innovative approaches to untangling students' differential appraisals of SE sources, and Leong (2021) took a case-study approach to investigating changes in MSE from moment to moment and across different tasks during a single lesson in mathematics.

5.3 Limitations

There are limitations to our study. One is our search strategy, i.e., the sampling of the study. We searched for papers that included our search terms in the title only, meaning we knowingly did not target the large number of studies that focus on students' MSE, but did not include the terms in their title. We included only research conducted in English, meaning research conducted in non-Western societies may not be appropriately represented.

6 Conclusion

Many trends uncovered in our review are promising, in terms of a large variety of thematic foci, coupled with novel (e.g., questionnaire followed by q-sorting procedure) and advanced (e.g., multilevel structural equation models) methodological approaches enabling new insights. As methodological advancement is rapid, new methods keep giving scope for phrasing new research questions. In the future we are likely to see further methodological-substantive synergies in the field of MSE. At the same time, we see the potential and need for future studies, such as to continue the focus on MSE as a multidimensional and dynamic concept, i.e., considering factors such as mathematical object specificity and (perceived) task demand, as well as designs that enable investigations into the nature and mechanisms of MSE changes over time. Qualitative or mixed-methods studies from the field of mathematics education focusing on, e.g., unpacking student appraisals of their experiences with mathematics, could constitute major future contributions to our understanding of students' MSE. Careful consideration of the theoretical background of the construct of MSE in relation to the aims of each individual study continues to be important in order to bring the field forward.

Author contributions All authors contributed to the study conception and design. Coding was performed by KESS and L-EM. The first draft of the manuscript was written by KESS and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open access funding provided by Western Norway University Of Applied Sciences. No funding was received for conducting this study, and the authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

References that were part of the review are marked (*). Papers of particular interest, published within the period of the review, have been highlighted (**).

Ahn, H. S., Usher, E. L., Butz, A. R., & Bong, M. (2016). Cultural differences in the understanding of modelling and feedback as sources of self-efficacy information. *British Journal of Educational Psychology, 86*(1), 112–136. https://doi.org/10.1111/bjep.12093

**Arens, A. K., Frenzel, A. C., & Goetz, T. (2022). Self-concept and self-efficacy in math: Longitudinal interrelations and reciprocal linkages with achievement. *Journal of Experimental Education, 90*(3), 1–19. https://doi.org/10.1080/00220973.2020.1786347. The authors investigated reciprocal effects between MSE and performance and grades in two separate models, including also the construct of self-concept. The dominant direction of effects was from test performance to SE and from self-concept to grades. Self-concept predicted SE, but not the other way around. The relationships between SE /self-concept and achievement did not vary according to student gender or age.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman and Co.

*Bartley, S. R., & Ingram, N. (2018). Parental modelling of mathematical affect: Self-efficacy and emotional arousal. *Mathematics Education Research Journal, 30*(3), 277–297. https://doi.org/10.1007/s13394-017-0233-3

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*, 1–40. https://doi.org/10.1023/A:1021302408382.

*Borgonovi, F., & Pokropek, A. (2019). Seeing is believing: Task-exposure specificity and the development of mathematics self-efficacy evaluations. *Journal of Educational Psychology, 111*(2), 268–283. https://doi.org/10.1037/edu0000280

Byars-Winston, A., Diestelmann, J., Savoy, J. N., & Hoyt, W. T. (2017). Unique effects and moderators of effects of sources on self-efficacy: A model-based meta-analysis. *Journal of Counseling Psychology, 64*(6), 645–658. https://doi.org/10.1037/cou0000219

*Cai, S., Liu, E., Yang, Y., & Liang, J. C. (2019). Tablet-based AR technology: Impacts on students' conceptions and approaches to learning mathematics according to their self-efficacy. *British Journal of Educational Technology, 50*(1), 248–263. https://doi.org/10.1111/bjet.12718

Diego-Mantecón, J. M., Blanco, T. F., Chamoso, J. M., & Cáceres, M. J. (2019). An attempt to identify the issues underlying the lack of consistent conceptualisations in the field of student mathematics-related beliefs. *PLoS ONE, 14*(11), 1–19. https://doi.org/10.1371/journal.pone.0224696

