Technische Universität München
TUM School of Social Sciences and Technology

Don’t Stop Believing? Investigating the Malleability of Teacher Self-Efficacy Through Meta-Analytic and Multi-Level Methods

Janina Katharina Helga Täschner

Vollständiger Abdruck der von der TUM School of Social Sciences and Technology der Technischen Universität München zur Erlangung einer Doktorin der Philosophie (Dr. phil.) genehmigten Dissertation.

Vorsitz: Prof. Dr. Claudia Nerdel

Prüfende der Dissertation:

1. Prof. Dr. Doris Holzberger
2. Prof. Dr. Andreas Gegenfurtner
3. Prof. Dr. Anna Keune

Die Dissertation wurde am 25.08.2025 bei der Technischen Universität München eingereicht und durch die TUM School of Social Sciences and Technology am 09.10.2025 angenommen.

Für alle mutigen Frauen.

Acknowledgements

Many individuals have shaped this doctoral journey, and I would like to take the opportunity to thank those who had a special impact — academically, professionally, and personally.

My first thanks goes to Prof. Dr. Doris Holzberger. Thank you for sharing your enthusiasm for the construct of teacher self-efficacy with me and for guiding me through my first steps in the fascinating world of academia. Your intellectual stimulation, hands-on support, and trust in my ideas gave me the freedom to explore unusual paths and follow my interests. I am truly grateful!

I would also like to thank Prof. Dr. Andreas Gegenfurtner and Prof. Dr. Anna Keune for taking the time to review my dissertation, and Prof. Dr. Claudia Nerdel for chairing the examination board.

A special thanks to Prof. Dr. Marjolein Zee for her supervision and support during my research stay in Amsterdam, for sharing her insights into the Dutch academic system, and for the continuing collaboration. The time for our paper is yet to come! I am also grateful to Dr. Kristina Stockinger, my mentor, for her timely check-ins, constant support, and fresh perspectives.

I want to express my appreciation to the entire team at the Professorship of Psychology of Teaching and Learning. Working in such a dedicated, focused, and warm environment was truly a pleasure. I am especially grateful to my student research assistants Kübra, Georgios, and Linyuan; your great work in preparing the data made a real difference. And to round off the people at TUM, a heartfelt thank you to Anna, Kaley, Sara, and Sarah for sharing offices, experiences, coffees, and for being a constant source of emotional support.

The passion for well-conducted meta-analyses has connected me with some truly inspiring minds, and I am grateful to each and every one of you. A special thank you to you, Lea, for our long walks, deep talks about meta-analyses, and your last-minute feedback and support. A heartfelt thank you also to Sog Yee for many inspiring conversations, for being a role model, and for your life-coach energy. Sabine, thank you for being a fantastic collaborator in all kinds of settings — from EXIT Games and shared accommodations to co-authored papers.

I am deeply grateful to my family, especially my parents, for their unwavering belief in me. You are a source of strength to me!

Finally, my greatest thanks go to my friends, who believed in me even more than I did at times. You not only lifted my spirits through musical, sporty and culinary activities, but also filled my freezer and fridge when I needed it most. A huge thank you to the Monaco-Mädels, to Anni, Lisa, Madeleine, Sophie, the Girlies and especially to König Franz - kennt den wer? With friends like you, I could even imagine writing a second dissertation…

Abstract

A high level of teacher self-efficacy (TSE) is considered a key predictor for teacher well-being, innovative, engaging, and effective teaching. Teacher self-efficacy is defined as teachers' belief in their abilities to successfully perform teaching-specific tasks. However, while the positive outcomes of TSE are well researched, little is known about its malleability. Based on Albert Bandura's socio-cognitive theory, this dissertation examines two different change processes and, in particular, central theoretical assumptions about the roles of the four sources of self-efficacy, the career stage, and the context for its malleability.

Study 1 examines intentional change processes through interventions using a meta-analysis. The data is based on 115 intervention studies involving 11,284 pre- and in-service teachers from 26 countries. Study 1 investigates whether interventions can promote TSE and to what extent the four sources of self-efficacy, i.e., mastery experiences, vicarious experiences, social persuasion, physiological and emotional reactions, and the teachers' career stage influence the intervention effects. The results show that teachers report significantly higher levels of self-efficacy after interventions than before. Interventions that contained elements of the four sources of self-efficacy were more effective than interventions without these elements. However, no source significantly outperformed the other. Pre- and in-service teachers benefited equally from the interventions. Further analyses showed that explicit opportunities for reflection increased the intervention effects.

