"Yes, I Can!" A Systematic Review and Meta-Analysis of Intervention Studies Promoting Teacher Self-Efficacy

Janina Täschner (iD)
*Technical University of Munich and Centre for
International Student Assessment*

Theresa Dicke (iD)
Australian Catholic University

Sarah Reinhold (iD)
Doris Holzberger (iD)
*Technical University of Munich and Centre for
International Student Assessment*

A high level of teacher self-efficacy is considered to be important for a successful and healthy teaching career. This preregistered meta-analysis focuses on whether and to what degree interventions can promote teacher self-efficacy. We included 115 studies representing 11,284 pre-service and in-service teachers in our meta-analysis. Interventions had a significant, positive effect on the promotion of teachers' self-efficacy (g = 0.47, RVE SE = 0.04, 95% CI = [0.40, 0.54]) with no significant differences between pre- and in-service teachers. A fine-grained coding and systematic review of the targeted sources of self-efficacy according to Bandura's sociocognitive theory revealed that overall interventions including mastery experiences did not significantly differ from those without. However, interventions targeting only mastery experiences were the most successful for pre-service teachers (g = 0.62, RVE SE = 0.11, CI = [0.35, 0.88]). Based on further moderator analyses, we recommend interventions to integrate reflective elements. Finally, future research should apply stricter study designs and more detailed intervention descriptions.

KEYWORDS: self-efficacy, Bandura, meta-analysis, teacher education, intervention

In times of increasing teacher shortages, concerning rates of teacher attrition, and increasing demands put on teachers, many institutions and research organizations are questioning how teachers' lifelong learning experiences should be designed to best enable teachers to deal with these challenges (Ball & Forzani,

2009; European Commission/EACEA/Eurydice, 2021; Grossman, Hammerness, & McDonald, 2009; Schleicher, 2018). A central variable in the discourse on dealing with challenges is teachers' self-efficacy (TSE), which is defined as teachers' belief in their abilities to successfully perform teaching-specific tasks (e.g., motivating disinterested students or calming a disruptive student down; Skaalvik & Skaalvik, 2007; Tschannen-Moran et al., 1998). A wide range of learning opportunities have emerged in recent years to strengthen teachers' self-efficacy. There are specific intervention programs targeting teacher self-efficacy in pre- and in-service teacher education (e.g., Chao et al., 2017; Dalioglu & Adiguzel, 2016; Hayes et al., 2020; McCullough et al., 2021; Tschannen-Moran & McMaster, 2009) or more global approaches such as placing core teaching practices and opportunities for practice at the heart of pre-service teacher education (Grossman, Hammerness, & McDonald, 2009; Michigan Department of Education, n.d.). Although these two exemplary approaches may seem very different at first glance, they are quite related when examined more closely. Both offer pre-service and in-service teachers a specific combination of experiences.

Albert Bandura's (1997) socio-cognitive theory provides a suitable framework to describe these different experiences. He proposes that self-efficacy beliefs are based on four sources of experiences, namely, mastery experiences, vicarious experiences, social persuasion, and physiological reactions. Mastery experiences are seen as having the strongest impact on the level of self-efficacy beliefs. However, empirical evidence for the differential effect of the four sources of self-efficacy on the promotion of teacher self-efficacy is inconclusive (D. B. Morris et al., 2017). Given the relevance of high levels of teacher self-efficacy beliefs (Klassen & Tze, 2014; Zee & Koomen, 2016), it remains an open research question how interventions can support pre-service and in-service teachers to strengthen their self-efficacy (Klassen et al., 2011; Moulding et al., 2014; Pendergast et al., 2011; Tschannen-Moran & Woolfolk Hoy, 2007).

The present meta-analysis systematically summarizes current research on interventions[1] promoting teacher self-efficacy. We investigate which characteristics of the interventions are influencing the interventions' effects and focus specifically on a repeatedly raised research gap, namely, evaluating the role of the experiences captured as sources of self-efficacy (Duffin et al., 2012; Klassen et al., 2011). Within a teacher professional development framework, we evaluate interventions' effects on different samples of teachers (pre-service teachers and in-service teachers) and examine whether the two groups react differently to different source combinations. Our overall goal is to contribute evidence-based knowledge to the question of how lifelong learning opportunities in teacher education and training should be designed so that teachers can deal with future challenges in a self-efficacious way.

The Concept and Relevance of Teacher Self-Efficacy

The construct of teacher self-efficacy is based on Bandura's socio-cognitive theory, in which he defined self-efficacy as "the belief in one's capabilities to organize and execute the course of action required to manage prospective situations" (Bandura, 1995, p. 2). Teacher self-efficacy can, therefore, be defined as an individual teacher's belief to be able to execute all the actions

needed for the job profile of a teacher (for similar definitions, see Holzberger et al., 2013; Tschannen-Moran et al., 1998).

The construct has inspired many studies, meta-analyses, and systematic reviews showing benefits of a high level of teacher self-efficacy for teachers, their teaching quality, and their students (Zee & Koomen, 2016). Teachers with a higher level of teacher self-efficacy report more job satisfaction (Klassen & Chiu, 2010; Stephanou et al., 2013; Toropova et al., 2020), more commitment (Chesnut & Burley, 2015), less emotional exhaustion and fewer stress symptoms (Aloe et al., 2014; Betoret, 2006; Dicke et al., 2018; Fernet et al., 2012; Skaalvik & Skaalvik, 2010, 2014; Wang et al., 2015). Students and even independent observers rate the teaching of self-efficacious teachers as high-quality instruction (Holzberger et al., 2013; Klassen & Tze, 2014; Ryan et al., 2015). In a longitudinal study, teachers' level of teacher self-efficacy in the first year was predictive of their instructional quality regarding classroom climate, classroom management, and cognitive activation 10 years later (Künsting et al., 2016). Finally, regarding concrete teaching practices, teacher self-efficacy is considered an important predictor of differentiated instruction (Neve et al., 2015; Suprayogi et al., 2017) and innovative behavior (Thurlings et al., 2015). In turn, students from teachers with higher teacher self-efficacy beliefs report more engagement in the class (Zee & Koomen, 2020), better relationships with these teachers (Holzberger et al., 2014), and more interest (Fauth et al., 2019; mediated by teaching quality). Students also show higher academic achievement when taught by teachers with higher levels of teacher self-efficacy (Caprara et al., 2006; Kim & Seo, 2018).

In light of the described evidence on the beneficial outcomes of high levels of teacher self-efficacy, research has consequently started to investigate how teacher self-efficacy can be promoted (e.g., Ansley et al., 2021; Groeschner et al., 2018; Hoogendijk et al., 2018; Junqueira & Matoti, 2013; Ross & Bruce, 2007).

Promoting Teacher Self-Efficacy Through the Four Sources of Self-Efficacy

So far, a huge variety of interventions exist that aim to promote TSE. These interventions differ in the content they deliver (e.g., classroom management skills in Dicke, Elling, et al., 2015; English grammar and vocabulary in S. Chan et al., 2021), the involvement of core teaching practices (e.g., developing positive student-teacher relationships in Aasheim et al., 2020; creating lesson plans in Karalar & Altan, 2018), the ways they are organized (e.g., specific training intervention in Dicke, Elling, et al., 2015; teaching practicum in Holdaway, 2017; individual coaching in Hoogendijk et al., 2018), the activities they implement (e.g., role plays in Dicke, Elling, et al., 2015; video observations in Gold et al., 2017), and the degree of authentic approximations of practice (e.g., interacting with avatars in Bosch & Ellis, 2021; real-time coaching in classroom in Hoogendijk et al., 2018).

Despite the wide variety of intervention studies aimed at promoting teacher self-efficacy beliefs, the four sources of self-efficacy (i.e., mastery experiences, vicarious experiences, social persuasion, and physiological reactions) described by Bandura (1997) may serve as a feasible and meaningful overall framework for categorizing the diverse interventions. While these sources were originally described outside a specific context, they can be transferred to

the teaching context (D. B. Morris et al., 2017; Tschannen-Moran et al., 1998). We therefore carefully reviewed intervention studies on promoting teacher self-efficacy and provide examples of the operationalization of the four sources in teacher education and training in Table 1. Despite their different contents, organizational structures, involved practices, and approximation levels, interventions can offer the same sources of self-efficacy. For example, an intervention specifically fostering classroom management skills (Dicke, Elling, et al., 2015), a student teaching practicum (Holdaway, 2017), and an intervention focusing on different ways of coaching pre-service teachers (Weber et al., 2019) all target the same sources of self-efficacy, namely, mastery experiences, vicarious experiences, and social persuasion. Through coding the different sources of self-efficacy targeted by the interventions, we are able to compare seemingly diverse interventions and to decompose successful strategies for fostering teacher self-efficacy. In the following, we introduce the four sources of self-efficacy according to Bandura's theory. For each source, we highlight examples from teacher education and training interventions based on our systematic review of relevant studies (see also Table 1).

Mastery experiences take place whenever someone gains operational experience by doing something.[2] In teacher education, mastery experiences may be teaching experiences in internships for pre-service teachers (e.g., Knoblauch & Chase, 2015; Liaw, 2017) or role plays in seminars for in-service teachers (e.g., Aasheim et al., 2020; Dicke, Elling, et al., 2015), for example. Individuals have *vicarious experiences* whenever they observe a model doing something, for example, pre-service teachers watching classroom videos (e.g., Bowlin et al., 2015; Kumschick et al., 2017; Thiel et al., 2020) or in-service teachers observing their colleagues or experts (e.g., Schipper et al., 2020; Tschannen-Moran & McMaster, 2009; Vansteelandt et al., 2020). *Social persuasion* covers all the situations where individuals receive verbal support and encouragement for a specific activity. In teacher education, social persuasion may occur when pre-service and in-service teachers receive feedback from their peers or supervisors or are coached by experts (e.g., Hoogendijk et al., 2018; Weber et al., 2019). The fourth source of self-efficacy pertains to *physiological and emotional reactions* (sometimes also called "physiological and emotional states"). People obtain information about whether they feel capable in certain situations by interpreting their affective reactions, such as an increased heart rate as a reflection of fear versus excitement. For example, when teachers interpret their palpitations as excitement at teaching a new topic, they feel capable and enter the classroom with more energy and enthusiasm. Emotional exhaustion or higher levels of negative emotions also predicted (changes in) teacher self-efficacy (Burić et al., 2020; Dicke, Parker, et al., 2015). In teacher education, meditation or autogenic training workshops, for example, address this source of self-efficacy (e.g., Ansley et al., 2021; de Carvalho et al., 2021).

The Central Role of Mastery Experiences as a Source to Promote Teacher Self-Efficacy

Bandura (1997) claimed that mastery experiences are "the most effective way of creating a strong sense of efficacy" (p. 3). Empirical studies on mastery experiences (e.g., teaching internships or microteaching) mostly found increases in

TABLE 1

Examples of Operationalization of Sources of Self-Efficacy in Intervention Studies in Teacher Education and Training

Source	Operationalization
Mastery experiences	• Role plays (Aasheim et al., 2020; Dicke, Elling, et al., 2015; Kelleci et al., 2018; Kurt, 2017) • Teaching internships or teaching practicums (Atay, 2007; Berg & Smith, 2018; Junqueira & Matoti, 2013; Putman, 2012a; Weber et al., 2019) • Teaching mini-lessons to peers (Yough, 2019) • Micro-teaching (Karalar & Altan, 2018; Sheehan & Moore, 2019; Wagler & Moseley, 2005)
Vicarious experiences	• Videotaped teaching models (van der Scheer & Visscher, 2016) • Observing peers or experienced teachers teaching (Cabaroglu, 2014; Grammatikopoulos et al., 2013; Wagler & Moseley, 2005) • Symbolic models (vignettes) (Garwood & Harris, 2020; Marlow et al., 2015; Thiel et al., 2020)
Social persuasion	• Discussions with colleagues or peers (Karimi, 2011; Stevens et al., 2013; Thiel et al., 2020; Vansteelandt et al., 2020) • Feedback from peers, experienced teachers, or academic staff (Weber et al., 2019; Yoo, 2016) • Coaching (from peers or experienced teacher) (Conroy et al., 2019; Hoogendijk et al., 2018; McCullough et al., 2021; Michos et al., 2022) • Collaboration with peers, experienced teacher, or academic staff (Schipper et al., 2018; Stevens et al., 2013)
Physiological reactions	• Yoga, meditation, breathing techniques, or other mindfulness exercises (Ansley et al., 2021; de Carvalho et al., 2021)

Note. These examples are based on the careful review of intervention descriptions and the source coding by the authors.

teacher self-efficacy (Fives et al., 2007; Mergler & Tangen, 2010; O'Neill & Stephenson, 2012; Pfitzner-Eden, 2016a; Rupp & Becker, 2021). However, at the same time, some studies found that the level of teacher self-efficacy did not change significantly after having mastery experiences (Haverback & Parault, 2008; Klassen et al., 2021; Knobloch, 2006). In general, the research is inconclusive when looking at the differential impact of each of the four sources on the level of teacher self-efficacy in general (D. B. Morris et al., 2017). Moreover, only a few studies explicitly compare the effect sizes of the four different sources separately. One example is the study by Tschannen-Moran and McMaster (2009), who explicitly tested different combinations of mastery experiences, vicarious experiences, and social persuasion in an intervention introducing a new instructional reading strategy. While there were increases in teacher self-efficacy in all

conditions, they found the largest gains in the one inheriting all three sources. With this result, it remains unclear whether mastery experiences alone represent the strongest source for teacher self-efficacy or whether mastery experiences develop their full potential in combination with other sources.

Further Influencing Factors

Previous research on teacher self-efficacy and the professional development of teachers, in general, has pointed to further characteristics of interventions that might influence the interventions' effects.

Knowledge. It is still controversial in research how knowledge impacts teacher self-efficacy and whether increases in knowledge are significantly associated with increases in self-efficacy (Dicke, Parker, et al., 2015; Lauermann & König, 2016; D. B. Morris et al., 2017). In qualitative studies, where teachers were asked what they base their self-efficacy on, knowledge is repeatedly cited as a root of self-efficacy (Bautista & Boone, 2015; Palmer, 2011; Phan & Locke, 2015). It is also suggested that knowledge is especially important for individuals with little availability to other sources (Gold et al., 2017; Schwarzer & Warner, 2011). Therefore, in our meta-analysis, we systematically review self-efficacy interventions and investigate whether interventions that target knowledge alone differ from interventions that target at least one of the four sources of self-efficacy, according to Bandura (1997).

Moment of Reflection. Reflection opportunities are seen as an important feature of interventions aiming to change teachers' attitudes or self-efficacy (Bardach et al., 2021; Dunst et al., 2015; Gaudreau et al., 2013; Kayapinar, 2016; Klassen et al., 2021; Prilop et al., 2019; Weber et al., 2018, 2019). Bauer et al. (2020) showed that reflection opportunities at the university were the only significant predictor of a change in student teachers' self-efficacy throughout a semester-long internship. As the four sources of self-efficacy exert their effect on self-efficacy beliefs through various cognitive processes such as reflecting, evaluating, and weighting of information (Bandura, 1997), it is presumable that explicit moments of reflection in interventions support these mostly subconscious processes and strengthen the interventions' effects. In our meta-analysis, we aim to test if interventions that provide a moment of reflection produce larger effects than interventions without a moment of reflection. This would allow us to determine whether a moment of reflection is an important characteristic for successfully promoting teacher self-efficacy.

Duration of Intervention and Period Between Measurement Points. Research on the characteristics of effective teacher training programs furthermore suggests that a certain amount of time is necessary to effect meaningful change (Desimone, 2009; Piwowar et al., 2013; Yoon et al., 2007). Yoon et al. (2007) even concluded that an intervention should last at least 14 hours to impact students' achievement. In a study on the implementation of differentiated instruction in the classroom, the teachers who had completed more professional development hours displayed

higher self-efficacy beliefs in relation to differentiation (Dixon et al., 2014). Since this study is correlational, it is unclear whether more intervention hours cause increases in teacher self-efficacy or whether teachers with higher levels of self-efficacy participate in more professional development activities. The meta-analysis presented here investigates whether there is a linear relationship between the intervention's duration in hours and its effects.

In addition to the duration of an intervention, we considered the period between the assessments of the interventions' effectiveness. In a study with university teachers, the authors pointed out that it takes a year to see positive effects on teacher self-efficacy and that shorter periods may lead to decreases (Postareff et al., 2007). This is in line with the theoretical background that teacher self-efficacy might be rather stable once established (Bandura, 1997) and that an intervention over a short period might compete with a longer period of previous experience. However, Bauer et al. (2020) emphasize that studies investigating different periods of internships could not find any additional effects of longer internships on pre-service teachers' competencies. We will therefore examine whether longer periods are connected with larger effect sizes.

Domains of Teacher Self-Efficacy. Teacher self-efficacy is considered to be domain-specific[3] (Schulte, 2008; Skaalvik & Skaalvik, 2007; Weber et al., 2019). The Teachers' Sense of Efficacy Scale TSES (TSES; also known as the Ohio State Teacher Efficacy Scale) developed by Tschannen-Moran and Woolfolk Hoy (2001) contains three domains of teacher self-efficacy, which capture core elements of teachers' daily life: classroom management, student engagement, and instructional strategies. While teachers might feel competent in one domain, they do not automatically feel competent in another (Bandura, 1997). Empirical findings indicate relationships of different strengths between the domains and different outcomes (e.g., Chen et al., 2020; Kim & Seo, 2018). There are also first indices from intervention studies with pre-service teachers that the domains differ in their malleability due to interventions (Avalos & Bascopé, 2014; Dalioglu & Adiguzel, 2016; Pfitzner-Eden, 2016a). We will therefore check whether the three domains of teacher self-efficacy differ in their malleability due to interventions.