Dowker, A., Sarkar, A., & Looi, C. Y. (2016). Mathematics anxiety: What have we learned in 60 years? *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2016.00508

*Du, C., Qin, K., Wang, Y., & Xin, T. (2021). Mathematics interest, anxiety, self-efficacy and achievement: Examining reciprocal relations. *Learning and Individual Differences, 91*(19), 1–8. https://doi.org/10.1016/j.lindif.2021.102060

Dweck, C. S. (2008). *Self-theories: Their role in motivation, personality and development*. Academic Press.https://doi.org/10.4324/9781315783048.

*Ford, C. J., Usher, E. L., Scott, V. L., & Chen, X. Y. (2023). The 'perfect' lens: Perfectionism and early adolescents' math self-efficacy development. *British Journal of Educational Psychology, 93*(1), 211–228. https://doi.org/10.1111/bjep.12550

**Gao, J. (2020). Sources of mathematics self-efficacy in Chinese students: A mixed-method study with Q-sorting procedure. *International Journal of Science and Mathematics Education, 18*(4), 713–732. https://doi.org/10.1007/s10763-019-09984-1. Gao used a survey followed by Q-sorting procedure to untangle the differential effects of exposure and importance of sources of MSE. The study demonstrated differential experiences and appraisals of sources of SE according to students' SE strength and gender and highlighted individual variations. Furthermore, it demonstrated how the indigenous Chinese concept "face" (i.e., the need to sustain dignity and self-esteem) might influence Chinese students' MSE.

Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., Oliver, P. H., & Guerin, D. W. (2007). Multivariate latent change modeling of developmental decline in academic intrinsic math motivation and achievement: Childhood through adolescence. *International Journal of Behavioral Development, 31*(4), 317–327. https://doi.org/10.1177/0165025407077752

*Grigg, S., Perera, H. N., McIlveen, P., & Svetleff, Z. (2018). Relations among math self efficacy, interest, intentions, and achievement: A social cognitive perspective. *Contemporary Educational Psychology, 53*(January), 73–86. https://doi.org/10.1016/j.cedpsych.2018.01.007

*Hiller, S. E., Kitsantas, A., Cheema, J. E., & Poulou, M. (2022). Mathematics anxiety and self-efficacy as predictors of mathematics literacy. *International Journal of Mathematical Education in Science and Technology, 53*(8), 2133–2151. https://doi.org/10.1080/0020739X.2020.1868589

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *In Educational Research Review, 17*, 63–84. https://doi.org/10.1016/j.edurev.2015.11.002

Hox, J., Moerbeek, M., & van de Schoot, R. (2017). *Multilevel analyses* (3rd ed.). Routledge. https://doi.org/10.4324/9781315650982.

*Huang, X., Zhang, J., & Hudson, L. (2019). Impact of math self-efficacy, math anxiety, and growth mindset on math and science career interest for middle school students: The gender moderating effect. *European Journal of Psychology of Education, 34*(3), 621–640. https://doi.org/10.1007/s10212-018-0403-z

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: gender and domain differences across grades one through twelve. *Child Development, 73*(2), 509–527. https://doi.org/10.1111/1467-8624.00421.

*Jiang, H., Turnbull, D., Wang, X., Chugh, R., Dou, Y., & Chen, S. (2022). How do mathematics interest and self-efficacy influence coding interest and self-efficacy? A structural equation modeling analysis. *International Journal of Educational Research, 115*(May), 102058. https://doi.org/10.1016/j.ijer.2022.102058

*Kaskens, J., Segers, E., Goei, S. L., van Luit, J. E. H., & Verhoeven, L. (2020). Impact of Children's math self-concept, math self-efficacy, math anxiety, and teacher competencies on math development. *Teaching and Teacher Education, 94*, 103096. https://doi.org/10.1016/j.tate.2020.103096

*Katranci, Y., & Şengül, S. (2019). The relationship between mathematical literacy and visual math literacy self-efficacy perceptions of middle school students. *Pegem Egitim Ve Ogretim Dergisi, 9*(4), 1113–1138. https://doi.org/10.14527/pegegog.2019.036