Study 2 examines the variation in self-efficacy across different contexts. More specifically, it analyzes the extent to which secondary school teachers, who typically teach several classes, have different levels of self-efficacy per class. It also investigates which class characteristics can predict the level of teacher self-efficacy. Data comes from 26 secondary school teachers and 1,326 students from 74 classes in the Netherlands. Intraclass correlations show clear differences in TSE between classes within teachers; up to 63% of the variance in TSE can be attributed to differences between classes. The class characteristics "misbehavior" was consistently negatively associated with the level of TSE. Moreover, class engagement and achievement were partly related to TSE, while the proportion of students with a migrant background had no influence on TSE.

Both studies provide evidence that teacher self-efficacy is a malleable construct. Except for the finding that TSE is similarly malleable through interventions in all considered career stages, the results are largely consistent with Bandura's theoretical assumptions. However, they must also be viewed in light of the limitations of the respective methodology and study design. All in all, the positive findings of Study 1 in particular call teacher educators, and trainers as well as educational administration to monitor teacher self-efficacy and to invest in support and promotion. The finding in Study 2 that teachers experience different levels of self-efficacy depending on the class they teach, underscores the need for future research to examine valid methods of measuring self-efficacy.

Zusammenfassung

Ein hohes Maß an Selbstwirksamkeit gilt als wichtiger Prädiktor für die mentale Gesundheit von Lehrkräften sowie für innovativen, motivierenden und effektiven Unterricht. Selbstwirksamkeit ist dabei definiert als die Überzeugung von Lehrkräften, mit ihren Fähigkeiten unterrichtsspezifische Aufgaben erfolgreich bewältigen zu können. Während die positiven Auswirkungen einer hohen Selbstwirksamkeit gut erforscht sind, ist über die Veränderbarkeit der Selbstwirksamkeit bislang wenig bekannt. Auf Basis der sozial-kognitiven Theorie von Albert Bandura nimmt sich diese Dissertation zwei unterschiedlichen Veränderungsprozessen an und untersucht insbesondere theoretische Annahmen zur Rolle der vier Quellen der Selbstwirksamkeit, der Karrierestufe und des Kontexts zur Veränderbarkeit.

Studie 1 untersucht gezielte Veränderungsprozesse durch Interventionen mithilfe einer Metaanalyse. Die Datenbasis bilden 115 Interventionsstudien mit 11.284 angehenden und aktiven Lehrkräften aus 26 Ländern. Studie 1 analysiert dabei, ob Interventionen die Selbstwirksamkeit von Lehrkräften gezielt fördern können und inwiefern die vier Quellen der Selbstwirksamkeit (eigene Erfolgserfahrungen, stellvertretende Erfahrungen, verbale Rückmeldungen, physiologische und emotionale Reaktionen) sowie die Karrierestufe (mit oder ohne Berufserfahrung) die Effekte der Interventionen beeinflussen. Die Ergebnisse zeigen, dass Lehrkräfte nach Interventionen eine signifikant höhere Selbstwirksamkeit angeben als davor. Interventionen, die sich den vier Quellen der Selbstwirksamkeit zuordnen ließen, waren wirksamer als Interventionen, bei denen diese Zuordnung nicht möglich war. Lehrkräfte mit und ohne Berufserfahrung profitierten gleichermaßen von den Interventionen. In weiteren Analysen erwiesen sich insbesondere explizite Gelegenheiten zur Reflexion als besonders förderlich für die Interventionseffekte.