The Promotion of Teacher Self-Efficacy in Different Career Stages

Theoretically, self-efficacy beliefs are assumed to be relatively stable once they are set (Bandura, 1997; Woolfolk Hoy & Spero, 2005). And indeed, longitudinal research indicates that after a significant increase in the first 5 years of teaching, in-service teachers' self-efficacy tends to be rather stable over several years (Künsting et al., 2016; Savolainen et al., 2022). However, other research shows that self-efficacy is certainly malleable over the teaching career. In comparative studies, teachers with more years of experience, that means in-service teachers in comparison with pre-service teachers or in-service teachers with more years in the teaching profession compared to in-service teachers with fewer years of working experience, report higher levels of TSE (D. W. Chan, 2008; Fackler et al., 2021; George et al., 2018; Klassen & Chiu, 2010, 2011; Wolters & Daugherty, 2007; Yeo et al., 2008). This may be because having more years of

experience provides teachers with a broader range of experiences in classroom management, student engagement, and instructional strategies on which they can base their self-efficacy ratings. Also, studies evaluating specific interventions or major environmental and societal changes as virtual teaching during the Covid-19 pandemic show changes in in-service TSE (e.g., Aasheim et al., 2020; Conroy et al., 2019; Pressley & Ha, 2021). Summed up, there is empirical evidence that teachers' years of experience do not prevent teacher self-efficacy changes due to major events or interventions. Nevertheless, it is presumable that interventions with in-service teachers yield smaller changes than interventions with pre-service teachers. A single intervention represents only one of many experiences for in-service teachers and might not cause an immense shift in their self-efficacy. To empirically evaluate this assumption, we analyze in our meta-analysis whether pre-service and in-service teachers react similarly to interventions promoting teacher self-efficacy.

The Interaction of Sources of Teacher Self-Efficacy and Career Stage

Bandura (1997) further describes that different sources may differ in importance depending on someone's experience level. Vicarious experiences and social persuasion are assumed to be important when people face new tasks or have little previous experience (Bandura, 1997). In turn, they may be more relevant for pre-service than in-service teachers. This also is in line with the idea of practice-based teacher education that highlights the importance of vicarious experiences (terminology there: representations of practice) and social persuasion (terminology there: approximation) for pre-service teachers as well (Ball & Forzani, 2009; Grossman, Compton, et al., 2009; Grossman, Hammerness, & McDonald, 2009). However, in interventions using video observation (a possible operationalization of vicarious experiences), both pre-service and in-service teachers showed significant increases in self-efficacy and outperformed the comparison groups without video observation (Groeschner et al., 2018; Weber et al., 2019). The picture is equally ambiguous when we look at social persuasion. While in a study by Tschannen-Moran and Woolfolk Hoy (2007), social persuasion (operationalized as support from colleagues) explained variance in teacher self-efficacy only for teachers with less than 3 years of experience, several studies show positive relations between social persuasion (e.g., feedback) and the level of teacher self-efficacy across all career stages (Klassen & Durksen, 2014; Moulding et al., 2014; Smith et al., 2020). Moreover, several researchers emphasize the importance of social persuasion (e.g., feedback by mentors or coaches) after mastery experiences for the development of pre-service teachers' self-efficacy (Klassen & Durksen, 2014; D. B. Morris et al., 2017; Pfitzner-Eden, 2016b; Tschannen-Moran et al., 1998; Tschannen-Moran & Woolfolk Hoy, 2007; Weß et al., 2020). However, Bardach et al. (2021) could not find an additional influence of social persuasion (operationalized as feedback) beyond mastery experiences on pre-service teachers' self-efficacy.

Finally, it is also unknown whether mastery experiences, generally highlighted as the most important source, are equally important for pre-service and in-service teachers' self-efficacy. Unlike pre-service teachers, in-service teachers have daily opportunities for mastery experiences—at least regarding the domains of classroom

management, student engagement, and instructional strategies. Similar to the economic principle of "diminishing returns," it is questionable whether additional mastery experiences in interventions will have any additional effect on in-service teachers. On the other hand, during interventions, in-service teachers might gain mastery experiences in areas they are not yet comfortable with (e.g., use of new technologies; dealing with heterogeneous classes), and those mastery experiences might, therefore, affect in-service teachers' self-efficacy as well.

In summary, it is still unclear whether the three sources[4]—vicarious experiences, social persuasion, and mastery experiences—have differential effects on pre-service and in-service teachers. It also remains to be examined whether social persuasion combined with mastery experiences is especially beneficial for pre-service compared to in-service teachers.

The Present Study

Although teacher self-efficacy is considered a central variable for teachers' health (e.g., Aloe et al., 2014), their commitment (e.g., Chesnut & Burley, 2015), their teaching quality (e.g., Klassen & Tze, 2014), and their students' academic adjustment (e.g., Kim & Seo, 2018), to date, it is not empirically clear how interventions can support teachers' self-efficacy beliefs and what factors influence the effectiveness of interventions (Klassen et al., 2011). Previous reviews and meta-analyses on teacher self-efficacy and its outcomes have claimed that future research should investigate the role of the four sources of self-efficacy in promoting teacher self-efficacy (Chesnut & Burley, 2015; Klassen et al., 2011; D. B. Morris et al., 2017). Therefore, we will analyze which sources are targeted in each intervention and how the respective combinations of sources influence the effectiveness of the interventions. We will also examine interventions with pre-service and in-service teachers to determine whether there are differences in the promotion of teacher self-efficacy depending on the career stage. Additionally, we will investigate whether the sources of self-efficacy and the career stage of teachers interact in promoting teacher self-efficacy. With this focus on the sources of self-efficacy and the malleability of self-efficacy in different career stages, we aim to contribute further insights into the theoretical tenets of self-efficacy by Bandura (1997). To make it possible for teacher-educators to base their decisions on the best available and relevant evidence, we will rate each study's quality and consider various intervention characteristics. This will allow us to draw a picture of the quality of the current research in promoting teacher self-efficacy and to identify research gaps.

Our research questions and expected findings are as follows:

Research Question 1: Can interventions promote teachers' self-efficacy?
We expect interventions to significantly and positively affect teachers' self-efficacy.
Research Question 2: Are there differences in the effects of the interventions depending on the sources of self-efficacy the intervention targets?
According to Bandura's (1997) theoretical tenets of mastery experiences as the most effective source, we expect interventions that include mastery experiences to produce larger effects than interventions without mastery experiences.

Research Question 3: Are there differences in the interventions' effects regarding the investigated sample's career stage?

Based on the claim from Bandura (1997) that teacher self-efficacy is more malleable in the beginning, we expect larger effects for pre-service teachers than for in-service teachers.

Research Question 4: Are there differences in the effects of the interventions based on the interaction of teachers' career stage and the source targeted by the intervention?

Based on the claim from Bandura (1997) that vicarious experiences and social persuasion are especially important for beginners, we expect interventions targeting vicarious experiences or social persuasion to produce higher effect sizes for pre-service teachers than for in-service teachers. In the literature, it is assumed that especially pre-service teachers may benefit from a combination of social persuasion and mastery experiences (Klassen & Durksen, 2014; D. B. Morris et al., 2017; Pfitzner-Eden, 2016b; Tschannen-Moran et al., 1998; Tschannen-Moran & Woolfolk Hoy, 2007; Weß et al., 2020). Consequently, we expect higher effects of this combination for pre-service teachers than for in-service teachers. In-service teachers have already gained various mastery experiences through their years of experience. Therefore, it is conceivable that additional mastery experiences will not translate into additional self-efficacy. We also expect that pre-service teachers will report higher changes in their teacher self-efficacy after interventions targeting mastery experiences alone or in combination than in-service teachers.

Research Question 5: Are there differences in the effects of the interventions concerning the study's quality?

Generally, studies with higher quality, as indicated by a randomized study design, for example, produce smaller effects than studies with lower quality (Cheung & Slavin, 2016). We hypothesize that studies with a higher study quality (indicated by a study design with randomization and a control group, a reliability coefficient higher than .8, and a detailed description of the intervention) will report smaller effect sizes than studies with lower study quality.

All further potential influencing factors that have been mentioned in the literature, that is, intervention aims to solely improve knowledge, moment of reflection included in the intervention, intervention's duration and period, measure of TSE and domains of teacher self-efficacy used in the study, are analyzed as robustness checks without specified hypotheses.

Method

This meta-analysis follows common guidelines for high-quality systematic reviews and meta-analysis (Alexander, 2020; Page et al., 2021; Pigott & Polanin, 2019). We preregistered the meta-analysis on OSF (https://osf.io/ev7tf/) and report changes after the preregistration along with our justifications and possible effects in Table S1 (in the online version of the journal). Within the preregistration process, we calculated a priori power analyses with the R-package meta-power (Griffin, 2021) to identify the required sample size and number of studies

for the present meta-analysis. Necessary estimates for the power analysis (i.e., effect size, type of effect size, sample size, number of studies, and amount of heterogeneity) were based on the results from a preliminary search that identified a set of eight studies that seemed eligible for our meta-analysis (Aasheim et al., 2020; Al-Awidi & Alghazo, 2012; Aykaç et al., 2019; Bautista & Boone, 2015; Çelebi et al., 2014; Gold et al., 2017; Kunz et al., 2019; Thurm & Barzel, 2020). The a priori power analysis identified a minimum of at least 20 studies with an average sample size of 60 in order to reach 92% power under the conservative assumptions of high heterogeneity (75%) and an overall effect size of $d = 0.4$.

Systematic Literature Search

We applied a variety of search strategies to find as many relevant studies as possible. First, we conducted a systematic literature search in the databases FIS Bildung, ERIC, and Web of Science,[5] following the guidelines from Gusenbauer and Haddaway (2020) and Siddaway et al. (2019). Depending on the database, the search terms scanned either database-specific descriptors or abstract, title, and keywords (see exact search strings in Table S2 in the online version of the journal). The final search terms underwent numerous refinements following preliminary scoping searches and consultations with review team members and an information specialist. The structure of our search terms was inspired by the PICO-scheme (i.e., Population, Intervention, Comparison, Outcome; Kugley et al., 2017; Lefebvre et al., 2021). Our target population is pre- and in-service teachers. We looked for all kinds of interventions, such as programs or training. The outcome of interest is self-efficacy. We did not specify a comparison for our search term. The search was conducted on December 15, 2020, and delivered 3,526 hits. We kept the search algorithm alive and included search alerts until March 1, 2022, delivering a further 627 results. The search was limited to publications since 1977 because Bandura first described the sources of self-efficacy in that year. The languages of references were limited to German and English. No limitations regarding the publication type were made in order to include grey literature and reduce publication bias.

Second, we screened conference books and one online paper repository from relevant educational societies in Germany, Europe, and the United States, namely, Gesellschaft für empirische Bildungsforschung (GEBF [the Society for Empirical Educational Research]), the European Association for Research on Learning and Instruction (EARLI), and the American Educational Research Association (AERA), resulting in 176 possible studies.

Third, we sent out calls for intervention studies via Twitter and 12 mailing lists (5 special interest groups of EARLI, 6 special interest groups of AERA, and 1 mailing list of the German Psychology Society [DGPS]), resulting in 19 possible studies.

After finishing the first round of eligibility screening with the references yielded by the described search strategies, we further conducted a backward and forward citation search for all included studies via the tool citationchaser (Haddaway et al., 2021). Finally, we screened the reference lists from 12 previous reviews and meta-analyses related to teacher self-efficacy or intervention studies for teachers (Aloe et al., 2014; Chesnut & Burley, 2015; Gegenfurtner et al., 2013;

Gesel et al., 2021; Iancu et al., 2018; Kim & Seo, 2018; Klassen & Kim, 2019; Klassen & Tze, 2014; Kraft et al., 2018; Mok & Staub, 2021; D. B. Morris et al., 2017; Zee & Koomen, 2016).

Screening Process

The flowchart (Page et al., 2021) in Figure 1 displays the progress and exclusion reasons from the initially identified references ($n = 4,355$) to the final data set ($n = 119$). All studies included in our meta-analysis meet the inclusion criteria listed in Table 2. Studies that did not fulfill one or more inclusion criteria were excluded. All decisions and reasons for exclusion are documented in the dataset SEIMA_inclusion_exclusion.csv on OSF (https://osf.io/65vfq). In Table S3 (in the online version of the journal), we specify exemplary excluded works.

Based on the inclusion criteria, the screening for eligibility was done in two rounds: In the first round, abstracts and titles were checked; and in the second, full texts were screened for inclusion. Before the screening started, the first author conducted a training with the other two coders (one of the authors and one independent coder) based on a detailed inclusion manual with examples for each criterion. Forty studies were established as the training set, whereas interrater reliability was checked, and unclear issues were discussed after 20 studies. If the interrater reliability was lower than .8, another round of training with a further 20 studies was done. After sufficient interrater reliability ($\kappa \geq .8$) was reached, the independent coding started. To ensure high interrater reliability and high quality during screening, further questions, tips, refinements of the inclusion criteria, and insecure cases were discussed in weekly meetings among the three coders. The screening was done with high sensitivity, so the study was included rather than excluded in unclear cases. Double screening of 462 studies resulted in high interrater reliability ($\kappa = .81$ between coder 1 and 2, $\kappa = .83$ between coder 1 and 3). Disagreements were resolved through discussion. We contacted the authors for studies that met all the inclusion criteria but lacked statistical information to calculate effect sizes. Responses were included until April 2022. Our final data set exceeds the requisites from our a priori power analyses.

Data Extraction

Two of the authors extracted the data necessary to characterize all the studies descriptively and to answer our research questions. A detailed coding manual (see https://osf.io/65vfq) provided instructions regarding the characteristics of (a) the publication, (b) the study, (c) the sample, (d) the intervention, and (e) the calculation of the effect sizes. Table 3 displays an overview of the variables extracted in each category. More detailed information on the study and intervention characteristics is provided below. In cases where information was missing, we coded the variable as not reported.

A trained student assistant double-coded all variables out of 80 studies that left scope for interpretation. Interrater agreements on single variables ranged from 76.5% (intervention description with details) to 95.1% (intervention targeted only knowledge). The values for Cohen's kappa ranged from .32 (intervention description) to .88 (career stage). The low value for intervention description was caused by a strongly imbalanced distribution within this variable (Belur et al., 2021). Disagreements were resolved during regular meetings among the three coders.
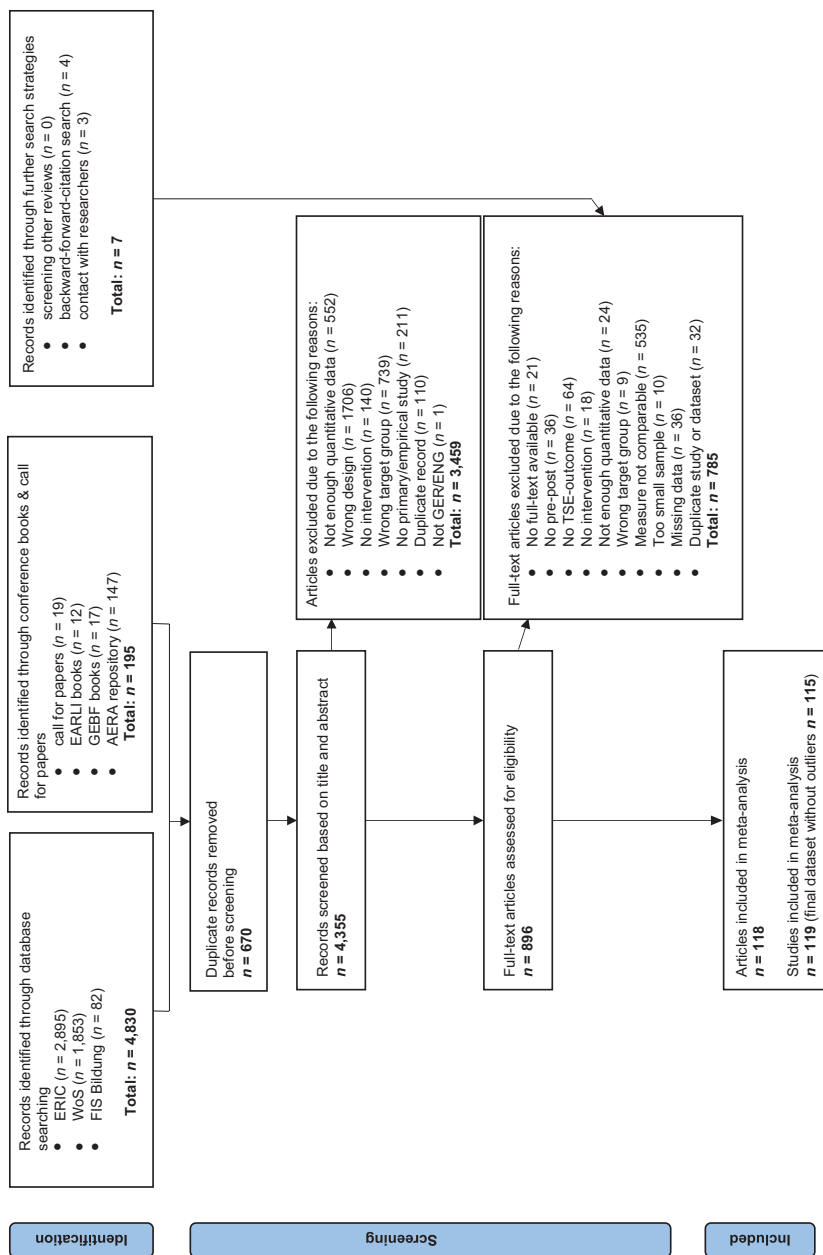
FIGURE 1. *Flowchart representing screening process.*