Klassen, R. M. (2004). A cross-cultural investigation of the efficacy beliefs of South Asian immigrant and Anglo Canadian nonimmigrant early adolescents. *Journal of Educational Psychology, 96*(4), 731–742. https://doi.org/10.1037/0022-0663.96.4.731

Klassen, R. M., & Usher, E. L. (2010). Self-efficacy in educational settings: Recent research and emerging directions. In T. C. Urdan & S. A. Karabenick (Eds.), *Advances in motivation and achievement* (Vol. 16, pp. 1–33). Emerald Group Publishing. https://doi.org/10.1108/S0749-7423(2010)000016A004

Klee, H. L., Miller, A. D., & Buehl, M. M. (2022). mathematics anxiety, self-concept, and self-efficacy: A multidimensional scaling consideration of measures. *Journal of Experimental Education*. https://doi.org/10.1080/00220973.2021.2024788

*Kohen, Z., Amram, M., Dagan, M., & Miranda, T. (2022). Self-efficacy and problem-solving skills in mathematics: The effect of instruction-based dynamic versus static visualization. *Interactive Learning Environments, 30*(4), 759–778. https://doi.org/10.1080/10494820.2019.1683588

**Krawitz, J., & Schukajlow, S. (2018). Do students value modelling problems, and are they confident they can solve such problems? Value and self-efficacy for modelling, word, and intra-mathematical problems. *ZDM, 50*(1–2), 143–157. https://doi.org/10.1007/s11858-017-0893-1. The authors analyzed German lower-secondary students' MSE for different topics (geometry vs algebra) and type of problems (modelling vs. word vs intra-mathematical problems) by using task-specific questionnaires. The analysis indicated lower MSE for modelling problems than for other types of problems. Students' MSE did not differ for different mathematical topics. The results underline the importance of interventions that increases MSE for solving modelling problems.

*Lau, C., Kitsantas, A., Miller, A. D., & Drogin Rodgers, E. B. (2018). Perceived responsibility for learning, self-efficacy, and sources of self-efficacy in mathematics: A study of international baccalaureate primary years programme students. *Social Psychology of Education, 21*(3), 603–620. https://doi.org/10.1007/s11218-018-9431-4

**Leong, Y. H. (2021). Contours of self-efficacy across nested mathematical domains: a case of a Singapore student with a history of low performance in mathematics. *Mathematics Education Research Journal*. https://doi.org/10.1007/s13394-021-00394-7. Leong investigated how the SE of a single year nine student in Singapore changed from moment to moment and across different tasks during a lesson in mathematics. The study demonstrated switches of SE between tasks and highlighted how lesson materials such as notes can potentially strengthen a student's MSE, as a reference point to both mastery experiences and social persuasion (here; teacher encouragement).

*Liu, Q., Liu, J., Cai, J., & Zhang, Z. (2020). The relationship between domain- and task-specific self-efficacy and mathematical problem posing: A large-scale study of eighth-grade students in China. *Educational Studies in Mathematics, 105*(3), 407–431. https://doi.org/10.1007/s10649-020-09977-w

*Liu, Y., Wang, C., Liu, J., & Liu, H. (2022). The role of cognitive activation in predicting mathematics self-efficacy and anxiety among internal migrant and local children. *Educational Psychology, 42*(1), 83–107. https://doi.org/10.1080/01443410.2021.1987388

Marsh, H. W., & Hau, K. T. (2007). Applications of latent-variable models in educational psychology: The need for methodological-substantive synergies. *Contemporary Educational Psychology, 32*(1), 151–170. https://doi.org/10.1016/j.cedpsych.2006.10.008

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of Lurking Jingle-Jangle Fallacies. *Journal of Educational Psychology, 111*(2), 331–353. https://doi.org/10.1037/edu0000281

Martin, A. J., Malmberg, L.-E., Pakarinen, E., Mason, L., & Mainhard, T. (2023). The potential of biophysiology for understanding motivation, engagement and learning experiences (editorial for special issue). *British Journal of Educational Psychology, 93*(1), 1–9. https://doi.org/10.1111/bjep.12584.