Studie 2 untersucht die Variation der Selbstwirksamkeit über verschiedene Kontexte hinweg. Genauer gesagt wird analysiert, inwieweit Lehrkräfte an Sekundarschulen, die typischerweise mehrere Klassen unterrichten, je Klasse eine unterschiedlich hohe Selbstwirksamkeit haben und welche Merkmale der Klassen die Höhe der Selbstwirksamkeit der Lehrkraft vorhersagen können. Die Daten stammen von 26 Sekundarschullehrkräften und 1.326 Schüler:innen aus 74 Klassen in den Niederlanden. Intraklassenkorrelationen zeigen deutliche Unterschiede in der Selbstwirksamkeit der Lehrkräfte zwischen ihren verschiedenen Klassen; bis zu 63 % der Gesamtvarianz der Selbstwirksamkeit ist auf Unterschiede zwischen den Klassen zurückzuführen. Das Klassenmerkmal „Fehlverhalten“ war durchwegs negativ mit der Selbstwirksamkeit verbunden. Das Engagement und die Leistung der Klassen standen nur teilweise in Zusammenhang mit der Selbstwirksamkeit der Lehrkraft, der Anteil an Schüler:innen mit Migrationshintergrund hatte keinen Einfluss auf die Selbstwirksamkeit.

Beide Studien liefern Belege dafür, dass die Selbstwirksamkeit von Lehrkräften ein veränderbares Konstrukt ist. Die Ergebnisse stehen mit Ausnahme des Befundes, dass die

Selbstwirksamkeit von Lehrkräften in allen berücksichtigten Karrierestufen durch Interventionen veränderbar ist, größtenteils im Einklang mit den theoretischen Annahmen von Bandura. Sie müssen aber auch vor den Grenzen der jeweiligen Methodik und Studiendesigns betrachtet werden. Alles in allem rufen insbesondere die positiven Befunde von Studie 1 Personen in der Lehrkräfteaus- und weiterbildung sowie Bildungsadministration und -politik dazu auf, die Selbstwirksamkeit von Lehrkräften im Blick zu behalten und zu fördern. Der Befund von Studie 2, dass Lehrkräfte je nach Klasse eine unterschiedlich hohe Selbstwirksamkeit erleben, unterstreicht den Bedarf, in zukünftiger Forschung zu prüfen, wie Messungen der Selbstwirksamkeit diese Kontextspezifität berücksichtigen können und sollen.

Table of Contents

List of Tables

List of Figures

1 Introduction

The study of self-perception has been a central topic since the early days of psychology in the late 19th century (James, 1890; Marsh et al., 2019; Reeve, 2018). In recent decades, *self-efficacy*, which is defined as the "belief in one's capabilities to organize and execute the course of action required to manage prospective situations" (Bandura, 1995, p. 2) has attracted increased attention in self-perception research: Researchers in various fields such as entrepreneurship, education, or health care have developed scales for measuring self-efficacy (e.g., Gruber-Baldini et al., 2017; Kilday et al., 2016; McGee et al., 2009) and pointed to its value for outcomes such as entrepreneurial persistence (e.g., Cardon & Kirk, 2015), learning engagement (Yanqing Wang et al., 2022), academic performance (Honicke & Broadbent, 2016), or health-related behaviors (e.g., Collado-Mateo et al., 2021; Sheeran et al., 2016).

Self-efficacy is a construct of particular importance in educational psychology when it comes to research on teachers. Teacher self-efficacy (TSE), an individual teacher's belief to be able to successfully perform teaching-specific tasks (e.g., Holzberger et al., 2013; Skaalvik & Skaalvik, 2007; Tschannen-Moran & Woolfolk Hoy, 2001), is not only regularly featured in large longitudinal studies such as the German National Educational Panel Study (Ortenburger et al., 2023) or large-scale assessments such as the Teaching and Learning International Survey (OECD, 2019), it is also often investigated as a predictor for various beneficial outcomes (e.g., Holzberger et al., 2013; Künsting et al., 2016; Toropova et al., 2021). Meta-analyses indicate that teachers with higher levels of TSE experience less feelings of burnout (e.g., Aloe, Amo, & Shanahan, 2014), report more job commitment (e.g., Chesnut & Burley, 2015), and show better teaching quality (Klassen & Tze, 2014). However, while the beneficial effects of a high level of teacher self-efficacy have been the focus of the research in the last decades, fundamental *theoretical assumptions about the malleability* of self-efficacy remain untested in teacher self-efficacy research (Klassen et al., 2011; Lazarides & Warner, 2016).