Identification

Records identified through database searching
- ERIC (*n* = 2,895)
- WoS (*n* = 1,853)
- FIS Bildung (*n* = 82)

Total: *n* = 4,830

Records identified through conference books & call for papers
- call for papers (*n* = 19)
- EARLI books (*n* = 12)
- GEBF books (*n* = 17)
- AERA repository (*n* = 147)

Total: *n* = 195

Records identified through further search strategies
- screening other reviews (*n* = 0)
- backward-forward-citation search (*n* = 4)
- contact with researchers (*n* = 3)

Total: *n* = 7

Screening

Duplicate records removed before screening
***n* = 670**

Records screened based on title and abstract
***n* = 4,355**

Full-text articles assessed for eligibility
***n* = 896**

Articles excluded due to the following reasons:
- Not enough quantitative data (*n* = 552)
- Wrong design (*n* = 1706)
- No intervention (*n* = 140)
- Wrong target group (*n* = 739)
- No primary/empirical study (*n* = 211)
- Duplicate record (*n* = 110)
- Not GER/ENG (*n* = 1)

Total: *n* = 3,459

Full-text articles excluded due to the following reasons:
- No full-text available (*n* = 21)
- No pre-post (*n* = 36)
- No TSE-outcome (*n* = 64)
- No intervention (*n* = 18)
- Not enough quantitative data (*n* = 24)
- Wrong target group (*n* = 9)
- Measure not comparable (*n* = 535)
- Too small sample (*n* = 10)
- Missing data (*n* = 36)
- Duplicate study or dataset (*n* = 32)

Total: *n* = 785

Included

Articles included in meta-analysis
***n* = 118**

Studies included in meta-analysis
***n* = 119** (final dataset without outliers ***n* = 115**)

15

TABLE 2

Inclusion Criteria and Corresponding Reasons for Exclusions

Inclusion criterion	Corresponding reason for exclusion as displayed in the flowchart (Figure 1)
1) The study is an empirical primary study that implemented some form of intervention (e.g., internship, classroom management training, coaching). Purely longitudinal studies were excluded.	• No intervention • No primary/empirical study

2) Sample characteristics:
  a. The sample consists mainly of pre-service teachers, teacher students, or in-service teachers for elementary, middle, high, or vocational schools. Studies with samples consisting mainly of teachers for pre-school, higher education, or special education were excluded.
  b. The sample size is at least 10 persons.
	• Wrong target group
	• Too small sample

3) Teacher self-efficacy is the outcome variable.	• Wrong design • No TSE outcome

4) The measure of teacher self-efficacy:
  a. Studies used the TSES from Tschannen-Moran and Woolfolk Hoy (2001) to measure teacher self-efficacy
  b. Studies with scales with at least one subscale comparable to the three subscales of the TSES were also included. In this case, all items of the scale used for measuring teacher self-efficacy must be available.
	• Measure not comparable

5) The study provides a pre- and post-measure of teacher self-efficacy.	• Wrong design • No pre-post
6) The study provides sufficient quantitative data to calculate the standardized mean change, that is, M, SD, n for pre- and post-test (mixed-methods studies are included).	• Not enough quantitative data • Missing data
7) The study is written in English or German.	• Not GER/ENG
8) The study is available in full text.	• No full-text available

Note. Inclusion criteria 2a, 4a, and 4b were applied during the full-text screening (second round of screening).

Study Characteristics

The *study design* was coded as a one-group study (if no other intervention or control group existed), a quasi-experimental study (at least two groups, but no

randomized assignment), or an experimental study (at least two groups and randomized assignment). The *control group* had to be a "no treatment" condition. If there was a group with an alternative treatment in the study, we coded this as a second treatment group and not as a control group. The *used scale of teacher self-efficacy* was coded with author and title as referenced in the study. The *reliability coefficients* of the used scales were extracted as stated in the studies. However, if the authors did not provide reliability calculations with their used sample, we coded the reliability coefficients as not reported.

Intervention Characteristics

Duration of intervention was coded as stated in the study and then recalculated into hours. We subsequently categorized the hours according to the quartiles into interventions lasting 0.5–6 hours, 7–18 hours, 19–30 hours, and 31–210 hours. The *period between measurement points* was also collected in an open format and later categorized into weeks. A semester was coded as 20 weeks. Periods of less than 1 week were coded as 0 weeks. Regarding the *intervention description*, we first rated whether there was a description (yes or no). Studies only reporting the title of the intervention received the code "intervention description: no." Only studies with a general description were eligible for the coding of the sources. Second, we rated whether the studies provided *details* on the interventions (yes or no), which we used for our study quality rating.

We coded *only knowledge* for interventions focusing only on knowledge acquisition (e.g., evaluating a lecture in Chao et al., 2017) without mentioning any practical exercises or experiences that could be attributed to the sources. The role of knowledge for self-efficacy is still debated in the literature (D. B. Morris et al., 2017). Since we wanted to find out which characteristics of interventions best promote teacher self-efficacy, we also included studies that focused solely on knowledge acquisition. An intervention coded as only knowledge was not further eligible for coding the four self-efficacy sources. *Moment of reflection* was coded as "yes" when the intervention explicitly mentioned some form of introspection (e.g., diary, writing reflection papers, using reflection apps).

Coding the Four Sources of Self-Efficacy. One crucial element of the intervention characteristics was the coding of the *four sources of self-efficacy*. The coding scheme of the four sources was developed according to Bandura's descriptions (1997), a guide for identifying sources of self-efficacy (Brand & Wilkins, 2007), and a review of the measurement of the sources of self-efficacy (D. B. Morris et al., 2017). Anchor examples for the relevant sources were provided in the coding manual (see https://osf.io/65vfq/). The four sources of self-efficacy were coded individually as "mentioned" or "not mentioned."

Mastery experiences were coded as "mentioned" when teachers made actual practical experiences on their own (e.g., taught lessons, developed lesson plans). However, when a study only mentioned "teachers were encouraged to practice," we coded "no mastery experiences mentioned." *Vicarious experiences* were coded as "mentioned" when the teachers observed some model (e.g., the instructor of the intervention modeling a method, observing a classroom teacher). When teachers

TABLE 3

Overview of extracted variables

Category	Variable	Possible values
Publication characteristics	Author	First author's name
	Year of publication	Continuous variable
	Type of publication	• Published • Dissertation • other
Study characteristics	Study design	• One-group study • Quasi-experimental • Experimental
	Control group	Yes/No
	Domain	• Overall self-efficacy (measured with TSES) • Instructional strategies self-efficacy • Classroom management self-efficacy • Student engagement self-efficacy
	Used scale of TSE	Name of scale & author of scale
	Reliability of the scale	• Continuous
Sample characteristics	School level	• Elementary school • Secondary school • Elementary & secondary school
	Career stage	• In-service teachers • Pre-service teachers • In-service and pre-service teachers
	Country	• Name of country of the sample
Intervention characteristics	Intervention description provided	Yes/No
	Intervention description provided with details	Yes/No
	Duration of the intervention	• Continuous (as stated, later categorized)
	Period between measurement points	• Continuous (as stated, later categorized)
	Intervention only targeted knowledge	Yes/No
	Moment of reflection in intervention	Yes/No
	Mapping of sources possible	Yes/No
	Mastery experiences in intervention	• Mentioned • Not mentioned
	Vicarious experiences in intervention	• Mentioned • Not mentioned
	Social persuasion in intervention	• Mentioned • Not mentioned
	Physiological reactions in intervention	• Mentioned • Not mentioned
Effect size calculation (each for pre- and post-test of experimental and control group)	Sample size	• Continuous
	Mean	• Continuous
	SD	• Continuous
	r	• Continuous

only watched videos transferring knowledge without modeling (e.g., watching a lecture), we coded "no vicarious experiences mentioned." *Social persuasion* was coded as "mentioned" whenever there was some kind of social interaction among the teachers (e.g., feedback from supervising teacher, discussing lessons in meetings with mentors or colleagues). *Physiological reactions* were coded as "mentioned" whenever the intervention mentioned some practice of reducing stress levels or working with attributions (e.g., meditations, autogenic training).

We provide detailed information on each article's study, sample, and intervention characteristics in Tables S4 and S5 in the online version of the journal.

Statistical Analyses

Computation and Aggregation of Effect Sizes

From our perspective, studies without a control group may also offer valuable information on the sources targeted in their intervention. Hence, our meta-analysis combines effect sizes from two different study designs (studies with a pre- and post-measure from a one-group study [i.e., only treatment group]) and studies with a pre-and post-measure from the treatment and the control group). The effect sizes for studies without a control group ($n = 83$) are calculated as the standardized mean change, defined as the difference between the post- and pre-test divided by the standard deviation from the pre-test (Becker, 1988; S. B. Morris & DeShon, 2002). The effect size for studies with a control group ($n = 36$) is calculated as the difference between the standardized mean change in the treatment group and the standardized mean change in the control group. In two studies, we transformed given *t*-values into the corresponding effect size measure with the formulas provided by Borenstein et al. (2021) and S. B. Morris and DeShon (2002) (Gresko, 2013; Liaw, 2017).

S. B. Morris and DeShon (2002) present theoretical and empirical prerequisites to combine effect sizes from different study designs. Following their guidelines, we (a) used raw score standardization for both effect sizes to avoid any bias induced by the interaction between subject and treatment and (b) empirically evaluated whether the effects derived from studies with a control group differed significantly from those derived from studies without a control group. Further, based on the theoretical tenets of self-efficacy, we (c) assumed that teacher self-efficacy might not change spontaneously (Bandura, 1997). Therefore, there should be no significant change in the control group. Thus, we (d) empirically investigated whether there was a significant change in TSE in the control groups. As the comparability of the effect sizes across different study designs is a prerequisite, we first conducted these checks.

We often had multiple correlated effect sizes per study because the studies reported values from various subscales or even several treatment groups. We followed the recommendations from Viechtbauer (n.d.) and computed two-level random effects models. Additionally, we applied cluster robust inference methods with a small sample correction, also known as robust variance estimation (RVE) (Pustejovsky & Tipton, 2022; Tanner-Smith et al., 2016; Tipton & Pustejovsky, 2015). All analyses were conducted in R using the packages metafor and clubSandwich (Pustejovsky, 2022; Viechtbauer, 2010).

Assessment of Outliers. To identify possible outliers, we calculated Cook's distance and DFBETAS values (Viechtbauer, 2020; Viechtbauer & Cheung, 2010). The interpretation of Cook's distance is similar to the Mahalanobis distance. We applied a general rule of thumb that every value larger than $4/n$ (in our case $4/119 = 0.034$) is considered an influential study. DFBETAS values show how much the overall effect size would change (in standard deviations) when the study is deleted. We considered any DFBETAS value larger than one as an influential case (Viechtbauer & Cheung, 2010).

Moderator Analyses
We applied multiple meta-regression models for the moderator analyses. Omnibus *F*-tests, implemented in the metafor package, indicated the overall significance of the moderators in each model. All categorical moderators were dummy coded. We calculated five meta-regression models to analyze the second research question, whether the targeted sources of self-efficacy influence the interventions' effects. In the first model, we tested interventions including mastery experiences, against interventions targeting other sources but not mastery experiences. In the second model, we evaluated all interventions targeting only one single source against interventions targeting *only mastery experiences*. In the third to fifth model, we tested interventions targeting only one specific source (i.e., only mastery experiences, only vicarious experiences, or only social persuasion[6]) against interventions targeting this specific source in combination with other sources. To analyze the third research question, whether there are differences in the intervention effects depending on the career stage, we computed one meta-regression model with career stage as a dichotomous moderator. Concerning research question four, how specific sources interact with different career stages, we calculated four meta-regression models. In the first model, we evaluated whether there is an interaction between interventions targeting vicarious experiences (alone or in combination) and teachers' career stage (pre-service or in-service). In the second model, we analyzed whether there is an interaction between interventions targeting social persuasion alone or in combination and the career stage. The third meta-regression model tested whether there are differences in the intervention effects depending on the career stage within the subgroup of interventions targeting mastery experiences *and* social persuasion together. We built a fourth meta-regression model where we analyzed whether there is an interaction between interventions targeting mastery experiences alone or in combination and teachers' career stage.

Assessment of Study Quality. Based on existing rating schemes of study quality (e.g., What Works Clearinghouse, 2020) we chose four variables of the study and intervention characteristics to evaluate the quality of the primary studies. As Wedderhoff and Bosnjak (2020) suggested, we evaluated each variable and built a sum score over the four variables (min. = 0, max. = 4). The four variables were (a) randomized assignment (yes = 1, no = 0), (b) existence of a control group (yes = 1, no = 0), (c) reliability of the measure ($\geq$ .8 = 1, < .8 = 0.5, not reported reliability = 0), and (d) details of the intervention described (yes = 1, no = 0). We

conducted five meta-regression models evaluating the sum score of study quality as a continuous moderator and each of the four variables as categorical moderators (Research Question 5).

Assessment of Publication Bias. First, we addressed the problem of publication bias through the literature search by including unpublished literature, such as dissertations and reports, in our meta-analysis ($n = 19$). Second, we visually checked for asymmetry in a funnel plot, which prints the effect sizes on the x-axis against the corresponding standard errors on the y-axis. We further examined a possible publication bias with statistical tests: We conducted a meta-regression on publication type as well as an Egger's regression test (Egger et al., 1997). The classical Egger's regression test uses the standard error as the predictor and is therefore equivalent to the precision effect test (PET) of Stanley and Doucouliagos (2014). We also conducted a precision effect test with *SE* (PEESE), where the sampling variance is the predictor. We followed the conditional PET-PEESE procedure (Carter et al., 2019; Stanley & Doucouliagos, 2014) to decide which model to use for a cautious interpretation of its intercept as a publication bias-adjusted overall effect size estimate.

Robustness Checks. We conducted six further meta-regression models to check whether our results were robust across the domains of self-efficacy, the used scale of self-efficacy, interventions targeting only knowledge, interventions involving moments of reflection, durations of interventions, and periods between measurement points.

Exploratory Analyses

Several studies also reported follow-up measures. We first calculated a meta-analysis including these effect sizes, too. Second, we ran a meta-regression on the measurement point, which allowed us to test the stability and consistency of the intervention effects over time from pre to post to follow-up measurement. We set the pre-post measures as the reference group and analyzed whether they significantly differed from the effect sizes derived from pre- to follow-up, and post- to follow-up.

Results

Checks of Requirements for Aggregation of Effect Sizes

As effect sizes stemmed from different study designs, we applied two empirical checks to provide evidence that we can combine these effect sizes. Our data showed no significant change in the control groups' teacher self-efficacy ($g = 0.04$, RVE *SE* $= 0.07$, 95% confidence interval [CI] $= [–0.10, 0.17]$, $p = .60$) from pre- to post-test. Thus, the effect sizes of studies with a control group mainly represent the change in the treatment groups' teacher self-efficacy and can be compared to the effect sizes of studies without a control group (only treatment group). We also could not find any significant difference ($p = .13$) between effect sizes based on studies with a control group ($g = 0.43$, RVE *SE* $= 0.07$, CI $= [0.30, 0.56]$) and effect sizes retrieved from studies without a control group ($g = 0.56$, RVE *SE* $= 0.05$, CI $= [0.45, 0.67]$). Therefore, we continued our analyses including effect sizes from both study designs.

Outlier Analyses

Cook's distance and DFBETAS both identified the same seven studies as influential (Cabaroglu, 2014; de Carvalho et al., 2021; Karalar & Altan, 2018; Kissau & Algozzine, 2015; Peebles & Mendaglio, 2014; Whitley et al., 2019; Yilmaz & Koca, 2017). A detailed inspection of the studies revealed problems with the effect sizes or the general credibility in four cases (Karalar & Altan, 2018; Peebles & Mendaglio, 2014; Whitley et al., 2019; Yilmaz & Koca, 2017). We decided to delete the four problematic studies and calculated all the following results without the four influential studies.

Characteristics of Included Studies

One hundred and nineteen studies published between 2005 and 2022 met our inclusion criteria and reported changes in teacher self-efficacy as a consequence of interventions. Four studies had to be excluded as they were identified as outliers. Table 4 presents an overview of the characteristics of the 115 included studies (i.e., publication year, publication type, study design, existence of control group, number of treatment groups, school level, and career stage). The total sample consisted of 11,284 pre-service and in-service teachers and studies stem from 26 different countries (USA: $n = 50$, Turkey: $n = 13$, Germany: $n = 11$). The majority of studies used the TSES ($n = 99$). The 115 included studies (n) reported 146 treatment groups (j), and we extracted a total of 318 effect size estimates (k).