Martin, A. J., Way, J., Bobis, J., & Anderson, J. (2015). Exploring the ups and downs of mathematics engagement in the middle years of school. *Journal of Early Adolescence, 35*(2), 199–244. https://doi.org/10.1177/0272431614529365

McNeish, D., & Hamaker, E. L. (2020). A primer on two level dynamic structural equation models for intensice longitudinal data in mplus. *Psychological Methods, 25*(5), 610–635. https://doi.org/10.1037/met0000250

Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology, 106*(1), 121–131. https://doi.org/10.1037/a0033546

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement, 2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*. https://doi.org/10.4324/9781315159416

Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology, 24*(2), 124–139. https://doi.org/10.1006/ceps.1998.0991

Pajares, F., & Miller, M. D. (1995). Mathematics self-efficacy and mathematics performances: The need for specificity of assessment. *Journal of Counseling Psychology, 42*(2), 190–198. https://doi.org/10.1037/0022-0167.42.2.190.

Pampaka, M., Kleanthous, I., Hutcheson, G. D., & Wake, G. (2011). Measuring mathematics self-efficacy as a learning outcome. *Research in Mathematics Education, 13*(2), 169–190. https://doi.org/10.1080/14794802.2011.585828

Pantziara, M. (2016). Student self-efficacy beliefs. In G. A. Goldin, M. S. Hannula, E. Heyd-Metzuyanim, A. Jansen, R. Kaasila, S. Lutovac, P. Di Martino, F. Morselli, J. A. Middleton, M. Pantziara, & Q. Zhang (Eds.), *Attitudes, beliefs, motivation and identity in mathematics education. An overview of the field and future directions* (pp. 7–11). Springer. https://link.springer.com/content/pdf/10.1007/978-3-319-32811-9.pdf.

*Pepper, D., Hodgen, J., Lamesoo, K., Kõiv, P., & Tolboom, J. (2018). Think aloud: Using cognitive interviewing to validate the PISA assessment of student self-efficacy in mathematics. *International Journal of Research & Method in Education, 41*(1), 3–16. https://doi.org/10.1080/1743727X.2016.1238891

*Rahyuningsih, S., Nurhusain, M., & Indrawati, N. (2022). Mathematical creative thinking ability and self-efficacy: A mixed-methods study involving indonesian students. *Uniciencia, 36*(1), 1–14. https://doi.org/10.15359/ru.36-1.20

*Rakoczy, K., Pinger, P., Hochweber, J., Klieme, E., Schütze, B., & Besser, M. (2019). Formative assessment in mathematics: Mediated by feedback's perceived usefulness and students' self-efficacy. *Learning and Instruction*. https://doi.org/10.1016/j.learninstruc.2018.01.004

*Sağkal, A. S., & Sönmez, M. T. (2022). The effects of perceived parental math support on middle school students' math engagement: The serial multiple mediation of math self-efficacy and math enjoyment. *European Journal of Psychology of Education, 37*(2), 341–354. https://doi.org/10.1007/s10212-020-00518-w

*Schukajlow, S., Achmetli, K., & Rakoczy, K. (2019). Does constructing multiple solutions for real-world problems affect self-efficacy ? *Educational Studies in Mathematics, 100*, 43–60. https://doi.org/10.1007/s10649-018-9847-y.

Schukajlow, S., Kaiser, G., & Stillman, G. (2018). Empirical research on teaching and learning of mathematical modelling: A survey on the current state-of-the-art. *ZDM, 50*, 5–18. https://doi.org/10.1007/s11858-018-0933-5.

Schukajlow, S., Rakoczy, K., & Pekrun, R. (2023). Emotions and motivation in mathematics education: Where we are today and where we need to go. *ZDM, 55*(2), 249–267. https://doi.org/10.1007/s11858-022-01463-2

**Soland, J. (2019). Can item response times provide insight into students' motivation and self-efficacy in math? An initial application of test metadata to understand students' social-emotional needs. *Educational Measurement: Issues and Practice, 38*(3), 86–96. https://doi.org/10.1111/emip.12260. Soland included item-response times from PISA data to illustrate how test metadata can provide additional insights into students' motivation and SE. While there was only a moderate association between the length of time students spent on the test items and their SE overall, item response time was strongly correlated with students' SE in the case of very difficult items. That is, students with higher SE spent longer on difficult items, relative to students with lower SE.