Most of the 318 effect sizes originate from the domain classroom management self-efficacy ($k = 94$), followed by the domain instructional strategies self-efficacy ($k = 82$). Seventy-five effect sizes belong to overall teacher self-efficacy, and 67 stem from the domain student engagement self-efficacy. The interventions lasted between 0.5 and 210 hours, the periods between pre- and post-measurement ranged from less than 1 week to 40 weeks, representing a full academic year. The mean average intervention duration was 27.94 hours (median 18 hours), and the mean average period was 16.8 weeks (median 15 weeks).

For 93 studies reporting on 123 treatment groups, we could code sources of self-efficacy ($k = 260$). In nine studies, we could not code sources of self-efficacy because the intervention targeted knowledge only ($j = 9$, $k = 24$). Almost half of all treatment groups ($j = 66$, $k = 142$) mentioned a moment of reflection. Table 5 displays the frequencies of the individual sources and source combinations in the interventions.

Mastery experiences ($k = 208$) and social persuasion ($k = 200$) are almost equally targeted in the interventions. Physiological reactions are rarely part of the interventions ($k = 14$) and are never targeted alone. In general, all four sources appear more often in combination with other sources than alone. The most often targeted sources were mastery experiences, vicarious experiences, and social persuasion in combination ($k = 111$), followed by mastery experiences and social persuasion ($k = 46$) and only mastery experiences ($k = 35$).

Overall Effect of Interventions on Teacher Self-Efficacy (Research Question 1)

Concerning our first research question about whether interventions can promote teacher self-efficacy at all, we found a positive significant effect of

TABLE 4

Frequencies of Publication, Study and Sample Characteristics

Variables	Frequencies	
	n	%
Publication year	*Min.* 2005, *Max.* 2022	*M* = 2015.9, *Md* = 2016
Publication type		
Published article	96	83.5
Dissertation	15	13.0
Other	4	3.4
Study design		
One-group study	58	50.4
Quasi-experimental study	29	25.2
Experimental study	28	24.3
Existence of control group		
Yes	36	31.3
No	79	68.7
Number of treatment groups	*Min.* 1, *Max.* 4	*M* = 1.3, *Md* = 1
School level of sample		
Elementary school	38	38.8
Secondary school	20	20.4
Elementary and secondary	40	40.8
Career stage of sample		
Pre-service teachers	59	51.3
In-service teachers	52	45.2
Pre- and in-service teachers	4	3.5
Sample size	*Min.* 11, *Max.* 1,322	*M* = 98.1 (*SD* = 134.6) *Md* = 62

Note. N = 115. Frequencies were calculated without the four studies that were identified as outliers. Some percentages do not sum up to 100% because of rounding.

interventions on teachers' self-efficacy (g = 0.47, RVE *SE* = 0.04, CI = [0.40, 0.54], p < .0001). The significant *Q*-statistic (Q [df = 317] = 3146.55, p < .0001) indicates that there is a substantial amount of heterogeneity between the studies. We found a similar overall effect in the exploratory analysis, which also included follow-up measures (g = 0.45, RVE *SE* = 0.04, CI = [0.38, 0.52], p < .0001). Further results from our exploratory analyses are reported in Table S6 (in the online version of the journal).

TABLE 5

Descriptive Overview of Frequencies of Source Combinations in the Interventions

	n	k
Individual sources mentioned (overall)		
Mastery experiences	78	208
Vicarious experiences	59	163
Social persuasion	76	200
Physiological reactions	8	14
Alone vs. in combination		
Mastery experiences only	14	35
Mastery experiences in combination	67	173
Vicarious experiences only	7	17
Vicarious experiences in combination	52	146
Social persuasion only	6	14
Social persuasion in combination	70	186
Physiological reactions only	0	0
Physiological reactions in combination	8	14
Source combinations[a]		
Mastery experiences + vicarious experiences + social persuasion (me/ve/sp)	40	111
Mastery experiences + social persuasion (me/sp)	21	46
Vicarious experiences + social persuasion (ve/sp)	6	16
Mastery experiences + vicarious experiences + social persuasion + physiological reactions (me/ve/sp/pr)	3	7
Mastery experiences + vicarious experiences (me/ve)	4	7
Vicarious experiences + social persuasion + physiological reactions (ve/sp/pr)	2	4
Mastery experiences + social persuasion + physiological reactions (me/sp/pr)	2	2
Vicarious experiences + physiological reactions (ve/pr)	1	1
Mastery experiences + physiological reactions (me/pr)	0	0
Social persuasion + physiological reactions (sp/pr)	0	0
Mastery experiences + vicarious experiences + physiological reactions (me/ve/pr)	0	0

Note. $N_n = 115$, $N_k = 318$.
a. source combinations are presented in descending order based on the number of effect sizes (k).

The Role of the Sources for Interventions' Success (Research Question 2)

First, we focused on the role of mastery experiences and evaluated whether, in general, they moderate the interventions' effects. Interventions that included mastery experiences ($g = 0.51$, RVE $SE = 0.05$, CI $= [0.42, 0.61]$) did not, descriptively or significantly ($p = .40$), differ from interventions without mastery

24

experiences (g = 0.54, RVE *SE* = 0.05, CI = [0.43, 0.66][7]). Next, we analyzed interventions targeting only one source and compared their effect sizes. Interventions targeting only vicarious experiences (k = 17) or only social persuasion (k = 14) did not differ significantly (p = .16) from interventions targeting only mastery experiences (k = 35).[8] Descriptively, interventions targeting only vicarious experiences showed the largest effects (g = 0.83, RVE *SE* = 0.20, CI = [0.34, 1.33]), followed by interventions targeting only mastery experiences (g = 0.43, RVE *SE* = 0.09, CI = [0.25, 0.62]) and interventions targeting only social persuasion (g = 0.35, RVE *SE* = 0.07, CI = [0.17, 0.53]).

In addition, we analyzed whether the interventions' effects differ between interventions targeting only one specific source and interventions targeting this source in combination with other sources. We found no significant difference for all three models contrasting mastery experiences, vicarious experiences, and social persuasion as single sources and in combination (p = .60, p = .53, p = .12). A descriptively slightly larger effect size can be found for interventions where mastery experiences are not accompanied by other sources (g = 0.58, RVE *SE* = 0.14, CI = [0.26, 0.91]), in contrast with interventions targeting mastery experiences in combination with other sources (g = 0.49, RVE *SE* = 0.06, CI = [0.37, 0.62]). The same picture appears when comparing interventions targeting only vicarious experiences (g = 0.61, RVE *SE* = 0.13, CI = [0.21, 1.01]) and interventions targeting vicarious experiences in combination with other sources (g = 0.51, RVE *SE* = 0.05, CI = [0.40, 0.62]). In contrast, interventions targeting only social persuasion (g = 0.35, RVE *SE* = 0.07, CI = [0.17, 0.53]) are less effective than interventions targeting social persuasion in combination (g = 0.51, RVE *SE* = 0.06, CI = [0.40, 0.63]). All the results from these moderation analyses are reported in Table 6. We report the average effect sizes for each source and source combination in Table S7 (in the online version of the journal).

Promoting Teacher Self-Efficacy in Different Career Stages
(Research Question 3)

Four studies had mixed samples with pre-service *and* in-service teachers without stating separate effect sizes for each group. As we cannot draw any conclusions about the role of the career stage from these studies, these four studies were excluded from the moderator analysis on different career stages. As reported in Table 6, our results showed no significant differences between pre-service and in-service teachers in their gains in teacher self-efficacy after the interventions (p = .50). Descriptively, the interventions' effects were slightly larger for samples with in-service teachers (g = 0.48, RVE *SE* = 0.05, CI = [0.37, 0.59]) than for pre-service teachers (g = 0.43, RVE *SE* = 0.05, CI = [0.34, 0.53]).

The Interaction of Career Stages and Sources (Research Question 4)

In order to examine whether different sources of self-efficacy may be differently important for teachers depending on their career stage, we checked four interaction models (see Table 7). The first model evaluated whether there is an interaction between career stages and interventions targeting vicarious experiences alone or in combination. There was no significant interaction (p = .92).

TABLE 6

Overall Effect of Interventions on TSE and Results From Meta-Regression Analyses Regarding the Moderators Targeted Sources and Career Stage (Research Questions 1–3)

	n	k	g	SE	95% CI	df	Q	p
Overall effect (random effects model)	115	318	0.47	0.04	[0.40, 0.54]	112	3146.55*	<.0001
Moderator variables								
Mastery experiences	97[a]	260					2004.9472*	.40
Targeted in intervention	78	208	0.51	0.05	[0.42, 0.61]	78.87		<.0001
Not targeted in intervention	22	52	0.03	0.03	[−0.07, 0.13]	2.9		.40
Single sources targeted	26[a]	66					412.91*	.16
Only mastery experiences (MEO)	14	35	0.43	0.09	[0.25, 0.62]	12.58		<.001
Only social persuasion (SPO)	6	14	−0.08	0.11	[−0.33, 0.17]	9.09		.48
Only vicarious experiences (VEO)	7	17	0.40	0.20	[−0.12, 0.92]	4.94		.11
Mastery experiences targeted	78[a]	208					1694.33*	.60
Alone	14	35	0.58	0.14	[0.26, 0.91]	7.66		<.01
In combination	67	173	−0.09	0.16	[−0.54, 0.36]	4.09		.60
Vicarious experiences targeted	59	163					1169.76*	.53
Alone	7	17	0.61	0.13	[0.21, 1.01]	3.44		.01
In combination	52	146	−0.10	0.13	[−0.89, 0.69]	1.47		.53
Social persuasion targeted	76	200					1545.56*	.12
Alone	6	14	0.35	0.07	[0.17, 0.53]	4.97		<.01
In combination	70	186	0.16	0.09	[−0.06, 0.39]	5.85		.12
Career stage	111	302					2865.49*	.50
In-service teachers (IST)	52	123	0.48	0.05	[0.37, 0.59]	49.9		<.0001
Pre-service teachers (PST)	59	179	−0.05	0.07	[−0.19, 0.09]	104.45		.50

Note. n = number of studies; k = number of effect sizes. Values are displayed in dependency; the first mentioned category is always the reference category.
a. Because some studies are represented in more than one category of this moderator, the total sum of the categories does not add up to the number given here.
*p < .001.

TABLE 7

Results of Meta-Regression Analyses Regarding the Interaction Between Career Stages and Targeted Sources (Research Question 4)

Interaction	n	k	g	SE	95% CI	df	Q	p
Interaction career stage × vicarious experiences	56[a]	156					1046.46*	.92
IST × vicarious experiences alone	3	3	0.62	0.26	[−0.84, 2.08]	1.6		.18
IST × vicarious experiences in combination	25	64	−0.17	0.27	[−1.85, 1.50]	1.47		.61
PST × vicarious experiences alone	3	11	−0.07	0.29	[−1.09, 0.96]	2.5		.84
PST × vicarious experiences in combination	27	78	0.11	0.29	[−1.13, 1.36]	1.99		.73
Interaction career stage × social persuasion	75	197					1439.06*	.01
IST × social persuasion alone	5	12	0.37	0.08	[0.14, 0.60]	3.97		.01
IST × social persuasion in combination	36	83	0.18	0.12	[−0.13, 0.49]	5.16		.19
PST × social persuasion alone	1	2	−0.12	0.08	[−0.34, 0.10]	4.39		.22
PST × social persuasion in combination	33	100	0.01	0.14	[−0.32, 0.33]	7.08		.96
Interaction career stage × mastery experiences	77[a]	205					1597.49*	.33
IST × mastery experiences alone	3	5	0.35	0.12	[−0.25, 0.94]	1.75		.12
IST × mastery experiences in combination	33	73	0.22	0.10	[−0.43, 0.86]	1.41		.21
PST × mastery experiences alone	11	30	0.27	0.16	[−0.24, 0.78]	3.10		.20
PST × mastery experiences in combination	33	97	−0.44	0.17	[−1.05, 0.17]	2.43		.10
Subgroup mastery experiences & social persuasion	21	46					461.66*	.24
IST × me/sp	10	18	0.73	0.26	[0.14, 1.31]	8.99		.02
PST × me/sp	11	28	−0.35	0.29	[−0.95, 0.25]	18.77		.24

Note. n = number of studies; k = number of effect sizes. The first mentioned category is always the reference category. IST = in-service teachers; PST= pre-service teachers; me/sp = mastery experiences and social persuasion. Numbers of studies and effect sizes differ from the numbers of the single sources in Table 6 because studies with a mixed sample of pre-service and in-service teachers were ignored in this analysis.

a. Because some studies are represented in more than one category of this moderator, the total sum of the categories does not add up to the number given here.

*p < .001.

However, in all categories, the degrees of freedom were below four, pointing to caution with interpretability (Tanner-Smith et al., 2016). Descriptively, in-service teachers profit more from interventions targeting vicarious experiences alone ($g = 0.62$, RVE *SE* = 0.26, CI = [–0.84, 2.08]) than from interventions targeting vicarious experiences in combination ($g = 0.44$, RVE *SE* = 0.07, CI = [0.31, 0.58]). The same is true for pre-service teachers. However, for them, the difference in teacher self-efficacy gains between interventions targeting vicarious experiences alone ($g = 0.55$, RVE *SE* = 0.11, CI = [0.16, 0.94]), and interventions targeting vicarious experiences in combination ($g = 0.49$, RVE *SE* = 0.07, CI = [0.35, 0.63]) is much smaller. There was no significant interaction between career stages and interventions targeting social persuasion alone or in combination ($p = .96$), although the moderator as a whole was significant ($p < .05$). As the distribution of the effect sizes across the different groups was extremely unbalanced, we are not interpreting this significance. Compared with the first model, however, the descriptive results look different. Pre-service and in-service teachers report more gains in teacher self-efficacy in interventions targeting social persuasion *in combination* (for pre-service teachers: $g = 0.44$, RVE *SE* = 0.06, CI = [0.31, 0.57]; for in-service teachers: $g = 0.55$, RVE *SE* = 0.09, CI = [0.37, 0.73]). The changes in interventions targeting social persuasion alone are $g = 0.25$ (RVE *SE* = 0.00, CI = [0.24, 0.27]) for pre-service teachers and $g = 0.37$ (RVE *SE* = 0.08, CI = [0.14, 0.60]) for in-service teachers.

In a third step, we evaluated whether pre-service teachers benefit more from the combination of mastery experiences and social persuasion than in-service teachers. We found no significant moderating effect ($p = .24$). In-service teachers ($g = 0.73$, RVE *SE* = 0.26, CI = [0.14, 1.31]) benefited descriptively more from this combination than pre-service teachers ($g = 0.38$, RVE *SE* = 0.12, CI = [0.11, 0.65]).

We also did not find any interaction in our fourth model evaluating the role of interventions targeting mastery experiences alone or in combination for the two different career stages ($p = .33$). Again, the degrees of freedom were below four in all categories, pointing to caution with interpretability (Tanner-Smith et al., 2016). At least descriptively, in-service teachers profit more from combined mastery experiences ($g = 0.57$, RVE *SE* = 0.09, CI = [0.38, 0.76]) than from mastery experiences alone ($g = 0.35$, RVE *SE* = 0.12, CI = [–0.25, 0.94]). Pre-service teachers, in contrast, benefit more from mastery experiences alone ($g = 0.62$, RVE *SE* = 0.11, CI = [0.35, 0.88]) than from combined mastery experiences ($g = 0.40$, RVE *SE* = 0.07, CI = [0.25, 0.54]).

Moderating Effect of Study Quality (Research Question 5)

Less than half of the 115 studies reached a study quality rating of at least two points out of four possible points ($n = 55$). Nine studies fulfilled all criteria and received the highest rating of study quality (four points). In comparison, 14 studies fulfilled none and received the lowest rating of study quality (zero points). Table S4 (online only) displays the study-level aggregated study quality rating. Figure S1 (online only) displays the distribution of the study quality ratings of the 318 effect sizes, Table S8 (online only) the results from the meta-regressions.

The sum score of study quality did not significantly moderate the interventions' effects ($p = 0.18$). Descriptively, there was a very small positive relationship between study quality and effect size (b = 0.05). The studies with the highest rating of study quality (four points) reached, on average, an effect of $g = 0.59$ (RVE *SE* = 0.10, CI = [0.39, 0.79]), while the studies with the lowest rating of study quality (zero points) reached, on average, an effect of $g = 0.37$ (RVE *SE* = 0.07, CI = [0.22, 0.52]). As Wedderhoff and Bosnjak (2020) suggested, we also analyzed the single items of our sum score of study quality. Neither the randomized assignment ($p = 0.35$), the existence of a control group ($p = 0.40$), nor the reliability of the used measure ($p = 0.66$) significantly influenced the effects (see Table 8). Descriptively, studies with randomized assignment ($g = 0.54$, RVE *SE* = 0.09, CI = [0.36, 0.72]), without a control group ($g = 0.49$, RVE *SE* = 0.04, CI = [0.40, 0.57]) and either no reported reliability ($g = 0.48$, RVE *SE* = 0.07, CI = [0.34, 0. 61]) or a reliability coefficient higher than 0.8 ($g = 0.49$, RVE *SE* = 0.05, CI =[0.39, 0.58]) found larger gains than studies without randomized assignment ($g = 0.45$, RVE *SE* = 0.04, CI = [0.37, 0.53]), with a control group ($g = 0.42$, RVE *SE* = 0.06, CI = [0.29, 0.55]) and reliability coefficients of lower than 0.8 ($g = 0.40$, RVE *SE* = 0.08, CI = [0.24, 0.57]).

However, the detailed description of interventions was a significant moderator ($p < 0.01$). Interventions that were described with some details reached significantly larger gains in teacher self-efficacy ($g = 0.55$, RVE *SE* = 0.05, CI = [0.45, 0.64]) than interventions without details in the description ($g = 0.35$, RVE *SE* = 0.05, CI = [0.25, 0.44]).

Robustness Checks

To adhere to page constraints, we present the results from the robustness checks in Supplement S9 (in the online version of the journal).

Publication Bias

The funnel plot (see Figure 2) looks asymmetrical. Studies with negative effect sizes and larger standard errors seem to be missing. Four of the five most extreme points (indicated by the numbered effect sizes) are located on the funnel plot's right side. The funnel plot represents the quite large sample sizes ($M = 98.1$) in the studies and the high heterogeneity among them.

Although not taking into account the dependency among our correlated effect sizes, we applied precision-effect tests (Stanley & Doucouliagos, 2014) to retrieve statistical information about the funnel plot's symmetry. The Egger regression test (1997) was significant ($p < .0001$), indicating asymmetry in the funnel plot. As the Egger regression test was significant, we conducted the PEESE. The PEESE was significant as well ($p < .0001$). Its limit estimate can be cautiously interpreted as the publication bias-adjusted average true effect (b = 0.29, CI = [0.24, 0.33]).

The moderator publication type was not significant ($p = .66$). We found the biggest effect sizes for the publication type "other" ($g = 0.55$, RVE *SE* = 0.25, CI = [–0.24, 1.33]. Published articles reported slightly larger effect sizes ($g = 0.48$, RVE *SE* = 0.04, CI = [0.40, 0.56]) than dissertations ($g = 0.39$, RVE *SE* = 0.09, CI = [0.21, 0.57]).
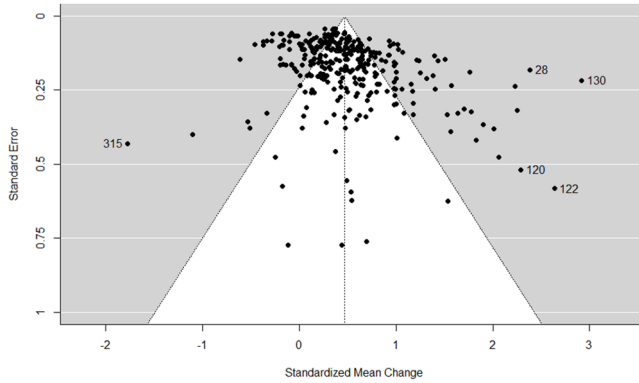
FIGURE 2. *Funnel plot of the effect sizes for interventions promoting teacher self-efficacy.* The five most extreme effect sizes are indicated by their respective number.

Discussion

This preregistered meta-analysis (https://osf.io/ev7tf) investigated whether interventions in pre-service teacher education and in-service teacher training can promote teacher self-efficacy. One hundred and nineteen studies identified by systematic search approaches were included to answer this question. We further analyzed whether the sample's career stage, the interventions' targeted sources, the interaction between these two, and specific study characteristics (e.g., study quality) moderated the effects. One main finding was that teacher self-efficacy can be promoted in interventions—equally for pre-service and in-service teachers. While the different source combinations did not produce statistically significant differences in the intervention effects, the effects were significantly larger when the intervention was described with details (one aspect of study quality), included moments of reflection, and had a certain duration (between 7 and 30 hours).

Research Question 1: Interventions as Ways to Support Teacher Self-Efficacy Promotion

This meta-analysis shows that teachers report, on average, significantly higher teacher self-efficacy beliefs after interventions than before. To evaluate the importance of this effect size, we first compare it with the results from similar meta-analyses; we then judge it concerning the costs of the interventions; and finally, we highlight its robustness across various methodologically relevant variables.

Compared to the only other meta-analysis on the promotion of teacher self-efficacy to date, our average effect size of $g = 0.47$ is somewhat smaller than the average effect of $g = 0.68$ in Mok et al. (2023). However, this difference is quite reasonable as we included a broader range of interventions and career stages in our meta-analysis, unlike Mok et al. (2023) who focused on individual support activities by mentors, coaches, or supervisors and examined only pre-service and beginning teachers' self-efficacy. As there are no other meta-analyses on the promotion of teacher self-efficacy, we further refer to general teacher professional

development and three meta-analyses on interventions for constructs related to and comparable with teacher self-efficacy, namely teacher burnout, teacher stress, and teacher well-being. Iancu et al. (2018) conducted a meta-analysis of 23 intervention studies aimed at reducing teacher burnout. Zarate et al. (2019) analyzed the effects of 18 mindfulness-based interventions on teacher stress and burnout, and Oliveira et al. (2021) evaluated the effects of 43 studies using social and emotional learning on teachers' well-being and stress. All three meta-analyses included only studies with control groups and in-service teachers. They found significant effects on enhancing well-being ($g = 0.35$ in Oliveira et al., 2021), reducing teacher burnout ($d = 0.18$ in Iancu et al., 2018; standardized mean difference (SMD) $= 0.33$ in Zarate et al., 2019), and reducing teacher stress (SMD $= 0.53$ in Zarate et al., 2019; $g = 0.34$ in Oliveira et al., 2021). Our average effect size of $g = 0.47$ and—for more comparability—the average effect size for studies with a control group of $g = 0.42$ exceed, with one exception, the effects of the mentioned meta-analyses.

Several authors advise that intervention effect sizes should be considered in relation to their costs (Harris, 2009; Kraft, 2020; Levin & Belfield, 2015). Many interventions in pre-service teacher education as internships or specific video-based seminars were part of the university program and did not incur any extra costs (e.g., Gold et al., 2017; Michos et al., 2022; Weber et al., 2019). Especially if pre-service teachers analyzed videos they took with their own devices, these interventions were almost free of costs. Interventions with in-service teachers incur costs through the hours teachers spend in other classrooms for lesson study, for example, or in paying external coaches. The costs for a lesson spent by teachers in another classroom in the United States can be roughly calculated as follows: Based on the assumptions of public school teachers' average salary of $65,090 per year (USA Facts, n.d.) and 36 weeks teaching 25 lessons per week, a single lesson costs $72.32. The literature on coaching costs reports ranges from $170 per contact hour to $400 per teacher (Barrett & Pas, 2020; Knight, 2012). These exemplary costs of interventions promoting and enhancing in-service teachers' self-efficacy are well worth the money when compared with the financial costs caused by teacher drop-out and teacher burnout or the societal costs caused by teaching positions that are not staffed. Studies calculate the costs caused by burnout among physicians to be at least $7,600 per year (Han et al., 2019; Shanafelt et al., 2017). This figure is comparable for teachers, for whom, unfortunately, no specific figures are available yet. The cost of replacing a teaching position is sometimes quoted as up to $20,000 per teacher (Learning Policy Institute, 2017). In this comparison, interventions to increase self-efficacy, and thus presumably to increase teachers' well-being and job satisfaction, are inexpensive and worth the money.

To prove the robustness of our overall effect, we tested a variety of methodologically relevant characteristics (study design, study quality, used scale, and domain of teacher self-efficacy) as moderators in meta-regressions. The effect size was robust across all mentioned methodological moderators, indicating that teacher self-efficacy can be promoted through interventions regardless of specific methodological choices within the studies. Although we included studies with different rigorous study designs, namely, with and without control group

comparison, the overall effect is comparable in size to the effect of methodologically rigorous experimental intervention studies (e.g., Thiel et al., 2020). This supports our claim that the overall effect size is credible and that interventions can support teacher self-efficacy development.

Research Question 2: The Role of Specific Sources for Promoting Teacher Self-Efficacy

Before conducting our analyses, each intervention was carefully reviewed with a detailed coding scheme to extract the targeted sources of self-efficacy.

First, it became obvious that some sources have never been targeted in the analyzed interventions. To be more specific, it is noticeable that physiological sources are present in all source combinations that never occurred within the interventions (e.g., only physiological reactions, combination of mastery experiences and physiological reactions). This finding aligns with a general underrepresentation of physiological factors in the research on motivation in education (Martin et al., 2023; Pekrun, 2023). It remains a point for future research to investigate the role of physiological reactions (and their attributions) in promoting teacher self-efficacy.

Within the pool of interventions targeting only one specific source, interventions targeting only vicarious experiences descriptively outreached mastery experiences. Although we must interpret this large effect with caution, as only seven studies contributed to this moderator and the broad corresponding confidence interval indicates high heterogeneity among these studies, we still think this large effect points to the important role of vicarious experiences for the promotion of teacher self-efficacy (e.g., Kumschick et al., 2017). This finding also blends in well with other research findings that highlight the general relevance of vicarious experiences in teacher education per se (e.g., representations of teaching practice for core practices of teaching in Ball & Forzani, 2009; Grossman, Compton, et al., 2009; Grossman, Hammerness, & McDonald, 2009; cognitive modeling for pre-service teachers' planning and instruction skills in Mok & Staub, 2021; or video observations for teachers' immersion in Seidel et al., 2011).

Against our theoretically based hypothesis (e.g., Bandura, 1997), interventions including mastery experiences did not produce larger effects than interventions targeting any other source combination without mastery experiences. Due to our reliance on the studies' intervention descriptions, we could only code *whether* mastery experiences were offered in interventions and not whether and to what extent these experiences were experienced as *successful* by the teachers. Therefore, it is presumable that some participants experienced failures, which may have reduced the effect sizes of interventions involving mastery experiences. Future research should explore how failure (unsuccessful mastery experiences) contributes to the promotion of teacher self-efficacy. Nevertheless, the effect size of $g = .43$ for only mastery experiences can still be interpreted as a substantial effect size for practicing various teaching-related tasks.

Third, in comparison with mastery and vicarious experiences, social persuasion plays a minor role in promoting teacher self-efficacy, as proposed in theory (Bandura, 1997; Schunk, 1991). Within the pool of interventions targeting only

one specific source, only social persuasion reached the descriptively lowest effect. This result is similar to findings from longitudinal studies reporting no effect of social support, especially only emotional support, on the development of teacher self-efficacy (e.g., Hartl & Holzberger, 2022). However, the source social persuasion covers a wide range of activities, including, for example, group discussions, mentoring and coaching, or feedback. Whereas the (low) effect size for social persuasion is similar in size to another meta-analysis investigating the effects of coaching, mentoring, and supervision on pre-service teachers' instruction skills (Mok & Staub, 2021), research further points to several conditions that may be decisive for the effectiveness of feedback, for example, feedback givers' expertise (Prilop, et al., 2021; Weber et al., 2018), the content delivered (Wisniewski et al., 2020), or the addressed level (e.g., task, process, etc.) (Hattie & Timperley, 2007). It could be that some "negative effects of feedback" have lowered the effect size of social persuasion.

The fact that we could not find significant differences between the sources also complements reflections on teacher education in general and studies on teacher self-efficacy showing that different elements of training (here: the different sources of self-efficacy) are intertwined and do not impact teacher self-efficacy independently but are instead all necessary for an overall effect (Bach, 2022; Ball & Forzani, 2009; Bardach et al., 2021; Grossman, Compton, et al., 2009; Grossman, Hammerness, & McDonald, 2009; Pfitzner-Eden, 2016b). As we had to rely on the intervention descriptions in the primary studies, we could only code whether a source was mentioned. However, we could not code the order, the amount, or the intensity of the mentioned sources of self-efficacy in the interventions nor the perceived valence (for example, whether the mastery experiences were perceived as successful mastery experiences). To better understand the underlying mechanisms, future research should apply more experimental studies that systematically investigate different combinations, orders, amounts, and perceived valences of the sources (e.g., Tschannen-Moran & McMaster, 2009).

Research Question 3: Promotion of Teacher Self-Efficacy Works in All Career Stages

Against our hypothesis, we found no significant difference in the promotion of pre-service and in-service teacher self-efficacy. The almost identical effect sizes for in-service and pre-service teachers raise the question of whether teacher self-efficacy is a less inherently stable construct than previously thought (Bauer et al., 2020; Rupp & Becker, 2021).

Although teacher self-efficacy seems malleable across the teaching career, we need to take a more differentiated view of this result. First, it is presumed that pre-service teachers tend to overestimate their self-efficacy before their first contact with reality (Klassen & Chiu, 2011; Pfitzner-Eden, 2016a; Schüle et al., 2017). The post-test measure, at which the pre-service teachers can base their self-rated teacher self-efficacy on real experiences, might provide a more reliable estimation of their self-efficacy. Therefore, the calculated standardized mean change for pre-service teachers might be influenced by a reality shift and might be underestimated. Second, due to our method, we had to rely on existing studies and cannot

completely dismiss the possibility that in-service teachers received different interventions than pre-service teachers and that differences in the interventions ruled out differences between career stages. It was also striking that only one study with pre-service teachers conducted a follow-up measurement; all other follow-up values came from studies with in-service teachers. Therefore, we could not investigate whether there are different mechanisms responsible for the promotion of TSE in different career stages. It is a task for future research to carefully examine identical interventions with pre-service and in-service teachers separately and to use follow-up measures in samples with pre-service teachers as well.

Third, due to our reliance on existing studies and samples, we could only differentiate globally between pre-service and in-service teachers. There are, for example, hints that teacher self-efficacy develops differently in different stages of in-service teachers. The findings from Swan et al. (2011) and Woolfolk Hoy and Spero (2005) lift out the first year of teaching as a critical phase where the teachers' self-efficacy decreases. Klassen and Chiu (2010) report a rise in teacher self-efficacy until the middle of the career (around 23 years of teaching experience), followed by a decrease in teacher self-efficacy. We support the claim from van der Scheer and Visscher (2016) that future research should investigate whether different subgroups of in-service teachers react the same to interventions (e.g., beginning, middle-career, and late-career teachers).

Research Question 4: The Interaction of Different Career Stages and Specific Sources

Contrary to our hypotheses, pre-service teachers did not benefit specifically from other sources than in-service teachers. Given that neither sources (Research Question 2) nor career stage (Research Question 3) significantly moderated the intervention effects when investigated separately, the lack of a significant interaction effect may not be surprising. At the same time, the descriptives do reveal interesting patterns: First, it is noticeable that in some of the tested combinations, only very few effect sizes were represented (e.g., $k = 2$ for only social persuasion and pre-service teachers, $k = 3$ for only vicarious experiences and in-service teachers, and $k = 5$ for only mastery experiences and in-service teachers). Interventions targeting only social persuasion are more studied for in-service teachers (e.g., Hoogendijk et al., 2018), whereas interventions targeting only vicarious experiences are more studied for pre-service teachers (e.g., Thiel et al., 2020). Future research should focus especially on the underresearched combinations (e.g., only vicarious experiences with in-service teachers) in order to investigate the interactions of different sources and career stages more systematically.

For pre-service teachers, interventions offering only mastery experiences produced the highest self-efficacy changes. This finding is partly in line with Bandura's (1997) claim of mastery experiences as the strongest source of self-efficacy. As also demanded by practice-based teacher education researchers (Ball & Forzani, 2009; Grossman, Compton, et al., 2009; Grossman, Hammerness, & McDonald, 2009), we conclude that teacher educators should prepare early and various opportunities for enactive experiences to support pre-service teachers' self-efficacy (e.g., developing a lesson plan at university). For

in-service teachers, the descriptively largest effect sizes appeared in interventions offering mastery experiences combined with social persuasion or only vicarious experiences. They seem more likely to benefit from experiences that are not part of their daily routine, such as classroom observation or coaching. As a practical implication for teacher training, we advise enriching in-service teachers' daily experiences with videotaping, collegial feedback sessions, or coaching (von Suchodoletz, et al., 2018). It remains a point for future research to investigate what role the regularity of such offers plays and which types of social persuasion (collegial feedback is different from feedback from a coach) impact in-service teachers' self-efficacy the most.

Research Question 5: The Relevance of Study Quality

We could not confirm our hypothesis that studies with a higher study quality score will report significantly smaller effect sizes than those with a lower score. In contrast, studies with higher quality scores reached at least descriptively larger gains in teacher self-efficacy than studies with lower quality scores. This result is encouraging, as it shows that the promotion of teacher self-efficacy may be particularly successful when interventions are carefully thought through and studies are not conducted ad hoc. This interpretation is further supported by the fact that out of the four variables of the sum score of study quality, only the variable detailed description of the intervention produced significantly different results. Interventions with detailed descriptions reached larger gains in TSE than those without details. We assume that a more detailed description in the article accompanied a more intense engagement with the participants. In summary, we have learned that methodological study aspects seem less relevant than content-related aspects for promoting TSE (similar to findings on teacher professional development in general; see Kennedy, 2016). However, our findings also point to a problem in the previous and current research. Forty-three effect sizes from 16 studies reached a study quality rating of zero. We kept these studies in our meta-analysis as we did not find a moderating effect on study quality. Nevertheless, we encourage future research to engage more in high-quality studies.