Street, K. E. S., Malmberg, L.-E., & Stylianides, G. J. (2017). Level, strength, and facet-specific self-efficacy inmathematics test performance. *ZDM - Mathematics Education*, *49*(3), 379–395. https://doi.org/10.1007/s11858-017-0833-0.

*Street, K. E. S., Malmberg, L.-E., & Stylianides, G. J. (2022). Changes in students' self-efficacy when learning a new topic in mathematics: A micro-longitudinal study. *Educational Studies in Mathematics, 111*(3), 515–541. https://doi.org/10.1007/s10649-022-10165-1

Talsma, K., Schüz, B., Schwarzer, R., & Norris, K. (2018). I believe, therefore I achieve (and vice versa): A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Learning and Individual Differences, 61*(January), 136–150. https://doi.org/10.1016/j.lindif.2017.11.015

*Tian, Y., Fang, Y., & Li, J. (2018). The effect of metacognitive knowledge on mathematics performance in self-regulated learning framework-multiple mediation of self-efficacy and motivation. *Frontiers in Psychology, 9*(DEC), 1–11. https://doi.org/10.3389/fpsyg.2018.02518

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., et al. (2018). Preferred Reporting items for systematic reviews and meta-analyses extension for scoping reviews (PRISMA-ScR) checklist SECTION. *Annals of Internal Medicine, 169*(7), 11–12. https://doi.org/10.7326/M18-0850.2

**Usher, E. L., Ford, C. J., Li, C. R., & Weidner, B. L. (2019). Sources of math and science self-efficacy in rural Appalachia: A convergent mixed methods study. *Contemporary Educational Psychology, 57*(October 2018), 32–53. https://doi.org/10.1016/j.cedpsych.2018.10.003. The study focused on the effects of mastery experiences and physiological states on students' MSE in a rural region in the US. The authors drew on multiple data sources to demonstrate general trends (earlier mastery experiences raised while physiological states lowered later SE) and details of these trends (the importance of the sources differed in terms of, e.g., the effect on SE and student gender).

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*(4), 751–796. https://doi.org/10.3102/0034654308321456.

Vandenbroucke, L., Spilt, J., Verschueren, K., Piccinin, C., & Baeyens, D. (2018). The classroom as a developmental context for cognitive development: A meta-analysis on the importance of teacher–student interactions for children's executive functions. *Review of Educational Research, 88*(1), 125–164. https://doi.org/10.3102/0034654317743200

Weidinger, A. F., Steinmayr, R., & Spinath, B. (2018). Changes in the relation between competence beliefs and achievement in math across elementary school years. *Child Development, 89*(2), e138–e156. https://doi.org/10.1111/cdev.12806

*Zamora-Araya, J. A., Montero-Rojas, E., Smith-Castro, V., Moreira-Mora, T. E., Zamora-Calvo, P., Quintero-Arias, K., & Matarrita-Muñoz, S. (2022). Gender, self-efficacy and performance in a mathematics test: The moderating role of the educational center. *Uniciencia, 36*(1), 1–17. https://doi.org/10.15359/ru.36-1.46

*Zhang, D., & Wang, C. (2020). The relationship between mathematics interest and mathematics achievement: Mediating roles of self-efficacy and mathematics anxiety. *International Journal of Educational Research, 104*(May), 101648. https://doi.org/10.1016/j.ijer.2020.101648

*Zhou, D., Du, X., Hau, K.-T., Luo, H., Feng, P., & Liu, J. (2020). Teacher-student relationship and mathematical problem-solving ability: Mediating roles of self-efficacy and mathematical anxiety. *Educational Psychology, 40*(4), 473–489. https://doi.org/10.1080/01443410.2019.1696947

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*(1), 82–91. https://doi.org/10.1006/ceps.1999.1016

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.