Limitations

Despite our large database and robust findings, we must consider several limitations. There were over 30 potentially relevant studies for which we contacted the authors to provide the needed statistical data. However, we either got no response or the answer that the data was no longer available. Thus, it is possible that we missed some relevant articles. Moreover, the funnel plot and the PEESE indicated that we had an above-average proportion of studies with large samples and positive effect sizes and missed smaller samples with larger standard errors and studies with negative effect sizes. Our overall effect size might thus overestimate the true effect. With our focus on intervention features and due to our meta-level, we could not consider the individual cognitive processes of the teachers or specific contextual backgrounds that may interact with the sources of self-efficacy.

Regardless of our efforts to make the coding of the targeted sources as objective as possible (for example, by regular meetings and anchor items in the coding

manual), coding the sources of self-efficacy in the interventions was highly infer-
ential. Despite our moderate to substantial interrater agreement, it is possible that
we did not capture every source combination perfectly because of the remaining
subjectivity within the coding of the sources of self-efficacy.

Although we tested a variety of theoretically based moderators, there remained
substantial heterogeneity in each of our meta-regressions. Thus, there may be
other variables than the considered ones that explain the between-study heteroge-
neity better.

Future Research

Our meta-analytical aggregation of results from 115 intervention studies,
capturing more than 11,000 pre- and in-service teachers, shows that teacher
self-efficacy can be promoted through interventions. At the same time, our
careful review reveals several research gaps for future research. It is still an
open question as to which factors are decisive for a successful intervention
promoting teacher self-efficacy. One starting point would be to conduct rigor-
ous experimental research in which the four sources of self-efficacy are sys-
tematically varied and analyzed. Future intervention studies should consider
the interplay between previous experiences, previous knowledge, contextual
class- and school-level variables, and the four sources of self-efficacy. As part
of investigating these interplays, exploring the effects of interventions in dif-
ferent in-service career stages, for example, beginning, middle-career, and
late-career teachers, would be useful. Additionally, other operationalizations of
"experience," for example, the familiarity with the intervention's content,
should be considered, too. Furthermore, future research should also take into
account that people differ in their cognitive processes of integrating informa-
tion from different sources of self-efficacy (e.g., multiplicative, additive;
Bandura, 1997) and consider such mechanisms as mediators. Additionally, dif-
ferentiations of the single sources should be considered in intervention studies.
There are, for example, indices from related fields that perceived similarity
with a model enhances the effect of vicarious experiences on self-efficacy
(Adams, 2004; Braaksma et al., 2002; Huang, 2017; Schunk, 1987). However,
this relationship has not yet been empirically explored within teacher educa-
tion. There is another research gap in investigating the consistency of interven-
tion effects with pre-service teachers. We suggest that future studies include
follow-up measures in their study designs. Finally, there is a need to evaluate
the effects of interventions on teacher self-efficacy and later teacher behavior
(e.g., classroom management).

Practical Implications

We derived some practical implications from our results that teacher self-
efficacy can be promoted in teacher education and training. Based on our
findings that mastery experiences and vicarious experiences play an important
role in promoting teacher self-efficacy, we recommend that university teach-
ers provide their students in teacher education with as many opportunities for
those sources as possible. Lists with teachers' core practices can be a good

inspiration (Michigan Department of Education, n.d.; TeachingWorks, n.d.). Many topics in teacher education curricula contain aspects of core practices and can be used to create mastery (= active practice) and vicarious experiences for the students. This can be, for example, that the students work out lesson ideas or engage in authentic role plays that they video-record. Our results show that in-service teachers' self-efficacy can be promoted as well. Therefore, we encourage school leaders and in-service teachers to enhance their own and their colleagues' self-efficacy by including time for mutual classroom visits (e.g., lesson study) and exchange with colleagues or external coaches. As the promotion of teacher self-efficacy, in general, benefits from a certain amount of time (at least 7 hours in our data), we recommend that responsible administrators in teacher education and training establish regular opportunities for reflections, mastery experiences, vicarious experiences, social persuasion, and exercises of attribution of physiological reactions to support the continuous development of self-efficacious teachers.

Conclusion

Promoting teacher self-efficacy is an emerging field, and this meta-analysis is the first systematic summary of 115 intervention studies. We included studies with pre- *and* in-service teachers in our sample and particularly investigated the role of the four sources of self-efficacy in the promotion of teacher self-efficacy. We thus developed a systematic code book on assigning intervention elements to the four sources of self-efficacy and conducted a systematic review of targeted sources in interventions. In general, we found a positive, significant effect of interventions on the level of teacher self-efficacy, regardless of the career stages of the teachers. This positive effect was robust across a variety of moderators. Neither of the tested source combinations significantly moderated the effects. The effect size was significantly larger for studies with a more detailed description of the intervention and for interventions that included moments of reflection. Our overall effect size was comparable to other meta-analyses testing interventions on teacher stress and well-being. Therefore, we conclude that teacher self-efficacy can indeed be promoted and hope that teacher education and teacher training will continue in their efforts to support the development of teachers' self-efficacy.

Funding

ORCID iDs

Janina Täschner ⓘD https://orcid.org/0000-0002-0221-493X
Theresa Dicke ⓘD https://orcid.org/0000-0001-8868-2242
Sarah Reinhold ⓘD https://orcid.org/0000-0002-5972-3428
Doris Holzberger ⓘD https://orcid.org/0000-0001-6535-289X

Notes

This meta-analysis was preregistered on OSF (https://osf.io/ev7tf). All materials, data, and analytic code underlying this study are available at https://osf.io/65vfq/. We would like to thank Leonie Gossner and Georgios Mitsostergios for their great support during the screening and coding process.

[1] In this meta-analysis, we define interventions as all organized learning opportunities that differ from the everyday experiences of pre- and in-service teachers.

[2] In the original work by Bandura (1997) mastery experiences are strongly connected to successes. Because this (individual) information is not available in intervention studies with multiple participants, we conceptualize mastery experiences as operational experiences and elaborate on possible biases due to this simplification in the discussion.

[3] In the self-efficacy literature, the terms "domain" and "dimension" appear interchangeably. We use the term *domain* in the following text to make clear that there are different areas of action in which teachers can feel capable.

[4] The fourth source of self-efficacy, physiological and emotional reactions, has received very little attention in research on teacher self-efficacy (D. B. Morris et al., 2017). In fact, the empirical evidence on physiological and emotional reactions is so limited that we cannot draw any conclusions about possible interactions with teachers' career stages.

[5] ERIC was accessed via EBSCOhost, the other databases were accessed directly.

[6] No intervention targeted only physiological reactions; therefore, we could not calculate a meta-regression for this case.

[7] In the text, we report the absolute effect size, the according standard error, and the confidence interval for each category. This was done by removing the intercept in the meta-regression.

[8] No intervention targeted only physiological reactions (see also Table 5).

[9] The reference list of all studies included in this meta-analysis can be found in S10 in the online version of the journal.

References[9]

Aasheim, M., Fossum, S., Reedtz, C., Handegård, B. H., & Martinussen, M. (2020). Examining the Incredible Years Teacher Classroom Management program in a regular Norwegian school setting: Teacher-reported behavior management practice, problem behavior in classroom and school environment, teacher self- and collective efficacy, and classroom climate. *SAGE Open*, *10*(2). https://doi.org/10.1177/2158244020927422

Adams, K. (2004). Modelling success: Enhancing international postgraduate research students' self-efficacy for research seminar presentations. *Higher Education Research & Development*, *23*(2), 115–130. https://doi.org/10.1080/07294360 42000206618

Al-Awidi, H. M., & Alghazo, I. M. (2012). The effect of student teaching experience on preservice elementary teachers' self-efficacy beliefs for technology integration in the UAE. *Educational Technology Research and Development*, *60*(5), 923–941. https://doi.org/10.1007/s11423-012-9239-4

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

Aloe, A. M., Amo, L. C., & Shanahan, M. E. (2014). Classroom management self-efficacy and burnout: A multivariate meta-analysis. *Educational Psychology Review*, *26*(1), 101–126. https://doi.org/10.1007/s10648-013-9244-0

Ansley, B. M., Houchins, D. E., Varjas, K., Roach, A., Patterson, D., & Hendrick, R. (2021). The impact of an online stress intervention on burnout and teacher efficacy. *Teaching and Teacher Education*, *98*, Article 103251. https://doi.org/10.1016/j.tate.2020.103251

Avalos, B., & Bascopé, M. (2014). Future teacher trajectory research: Its contribution to teacher education and policy. *Education as Change*, *18*(1), 19–32. https://doi.org/10.1080/16823206.2013.877353

Aykaç, N., Ulubey, Ö., Çelik, Ö., & Korkut, P. (2019). The effects of drama on pre-service teachers' affective traits about teaching. *International Journal of Contemporary Educational Research*, *6*(2), 338–351. https://doi.org/10.33200/ijcer.587566

Bach, A. (2022). *Selbstwirksamkeit im Lehrberuf. Entstehung und Veränderung sowie Effekte auf Gesundheit und Unterricht [Self-efficacy in the teaching profession. Emergence and change as well as effects on health and teaching]*. Waxmann Verlag GmbH. https://doi.org/10.31244/9783830995166

Ball, D., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, *60*(5), 497–511. https://doi.org/10.1177/0022487109348479

Bandura, A. (1995). Exercise of personal and collective efficacy in changing societies. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 1–45). Cambridge University Press.

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

Bardach, L., Klassen, R. M., Durksen, T. L., Rushby, J. V., Bostwick, K. C., & Sheridan, L. (2021). The power of feedback and reflection: Testing an online scenario-based learning intervention for student teachers. *Computers & Education*, *169*, Article 104194. https://doi.org/10.1016/j.compedu.2021.104194

Barrett, C. A., & Pas, E. T. (2020). A cost analysis of traditional professional development and coaching structures in schools. *Prevention Science*, *21*(5), 604–614. https://doi.org/10.1007/s11121-020-01115-5

Bauer, M., Traub, S., & Kunina-Habenicht, O. (2020). Analysis of the stability and variability of motivational beliefs of teacher candidates during their long-term internship. *Zeitschrift für Erziehungswissenschaft*, *23*(2), 367–391. https://doi.org/10.1007/s11618-020-00940-2

Bautista, N. U., & Boone, W. J. (2015). Exploring the impact of TeachME™ lab virtual classroom teaching simulation on early childhood education majors' self-efficacy beliefs. *Journal of Science Teacher Education*, *26*(3), 237–262. https://doi.org/10.1007/s10972-014-9418-8

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, *41*(2), 257–278. https://doi.org/10.1111/j.2044-8317.1988.tb00901.x

Belur, J., Tompson, L., Thornton, A., & Simon, M. (2021). Interrater reliability in systematic review methodology: Exploring variation in coder decision-making. *Sociological Methods & Research*, *50*(2), 837–865. https://doi.org/10.1177/0049124118799372

Betoret, F. D. (2006). Stressors, self-efficacy, coping resources, and burnout among secondary school teachers in Spain. *Educational Psychology*, *26*(4), 519–539. https://doi.org/10.1080/01443410500342492

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009/2021). *Introduction to meta-analysis* (2nd ed.). John Wiley & Sons. https://doi.org/10.1002/9780470743386

Bosch, C., & Ellis, T. (2021). Using avatars to address teacher self-efficacy. *Journal of Global Education and Research*, *5*(1), 15–35. https://doi.org/10.5038/2577-509X.5.1.1069

Bowlin, T. M., Bell, S. M., Coleman, M. B., & Cihak, D. F. (2015). Co-teaching and collaboration: Preservice teachers' knowledge, attitudes, and perceived sense of efficacy in teaching students with disabilities. *Journal of the American Academy of Special Education Professionals*, 37–57.

Braaksma, M. A. H., Rijlaarsdam, G., & van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology*, *94*(2), 405–415. https://doi.org/10.1037/0022-0663.94.2.405

Brand, B. R., & Wilkins, J. L. M. (2007). Using self-efficacy as a construct for evaluating science and mathematics methods courses. *Journal of Science Teacher Education*, *18*(2), 297–317. https://doi.org/10.1007/s10972-007-9038-7

Burić, I., Slišković, A., & Sorić, I. (2020). Teachers' emotions and self-efficacy: A test of reciprocal relations. *Frontiers in Psychology*, *11*, Article 1650. https://doi.org/10.3389/fpsyg.2020.01650

Cabaroglu, N. (2014). Professional development through action research: Impact on self-efficacy. *System*, *44*, 79–88. https://doi.org/10.1016/j.system.2014.03.003

Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, *44*(6), 473–490. https://doi.org/10.1016/j.jsp.2006.09.001

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science*, *2*(2), 115–144. https://doi.org/10.1177/2515245919847196

Çelebi, C., Krahé, B., & Spörer, N. (2014). Gestärkt in den Lehrerberuf: Eine Förderung berufsbezogener Kompetenzen von Lehramtsstudierenden [Strengthened for the teaching profession: Promoting professional competencies in teacher training students]. *Zeitschrift für Pädagogische Psychologie*, *28*(3), 115–126. https://doi.org/10.1024/1010-0652/a000128

Chan, D. W. (2008). General, collective, and domain-specific teacher self-efficacy among Chinese prospective and in-service teachers in Hong Kong. *Teaching and Teacher Education*, *24*(4), 1057–1069. https://doi.org/10.1016/j.tate.2007.11.010

Chan, S., Maneewan, S., & Koul, R. (2021). Cooperative learning in teacher education: Its effects on EFL pre-service teachers' content knowledge and teaching self-efficacy. *Journal of Education for Teaching*, *47*(5), 654–667. https://doi.org/10.1080/02607476.2021.1931060

Chao, C. N. G., Chow, W. S. E., Forlin, C., & Ho, F. C. (2017). Improving teachers' self-efficacy in applying teaching and learning strategies and classroom management to students with special education needs in Hong Kong. *Teaching and Teacher Education*, *66*, 360–369. https://doi.org/10.1016/j.tate.2017.05.004

Chen, R. J.-C., Lin, H.-C., Hsueh, Y.-L., & Hsieh, C.-C. (2020). Which is more influential on teaching practice, classroom management efficacy or instruction efficacy? Evidence from TALIS 2018. *Asia Pacific Education Review*, *21*(4), 589–599. https://doi.org/10.1007/s12564-020-09656-8

Chesnut, S. R., & Burley, H. (2015). Self-efficacy as a predictor of commitment to the teaching profession: A meta-analysis. *Educational Research Review*, *15*, 1–16. https://doi.org/10.1016/j.edurev.2015.02.001

Cheung, A. C. K., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, *45*(5), 283–292. https://doi.org/10.3102/0013189X16656615

Conroy, M. A., Sutherland, K. S., Algina, J., Ladwig, C., Werch, B., Martinez, J., Jessee, G., & Gyure, M. (2019). Outcomes of the BEST in CLASS intervention on teachers' use of effective practices, self-efficacy, and classroom quality. *School Psychology Review*, *48*(1), 31–45. https://doi.org/10.17105/SPR-2018-0003.V48-1

Dalioglu, S. T., & Adiguzel, O. C. (2016). Teacher candidates' self-efficacy beliefs and possible selves throughout the teaching practice period in Turkey. *Asia Pacific Education Review*, *17*(4), 651–661. https://doi.org/10.1007/s12564-016-9458-1

de Carvalho, J. S., Oliveira, S., Roberto, M. S., Gonçalves, C., Bárbara, J. M., Castro, A. F., de Pereira, R., Franco, M., Cadima, J., Leal, T., Lemos, M. S., & Marques-Pinto, A. (2021). Effects of a mindfulness-based intervention for teachers: A study on teacher and student outcomes. *Mindfulness*, *12*(7), 1719–1732. https://doi.org/10.1007/s12671-021-01635-3

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181–199. https://doi.org/10.3102/0013189X08331140

Dicke, T., Elling, J., Schmeck, A., & Leutner, D. (2015). Reducing reality shock: The effects of classroom management skills training on beginning teachers. *Teaching and Teacher Education*, *48*, 1–12. https://doi.org/10.1016/j.tate.2015.01.013

Dicke, T., Parker, P. D., Holzberger, D., Kunina-Habenicht, O., Kunter, M., & Leutner, D. (2015). Beginning teachers' efficacy and emotional exhaustion: Latent changes, reciprocity, and the influence of professional knowledge. *Contemporary Educational Psychology*, *41*, 62–72. https://doi.org/10.1016/j.cedpsych.2014.11.003

Dicke, T., Stebner, F., Linninger, C., Kunter, M., & Leutner, D. (2018). A longitudinal study of teachers' occupational well-being: Applying the job demands-resources model. *Journal of Occupational Health Psychology*, *23*(2), 262–277. https://doi.org/10.1037/ocp0000070

Dixon, F. A., Yssel, N., McConnell, J. M., & Hardin, T. (2014). Differentiated instruction, professional development, and teacher efficacy. *Journal for the Education of the Gifted*, *37*(2), 111–127. https://doi.org/10.1177/0162353214529042

Duffin, L. C., French, B. F., & Patrick, H. (2012). The Teachers' Sense of Efficacy Scale: Confirming the factor structure with beginning pre-service teachers. *Teaching and Teacher Education*, *28*(6), 827–834. https://doi.org/10.1016/j.tate.2012.03.004

Dunst, C. J., Bruder, M. B., & Hamby, D. W. (2015). Metasynthesis of in-service professional development research: Features associated with positive educator and student outcomes. *Educational Research and Reviews*, *10*(12), 1731–1744. https://doi.org/10.5897/ERR2015.2306

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

European Commission/EACEA/Eurydice. (2021). *Teachers in Europe: Careers, development and well-being*. Education, Audiovisual and Culture Executive Agency (EACEA, Education and Youth Policy Analysis). https://eurydice.ba/wp-content/uploads/2021/03/Teachers-in-Europe-Careers-Development-and-Well-being.pdf

Fackler, S., Malmberg, L.-E., & Sammons, P. (2021). An international perspective on teacher self-efficacy: Personal, structural and environmental factors. *Teaching and Teacher Education*, *99*, Article 103255. https://doi.org/10.1016/j.tate.2020.103255

Fauth, B., Decristan, J., Decker, A.-T., Büttner, G., Hardy, I., Klieme, E., & Kunter, M. (2019). The effects of teacher competence on student outcomes in elementary science education: The mediating role of teaching quality. *Teaching and Teacher Education*, *86*, Article 102882. https://doi.org/10.1016/j.tate.2019.102882

Fernet, C., Guay, F., Senécal, C., & Austin, S. (2012). Predicting intraindividual changes in teacher burnout: The role of perceived school environment and motivational factors. *Teaching and Teacher Education*, *28*(4), 514–525. https://doi.org/10.1016/j.tate.2011.11.013

Fives, H., Hamman, D., & Olivarez, A. (2007). Does burnout begin with student-teaching? Analyzing efficacy, burnout, and support during the student-teaching semester. *Teaching and Teacher Education*, *23*(6), 916–934. https://doi.org/10.1016/j.tate.2006.03.013

Gaudreau, N., Royer, E., Frenette, E., Beaumont, C., & Flanagan, T. (2013). Classroom behaviour management: The effects of in-service training on elementary teachers' self-efficacy beliefs. *McGill Journal of Education*, *48*(2), 359–382.

Gegenfurtner, A., Veermans, K., & Vauras, M. (2013). Effects of computer support, collaboration, and time lag on performance self-efficacy and transfer of training: A longitudinal meta-analysis. *Educational Research Review*, *8*, 75–89. https://doi.org/10.1016/j.edurev.2012.04.001

George, S. V., Richardson, P. W., & Watt, H. M. G. (2018). Early career teachers' self-efficacy: A longitudinal study from Australia. *Australian Journal of Education*, *62*(2), 217–233. https://doi.org/10.1177/0004944118779601

Gesel, S. A., LeJeune, L. M., Chow, J. C., Sinclair, A. C., & Lemons, C. J. (2021). A meta-analysis of the impact of professional development on teachers' knowledge, skill, and self-efficacy in data-based decision-making. *Journal of Learning Disabilities*, *54*(4), 269–283. https://doi.org/10.1177/0022219420970196

Gold, B., Hellermann, C., & Holodynski, M. (2017). Effekte videobasierter Trainings zur Förderung der Selbstwirksamkeitsüberzeugungen über Klassenführung im Grundschulunterricht [Effects of video-based trainings for promoting self-efficacy in elementary classroom management]. *Zeitschrift für Erziehungswissenschaft*, *20*(1), 115–136. https://doi.org/10.1007/s11618-017-0727-5

Gresko, J. A. (2013). *Supporting first year alternatively certified urban and rural intern teachers through a multicomponent distance induction program* (Publication No. 3558741) [Doctoral dissertation, Arizona State University]. ProQuest Dissertations and Theses Global.

Griffin, J. W. (2021). Calculating statistical power for meta-analysis using metapower. *Quantitative Methods for Psychology*, *17*(1), 24–39. https://doi.org/10.20982/tqmp.17.1.p024

Groeschner, A., Schindler, A.-K., Holzberger, D., Alles, M., & Seidel, T. (2018). How systematic video reflection in teacher professional development regarding classroom discourse contributes to teacher and student self-efficacy. *International Journal of Educational Research*, *90*, 223–233. https://doi.org/10.1016/j.ijer.2018.02.003

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, *111*(9), 2055–2100. https://doi.org/10.1177/016146810911100905

Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. *Teachers and Teaching*, *15*(2), 273–289. https://doi.org/10.1080/13540600902875340

Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, *11*(2), 181–217. https://doi.org/10.1002/jrsm.1378

Haddaway, N. R., Grainger, M. J., & Gray, C. T. (2021). *citationchaser: An R package and Shiny app for forward and backward citations chasing in academic searching* [Computer software]. Zenodo.

Han, S., Shanafelt, T. D., Sinsky, C. A., Awad, K. M., Dyrbye, L. N., Fiscus, L. C., Trockel, M., & Goh, J. (2019). Estimating the attributable cost of physician burnout in the United States. *Annals of Internal Medicine*, *170*(11), 784–790. https://doi.org/10.7326/M18-1422

Harris, D. N. (2009). Toward policy-relevant benchmarks for interpreting effect sizes: Combining effects with costs. *Educational Evaluation and Policy Analysis*, *31*(1), 3–29. https://doi.org/10.3102/0162373708327524

Hartl, A., & Holzberger, D. (2022). Identifying teachers' motivational profiles and their changes from teacher education into practice: A longitudinal study. *Zeitschrift für Erziehungswissenschaft*, *25*(2), 427–451. https://doi.org/10.1007/s11618-022-01093-0

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Haverback, H. R., & Parault, S. J. (2008). Pre-service reading teacher efficacy and tutoring: A review. *Educational Psychology Review*, *20*(3), 237–255. https://doi.org/10.1007/s10648-008-9077-4

Hayes, R., Titheradge, D., Allen, K., Allwood, M., Byford, S., Edwards, V., Hansford, L., Longdon, B., Norman, S., Norwich, B., Russell, A. E., Price, A., Ukoumunne, O. C., & Ford, T. (2020). The Incredible Years® Teacher Classroom Management programme and its impact on teachers' professional self-efficacy, work-related stress, and general well-being: Results from the STARS randomized controlled trial. *British Journal of Educational Psychology*, *90*(2), 330–348. https://doi.org/10.1111/bjep.12284

Holdaway, A. S. (2017). *Evaluating the effects of a formative classroom management coaching program for pre-service teachers* [Doctoral dissertation, Ohio University]. OhioLINK Electronic Theses & Dissertations Center. http://rave.ohiolink.edu/etdc/view?acc_num=ohiou1500565483566359

Holzberger, D., Philipp, A., & Kunter, M. (2013). How teachers' self-efficacy is related to instructional quality: A longitudinal analysis. *Journal of Educational Psychology*, *105*(3), 774–786. https://doi.org/10.1037/a0032198

Holzberger, D., Philipp, A., & Kunter, M. (2014). Predicting teachers' instructional behaviors: The interplay between self-efficacy and intrinsic needs. *Contemporary Educational Psychology*, *39*(2), 100–111. https://doi.org/10.1016/j.cedpsych.2014.02.001

Hoogendijk, C., Tick, N. T., Hofman, W. H. A., Holland, J. G., Severiens, S. E., Vuijk, P., & van Veen, A. F. D. (2018). Direct and indirect effects of Key2Teach on teachers' sense of self-efficacy and emotional exhaustion, a randomized controlled trial. *Teaching and Teacher Education*, *76*, 1–13. https://doi.org/10.1016/j.tate.2018.07.014

Huang, X. (2017). Example-based learning: Effects of different types of examples on student performance, cognitive load and self-efficacy in a statistical learning task. *Interactive Learning Environments*, *25*(3), 283–294. https://doi.org/10.1080/10494820.2015.1121154

Iancu, A. E., Rusu, A., Măroiu, C., Păcurar, R., & Maricuțoiu, L. P. (2018). The effectiveness of interventions aimed at reducing teacher burnout: A meta-analysis. *Educational Psychology Review*, *30*(2), 373–396. https://doi.org/10.1007/s10648-017-9420-8

Junqueira, K. E., & Matoti, S. N. (2013). A comparative study of pre-service teachers' teaching efficacy beliefs before and after work-integrated learning: Part two. *Africa Education Review*, *10*(Suppl. 1). https://doi.org/10.1080/18146627.2013.855423

Karalar, H., & Altan, B. A. (2018). New technologies in microteachings: Is it possible for web 2.0 tools to affect prospective teachers' web pedagogical content knowledge and self-efficacy? *International Online Journal of Education and Teaching*, *5*(3), 535–551.

Kayapinar, U. (2016). A study on reflection in in-service teacher development: Introducing reflective practitioner development model. *Educational Sciences: Theory and Practice*, *16*(5), 1671–1691. https://doi.org/10.12738/estp.2016.5.0077

Kelleci, O., Kulaksiz, T., & Pala, F. K. (2018). The effect of social network-supported microlearning on teachers' self-efficacy and teaching skills. *World Journal on Educational Technology: Current Issues*, *10*(2), 115–130.

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, *86*(4), 945–980. https://doi.org/10.3102/0034654315626800

Kim, K. R., & Seo, E. H. (2018). The relationship between teacher efficacy and students' academic achievement: A meta-analysis. *Social Behavior and Personality: An International Journal*, *46*(4), 529–540. https://doi.org/10.2224/sbp.6554

Kissau, S., & Algozzine, B. (2015). The impact of mode of instructional delivery on second language teacher self-efficacy. *ReCALL*, *27*(2), 239–256. https://doi.org/10.1017/S0958344014000391

Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology*, *102*(3), 741–756. https://doi.org/10.1037/a0019237

Klassen, R. M., & Chiu, M. M. (2011). The occupational commitment and intention to quit of practicing and pre-service teachers: Influence of self-efficacy, job stress, and teaching context. *Contemporary Educational Psychology*, *36*(2), 114–129. https://doi.org/10.1016/j.cedpsych.2011.01.002

Klassen, R. M., & Durksen, T. L. (2014). Weekly self-efficacy and work stress during the teaching practicum: A mixed methods study. *Learning and Instruction*, *33*, 158–169. https://doi.org/10.1016/j.learninstruc.2014.05.003

Klassen, R. M., & Kim, L. E. (2019). Selecting teachers and prospective teachers: A meta-analysis. *Educational Research Review*, *26*, 32–51. https://doi.org/10.1016/j.edurev.2018.12.003

Klassen, R. M., Rushby, J. V., Maxwell, L., Durksen, T. L., Sheridan, L., & Bardach, L. (2021). The development and testing of an online scenario-based learning activity to prepare preservice teachers for teaching placements. *Teaching and Teacher Education*, *104*, Article 103385. https://doi.org/10.1016/j.tate.2021.103385

Klassen, R. M., & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review*, *12*, 59–76. https://doi.org/10.1016/j.edurev.2014.06.001

Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998–2009: Signs of progress or unfulfilled promise? *Educational Psychology Review*, *23*(1), 21–43. https://doi.org/10.1007/s10648-010-9141-8

Knight, D. S. (2012). Assessing the cost of instructional coaching. *Journal of Education Finance*, *38*(1), 52–80.

Knoblauch, D., & Chase, M. A. (2015). Rural, suburban, and urban schools: The impact of school setting on the efficacy beliefs and attributions of student teachers. *Teaching and Teacher Education*, *45*, 104–114. https://doi.org/10.1016/j.tate.2014.10.001

Knobloch, N. A. (2006). Exploring relationships of teachers' sense of efficacy in two student teaching programs. *Journal of Agricultural Education*, *47*(2), 36–47. https://doi.org/10.5032/jae.2006.02036

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189X20912798

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, *88*(4), 547–588. https://doi.org/10.3102/0034654318759268

Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A.-M. K., Hammerstrøm, K., & Sathe, N. (2017). Searching for studies: A guide to information retrieval for Campbell systematic reviews. *Campbell Systematic Reviews*, *13*(1), 1–73. https://doi.org/10.4073/cmg.2016.1

Kumschick, I. R., Piwowar, V., Ophardt, D., Barth, V. L., Krysmanski, K., & Thiel, F. (2017). Optimierung einer videobasierten Lerngelegenheit im Problem Based Learning Format durch Cognitive Tools. Eine Interventionsstudie mit Lehramtsstudierenden [Optimizing a video-based learning environment in a problem based learning format using cognitive tools: An intervention study with preservice teachers]. *Zeitschrift für Erziehungswissenschaft*, *20*(1), 93–113. https://doi.org/10.1007/s11618-017-0728-4

Künsting, J., Neuber, V., & Lipowsky, F. (2016). Teacher self-efficacy as a long-term predictor of instructional quality in the classroom. *European Journal of Psychology of Education*, *31*(3), 299–322. https://doi.org/10.1007/s10212-015-0272-7

Kunz, H. D., Sandmeier, A., Hänggi, Y., Safi, N., & Cina, A. (2019). Training zum Umgang mit Unterrichtsstörungen: Effekte auf die Gesundheit von Lehrkräften [Training to deal with classroom disruptions: Effects on teacher health]. *Zeitschrift für Erziehungswissenschaft*, *22*(4), 925–944. https://doi.org/10.1007/s11618-019-00900-5

Lauermann, F., & König, J. (2016). Teachers' professional competence and wellbeing: Understanding the links between general pedagogical knowledge, self-efficacy and burnout. *Learning and Instruction*, *45*, 9–19. https://doi.org/10.1016/j.learninstruc.2016.06.006

Learning Policy Institute. (2017, September 13). *What's the cost of teacher turnover?* https://learningpolicyinstitute.org/product/the-cost-of-teacher-turnover

Lefebvre, C., Glanville, J., Briscoe, S., Littlewood, A., Marshall, C., Metzendorf, M.-I., Noel-Storr, A., Rader, T., Shokraneh, F., Thomas, J., & Wieland, L. (2021). Chapter 4: Searching for and selecting studies. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane handbook for systematic reviews of interventions* (Version 6.1, updated Sept. 2020). Cochrane. www.training.cochrane.org/handbook

Levin, H. M., & Belfield, C. (2015). Guiding the development and use of cost-effectiveness analysis in education. *Journal of Research on Educational Effectiveness*, *8*(3), 400–418. https://doi.org/10.1080/19345747.2014.915604

Liaw, E.-C. (2017). Application of computer-mediated communication on teacher efficacy during school-based field experience. *Teaching and Teacher Education*, *65*, 81–90. https://doi.org/10.1016/j.tate.2017.03.002

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

McCullough, S. N., Granger, K. L., Sutherland, K. S., Conroy, M. A., & Pandey, T. (2021). A preliminary study of BEST in CLASS-elementary on teacher self-efficacy, burnout, and attributions. *Behavioral Disorders*, *47*(2), 84–94. https://doi.org/10.1177/01987429211010672

Mergler, A. G., & Tangen, D. (2010). Using microteaching to enhance teacher efficacy in pre-service teachers. *Teaching Education*, *21*(2), 199–210. https://doi.org/10.1080/10476210902998466

Michigan Department of Education. (n.d.). *Core teaching practices*. https://www.michigan.gov/-/media/Project/Websites/mde/educator_services/prep/core_teaching_practices.pdf?rev=13216382a9a949c293262f9c5e0fad99

Michos, K., Cantieni, A., Schmid, R., Müller, L., & Petko, D. (2022). Examining the relationship between internship experiences, teaching enthusiasm, and teacher self-efficacy when using a mobile portfolio app. *Teaching and Teacher Education*, *109*, Article 103570. https://doi.org/10.1016/j.tate.2021.103570

Mok, S. Y., Rupp, D., & Holzberger, D. (2023). What kind of individual support activities in interventions foster pre-service and beginning teachers' self-efficacy? A meta-analysis. *Educational Research Review*, *100552*. https://doi.org/10.1016/j.edurev.2023.100552

Mok, S. Y., & Staub, F. C. (2021). Does coaching, mentoring, and supervision matter for pre-service teachers' planning skills and clarity of instruction? A meta-analysis of (quasi-)experimental studies. *Teaching and Teacher Education*, *107*, Article 103484. https://doi.org/10.1016/j.tate.2021.103484

Morris, D. B., Usher, E. L., & Chen, J. A. (2017). Reconceptualizing the sources of teaching self-efficacy: A critical review of emerging literature. *Educational Psychology Review*, *29*(4), 795–833. https://doi.org/10.1007/s10648-016-9378-y

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*(1), 105–125. https://doi.org/10.1037//1082-989X.7.1.105

Moulding, L. R., Stewart, P. W., & Dunmeyer, M. L. (*2014*). Pre-service teachers' sense of efficacy: Relationship to academic ability, student teaching placement characteristics, and mentor support. *Teaching and Teacher Education*, *41*, 60–66. https://doi.org/10.1016/j.tate.2014.03.007

Neve, D., de Devos, G., & Tuytens, M. (2015). The importance of job resources and self-efficacy for beginning teachers' professional learning in differentiated instruction. *Teaching and Teacher Education*, *47*, 30–41. https://doi.org/10.1016/j.tate.2014.12.003

Oliveira, S., Roberto, M. S., Pereira, N. S., Marques-Pinto, A., & Veiga-Simão, A. M. (2021). Impacts of social and emotional learning interventions for teachers on teachers' outcomes: A systematic review with meta-analysis. *Frontiers in Psychology*, *12*, Article 677217. https://doi.org/10.3389/fpsyg.2021.677217

O'Neill, S., & Stephenson, J. (2012). Exploring Australian pre-service teachers' sense of efficacy, its sources, and some possible influences. *Teaching and Teacher Education*, *28*(4), 535–545. https://doi.org/10.1016/j.tate.2012.01.008

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . .Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, *10*(1), Article 89. https://doi.org/10.1186/s13643-021-01626-4

Palmer, D. H. (2011). Sources of efficacy information in an in-service program for elementary teachers. *Science Education*, *95*(4), 577–600. https://doi.org/10.1002/sce.20434

Peebles, J. L., & Mendaglio, S. (2014). The impact of direct experience on preservice teachers' self-efficacy for teaching in inclusive classrooms. *International Journal of Inclusive Education*, *18*(12), 1321–1336. https://doi.org/10.1080/13603116.2014.899635

Pekrun, R. (2023). Mind and body in students' and teachers' engagement: New evidence, challenges, and guidelines for future research. *British Journal of Educational Psychology*. Advance online publication. https://doi.org/10.1111/bjep.12575

Pendergast, D., Garvis, S., & Keogh, J. (2011). Pre-service student-teacher self-efficacy beliefs: An insight into the making of teachers. *Australian Journal of Teacher Education*, *36*(12). https://doi.org/10.14221/ajte.2011v36n12.6

Pfitzner-Eden, F. (2016a). I feel less confident so I quit? Do true changes in teacher self-efficacy predict changes in preservice teachers' intention to quit their teaching degree? *Teaching and Teacher Education*, *55*, 240–254. https://doi.org/10.1016/j.tate.2016.01.018

Pfitzner-Eden, F. (2016b). Why do I feel more confident? Bandura's sources predict preservice teachers' latent changes in teacher self-efficacy. *Frontiers in Psychology*, *7*, Article 1486. https://doi.org/10.3389/fpsyg.2016.01486

Phan, N. T. T., & Locke, T. (2015). Sources of self-efficacy of Vietnamese EFL teachers: A qualitative study. *Teaching and Teacher Education*, *52*, 73–82. https://doi.org/10.1016/j.tate.2015.09.006

Pigott, T. D., & Polanin, J. R. (2019). Methodological guidance paper: high-quality meta-analysis in a systematic review. *Review of Educational Research*, 1–23. https://doi.org/10.3102/0034654319877153

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

Postareff, L., Lindblom-Ylänne, S., & Nevgi, A. (2007). The effect of pedagogical training on teaching in higher education. *Teaching and Teacher Education*, *23*(5), 557–571. https://doi.org/10.1016/j.tate.2006.11.013

Pressley, T., & Ha, C. (2021). Returning to teaching during COVID-19: An empirical study on teacher self-efficacy. *Teaching and Teacher Education*, Article 103465. https://doi.org/10.1016/j.tate.2021.103465

Prilop, C. N., Weber, K. E., & Kleinknecht, M. (2019). How digital reflection and feedback environments contribute to pre-service teachers' beliefs during a teaching practicum. *Studies in Educational Evaluation*, *62*, 158–170. https://doi.org/10.1016/j.stueduc.2019.06.005

Prilop, C. N., Weber, K. E., Prins, F. J., & Kleinknecht, M. (2021). Connecting feedback to self-efficacy: Receiving and providing peer feedback in teacher education. *Studies in Educational Evaluation*, *70*, Article 101062. https://doi.org/10.1016/j.stueduc.2021.101062

Pustejovsky, J. E. (2022). *clubSandwich: Cluster-Robust (Sandwich) Variance Estimators with Small-Sample Corrections* [Computer software]. https://cran.r-project.org/web/packages/clubSandwich/clubSandwich.pdf

Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with robust variance estimation: Expanding the range of working models. *Prevention Science*, *23*(3), 425–438. https://doi.org/10.1007/s11121-021-01246-3

Ross, J. A., & Bruce, C. D. (2007). Professional development effects on teacher efficacy: Results of randomized field trial. *Journal of Educational Research*, *101*(1), 50–60.

Rupp, D., & Becker, E. S. (2021). Situational fluctuations in student teachers' self-efficacy and its relation to perceived teaching experiences and cooperating teachers' discourse elements during the teaching practicum. *Teaching and Teacher Education*, *99*, Article 103252. https://doi.org/10.1016/j.tate.2020.103252

Ryan, A. M., Kuusinen, C. M., & Bedoya-Skoog, A. (2015). Managing peer relations: A dimension of teacher self-efficacy that varies between elementary and middle school teachers and is associated with observed classroom quality. *Contemporary Educational Psychology*, *41*, 147–156. https://doi.org/10.1016/j.cedpsych.2015.01.002

Savolainen, H., Malinen, O. P., & Schwab, S. (2022). Teacher efficacy predicts teachers' attitudes towards inclusion – a longitudinal cross-lagged analysis. *International Journal of Inclusive Education*, *26*(9), 958–972. https://doi.org/10.1080/13603116.2020.1752826

Schipper, T. M., de Vries, S., Goei, S. L., & van Veen, K. (2020). Promoting a professional school culture through lesson study? An examination of school culture, school conditions, and teacher self-efficacy. *Professional Development in Education*, *46*(1), 112–129. https://doi.org/10.1080/19415257.2019.1634627

Schipper, T. M., Goei, S. L., de Vries, S., & van Veen, K. (2018). Developing teachers' self-efficacy and adaptive teaching behaviour through lesson study. *International Journal of Educational Research*, *88*, 109–120. https://doi.org/10.1016/j.ijer.2018.01.011

Schleicher, A. (2018). *Valuing our teachers and raising their status*. OECD. https://doi.org/10.1787/9789264292697-en

Schüle, C., Besa, K.-S., Schriek, J., & Arnold, K.-H. (2017). Die Veränderung der Lehrerselbstwirksamkeitsüberzeugung in Schulpraktika [The development of student teacher self-efficacy in student teaching field experiences]. *Zeitschrift für Bildungsforschung*, *7*(1), 23–40. https://doi.org/10.1007/s35834-016-0177-9

Schulte, K. (2008). *Selbstwirksamkeitserwartungen in der Lehrerbildung – Zur Struktur und dem Zusammenhang von Lehrer-Selbstwirksamkeitserwartungen, Pädagogischem Professionswissen und Persönlichkeitseigenschaften bei Lehramtsstudierenden und Lehrkräften [Self-efficacy beliefs in teacher education - About the structure and relationship of self-efficacy beliefs, general pedagogical knowledge and personality in student-teachers and teachers]* [Unpublished doctoral dissertation]. Georg-August-Universität, Göttingen.

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research*, *57*(2), 149–174. https://doi.org/10.2307/1170234

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, *26*(3–4), 207–231. https://doi.org/10.1080/00461520.1991.9653133

Schwarzer, R., & Warner, L. (2011). Forschung zur Selbstwirksamkeit bei Lehrerinnen und Lehrern [Research on self-efficacy among teachers]. In E. Terhart, H. Bennewitz, & M. Rothland (Eds.), *Handbuch der Forschung zum Lehrerberuf* (pp. 496–510). Waxmann.

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

Shanafelt, T., Goh, J., & Sinsky, C. (2017). The business case for investing in physician well-being. *JAMA Internal Medicine*, *177*(12), 1826–1832. https://doi.org/10.1001/jamainternmed.2017.4340

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, *70*(1), 747–770. https://doi.org/10.1146/annurev-psych-010418-102803

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology*, *99*(3), 611–625. https://doi.org/10.1037/0022-0663.99.3.611

Skaalvik, E. M., & Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education*, *26*(4), 1059–1069. https://doi.org/10.1016/j.tate.2009.11.001

Skaalvik, E. M., & Skaalvik, S. (2014). Teacher self-efficacy and perceived autonomy: Relations with teacher engagement, job satisfaction, and emotional exhaustion. *Psychological Reports*, *114*(1), 68–77. https://doi.org/10.2466/14.02.PR0.114k14w0

Smith, E. C., Starratt, G. K., McCrink, C. L., & Whitford, H. (2020). Teacher evaluation feedback and instructional practice self-efficacy in secondary school teachers. *Educational Administration Quarterly*, *56*(4), 671–701. https://doi.org/10.1177/0013161X19888568

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*(1), 60–78. https://doi.org/10.1002/jrsm.1095

Stephanou, G., Gkavras, G., & Doulkeridou, M. (2013). The role of teachers' self- and collective-efficacy beliefs on their job satisfaction and experienced emotions in school. *Psychology*, *4*(3A), 268–278. https://doi.org/10.4236/psych.2013.43A040

Suprayogi, M. N., Valcke, M., & Godwin, R. (2017). Teachers and their implementation of differentiated instruction in the classroom. *Teaching and Teacher Education*, *67*, 291–301. https://doi.org/10.1016/j.tate.2017.06.020

Swan, B., Wolf, K., & Cano, J. (2011). Changes in teacher self–efficacy from the student teaching experience through the third year of teaching. *Journal of Agricultural Education*, *52*(2), 128–139. https://doi.org/10.5032/jae.2011.02128

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: A tutorial in R. *Journal of Developmental and Life-Course Criminology*, *2*(1), 85–112. https://doi.org/10.1007/s40865-016-0026-5

TeachingWorks. (n.d.). *High-leverage practices*. Retrieved July 6, 2023, from https://www.teachingworks.org/high-leverage-practices/

Thiel, F., Böhnke, A., Barth, V. L., & Ophardt, D. (2020). How to prepare preservice teachers to deal with disruptions in the classroom? Differential effects of learning

with functional and dysfunctional video scenarios. *Professional Development in Education*, *49*(1), 108–122. https://doi.org/10.1080/19415257.2020.1763433

Thurlings, M., Evers, A. T., & Vermeulen, M. (2015). Toward a model of explaining teachers' innovative behavior. *Review of Educational Research*, *85*(3), 430–471. https://doi.org/10.3102/0034654314557949

Thurm, D., & Barzel, B. (2020). Effects of a professional development program for teaching mathematics with technology on teachers' beliefs, self-efficacy and practices. *ZDM*, *52*(7), 1411–1422. https://doi.org/10.1007/s11858-020-01158-6

Tipton, E., & Pustejovsky, J. E. (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *Journal of Educational and Behavioral Statistics*, *40*(6), 604–634. https://doi.org/10.3102/1076998615606099

Toropova, A., Myrberg, E., & Johansson, S. (2020). Teacher job satisfaction: The importance of school working conditions and teacher characteristics. *Educational Review*, *73*(1), 71–97. https://doi.org/10.1080/00131911.2019.1705247

Tschannen-Moran, M., & McMaster, P. (2009). Sources of self-efficacy: Four professional development formats and their relationship to self-efficacy and implementation of a new teaching strategy. *Elementary School Journal*, *110*(2), 228–245. https://doi.org/10.1086/605771

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783–805. https://doi.org/10.1016/S0742-051X(01)00036-1

Tschannen-Moran, M., & Woolfolk Hoy, A. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and Teacher Education*, *23*(6), 944–956. https://doi.org/10.1016/j.tate.2006.05.003

Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, *68*(2), 202–248. https://doi.org/10.3102/00346543068002202

USA Facts. (n.d.). *Average public school teacher salary*. Retrieved March 7, 2023, from https://usafacts.org/data/topics/people-society/education/k-12-education/public-school-teacher-salary-average/#:~:text=The%20average%20public%20school%20teacher,1.49%25%20or%20%24957%20from%202020

van der Scheer, E. A., & Visscher, A. J. (2016). Effects of an intensive data-based decision making intervention on teacher efficacy. *Teaching and Teacher Education*, *60*, 34–43. https://doi.org/10.1016/j.tate.2016.07.025

Vansteelandt, I., Mol, S. E., Vanderlinde, R., Lerkkanen, M.-K., & van Keer, H. (2020). In pursuit of beginning teachers' competence in promoting reading motivation: A mixed-methods study into the impact of a continuing professional development program. *Teaching and Teacher Education*, *96*, Article 103154. https://doi.org/10.1016/j.tate.2020.103154

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W. (2020). Model checking in meta-analysis. In C. H. Schmid, T. Stijnen, & I. R. White (Eds.), *Handbook of meta-analysis* (pp. 219–254). Chapman and Hall/CRC. https://doi.org/10.1201/9781315119403-11

Viechtbauer, W. (n.d.). *General workflow for meta-analyses involving complex dependency structures*. Retrieved November 25, 2022, from https://wviechtb.github.io/metafor/reference/misc-recs.html#general-workflow-for-meta-analyses-involving-complex-dependency-structures

Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*(2), 112–125. https://doi.org/10.1002/jrsm.11

von Suchodoletz, A., Jamil, F. M., Larsen, R. A., & Hamre, B. K. (2018). Personal and contextual factors associated with growth in preschool teachers' self-efficacy beliefs during a longitudinal professional development study. *Teaching and Teacher Education*, *75*, 278–289. https://doi.org/10.1016/j.tate.2018.07.009

Wang, H., Hall, N. C., & Rahimi, S. (2015). Self-efficacy and causal attributions in teachers: Effects on burnout, job satisfaction, illness, and quitting intentions. *Teaching and Teacher Education*, *47*, 120–130. https://doi.org/10.1016/j.tate.2014.12.005

Weber, K. E., Gold, B., Prilop, C. N., & Kleinknecht, M. (2018). Promoting pre-service teachers' professional vision of classroom management during practical school training: Effects of a structured online- and video-based self-reflection and feedback intervention. *Teaching and Teacher Education*, *76*, 39–49. https://doi.org/10.1016/j.tate.2018.08.008

Weber, K. E., Prilop, C. N., & Kleinknecht, M. (2019). Effects of blended and video-based coaching approaches on preservice teachers' self-efficacy and perceived competence support. *Learning Culture and Social Interaction*, *22*, Article 100318. https://doi.org/10.1016/j.lcsi.2019.100318

Wedderhoff, N., & Bosnjak, M. (2020). Erfassung der Primärstudienqualität in psychologischen Meta-Analysen. Eine systematische Übersichtsarbeit [Assessment of study quality in psychological meta-analyses: A systematic review]. *Psychologische Rundschau*, *71*(2), 119–126. https://doi.org/10.1026/0033-3042/a000484

Weß, R., Priemer, B., Weusmann, B., Ludwig, T., Sorge, S., & Neumann, I. (2020). Der Verlauf von lehrbezogenen Selbstwirksamkeitserwartungen angehender MINT-Lehrkräfte im Studium [The development of pre-service STEM teachers' teaching-related self-efficacy in their course of study]. *Zeitschrift für Pädagogische Psychologie*, *34*(3–4), 1–18. https://doi.org/10.1024/1010-0652/a000272

What Works Clearinghouse. (2020). *What Works Clearinghouse™ standards handbook* (Version 4.1). U.S. Department of Education, Institute of Education Sciences, National Center. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf

Whitley, V. P., Park, T. D., Warner, W. J., & Horne, E. T. (2019). Relationship between career and technical education student teachers' self-efficacy and edTPA performance. *Career and Technical Education Research*, *44*(2), 88–113. https://doi.org/10.5328/cter44.2.88

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, *10*, Article 3087. https://doi.org/10.3389/fpsyg.2019.03087

Wolters, C. A., & Daugherty, S. G. (2007). Goal structures and teachers' sense of efficacy: Their relation and association to teaching experience and academic level. *Journal of Educational Psychology*, *99*(1), 181–193. https://doi.org/10.1037/0022-0663.99.1.181

Woolfolk Hoy, A., & Spero, R. B. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education*, *21*(4), 343–356. https://doi.org/10.1016/j.tate.2005.01.007

Yeo, L. S., Ang, R. P., Chong, W. H., Huan, V. S., & Quek, C. L. (2008). Teacher efficacy in the context of teaching low achieving students. *Current Psychology*, *27*(3), 192–204. https://doi.org/10.1007/s12144-008-9034-x

Yilmaz, I., & Koca, F. (2017). Investigation the effects of field-based practicum experience on pre-service physical education teachers' self-efficacy beliefs. *International Journal of Eurasia Social Sciences*, *8*(30), 1686–1695.

Yoon, K. S., Duncan, T., Wen-Yu Lee, S., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* [Issues & Answers Report]. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. https://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/rel_2007033.pdf

Zarate, K., Maggin, D. M., & Passmore, A. (2019). Meta-analysis of mindfulness training on teacher well-being. *Psychology in the Schools*, *56*(10), 1700–1715. https://doi.org/10.1002/pits.22308

Zee, M., & Koomen, H. M. Y. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, *86*(4), 981–1015. https://doi.org/10.3102/0034654315626801

Zee, M., & Koomen, H. (2020). Engaging children in the upper elementary grades: unique contributions of teacher self-efficacy, autonomy support, and student-teacher relationships. *Journal of Research in Childhood Education*, *34*(4), 477–495. https://doi.org/10.1080/02568543.2019.1701589

Authors

JANINA TÄSCHNER is a PhD candidate and research associate at the TUM School of Social Sciences and Technology, Technical University of Munich and the Centre for International Student Assessment, Germany; e-mail: *janina.taeschner@tum.de*. Her research focuses on teacher education and training with special regard to motivational characteristics of teachers.

THERESA DICKE, PhD, is an associate professor at the Institute for Positive Psychology, Australian Catholic University; e-mail: *Theresa.Dicke@acu.edu.au*. Her research focuses on the performance, motivation, and well-being of students, teachers, and school principals.

SARAH REINHOLD was a research associate at the TUM School of Social Sciences and Technology, Technical University of Munich and the Centre for International Student Assessment, Germany until 2022. Her research focused on individual and institutional influences on study choice.

DORIS HOLZBERGER is a professor for Learning and Instruction at TUM School of Social Sciences and Technology, Technical University of Munich and the Centre for International Student Assessment, Germany; e-mail: *doris.holzberger@tum.de*. Her research interests focus on characteristics of successful teachers (especially teacher motivation) and their instructional quality.