The first assumption under investigation concerns the *sources of self-efficacy*. Bandura (1997) outlined in his socio-cognitive theory that self-efficacy is mainly shaped by information from four different kinds of experiences, the so-called four sources of self-efficacy: mastery experiences, vicarious experiences, social persuasion, and physiological and emotional reactions. Among these sources, mastery experiences should have the strongest impact on self-efficacy, and vicarious experiences and social persuasion should be especially important at the beginning of learning processes. However, so far, it is unclear whether and how these four sources are addressed in interventions on TSE, and whether they could serve as evidence-based design principles for the targeted promotion of TSE. This gap leaves many intervention approaches somewhat disconnected to self-efficacy theory. Moreover, it is also unknown whether the theoretically assumed differential effects of the sources also empirically hold in interventions (Klassen et al., 2011; D. B. Morris et al., 2017).

The second theoretical assumption relates to the malleability of teacher self-efficacy in different *career stages*: Bandura (1997) described self-efficacy beliefs as rather stable once set, leading to the justified expectation of less changeability in more experienced teachers. Empirical research has produced mixed results so far: On the one hand, longitudinal studies show in line with the theoretical assumption high stability of in-service teachers' self-efficacy (e.g., Bosse et al., 2016; Burić et al., 2022; Künsting et al., 2016; Praetorius et al., 2017) and significant changes in pre-service teacher self-efficacy (e.g., Burger, 2024; George et al., 2018; Ma et al., 2022; Mintz et al., 2020). On the other hand, there is some evidence that teacher self-efficacy can change in later career stages too, as, for example school closures during COVID-19, led to changes in in-service teachers' self-efficacy (Pressley, 2021). However, whether pre- and in-service teachers benefit equally from interventions intentionally aiming to promote teacher self-efficacy is unknown.

A detailed examination of the roles of the four sources of self-efficacy and teachers' career stages in the context of interventions is both theoretically and practically relevant: it helps determine whether the four sources offer a useful framework for designing effective interventions, and whether teacher self-efficacy can be equally promoted across all career stages or should, for instance, be targeted early in the career if it proves to be more malleable at that point. Such insights are essential for developing evidence-based approaches in both pre-service teacher education and in-service professional development. Moreover, they can inform strategic decisions by educational administrators and policy-makers regarding the allocation of resources and the design of support structures.

The third assumption of Bandura (1997) regarding the malleability of teacher self-efficacy refers to the proposed *context-specificity* of self-efficacy beliefs. This suggests that self-efficacy beliefs are shaped by situational demands and may vary across contexts (Bandura, 1997). So far, the proposed context-specificity has mainly been researched with regard to differences between teachers (e.g., Çoban et al., 2023; Fackler et al., 2021; Holzberger & Prestele, 2021), although self-efficacy beliefs may differ *within persons across contexts* (Bandura, 1997; Tschannen-Moran et al., 1998). However, this variability across contexts within teachers has seldom been explored (but see, for example, Raudenbush et al., 1992; Zee, de Jong, & Koomen, 2016). This gap is particularly relevant in secondary education, where teachers typically instruct multiple classes that differ in their demographic composition, their behavior, and their achievement levels. Understanding whether teachers adapt their self-efficacy beliefs to the specific conditions of each class — and thus whether TSE is responsive to contextual variation within individuals — is crucial for both theory and practice. If TSE varies across classes taught by the same teacher, future research might need to clarify how and when contextual factors should be integrated into self-efficacy measurement (see also Thommen et al., 2022). Likewise, teacher education and

training programs[1] may need to place greater emphasis on the specific classroom contexts that shape the experiences of both pre- and in-service teachers.

The following dissertation aims to address these theoretically and practically important research gaps about the malleability of teacher self-efficacy by conducting two empirical studies. The first study focused on the first two research gaps regarding the impact of the sources of self-efficacy and the career stage on the malleability of teacher self-efficacy. It is a meta-analysis summarizing the effects of 115 intervention studies from 26 countries on teacher self-efficacy. All included studies provided a pre- and post-measure of teacher self-efficacy; one-third additionally reported data from a control group without an intervention. Data came from 11,284 teachers, almost equally split between pre- and in-service teachers. It was analysed to what extent teacher self-efficacy can be promoted by interventions and how the sources of self-efficacy and the teachers' career stage affect the intervention effects. Further empirically relevant intervention and study characteristics were considered as robustness checks. The second study dealt with the third assumption, investigating the context-specificity of teacher self-efficacy. It is a primary study conducted in the Netherlands involving 26 experienced secondary school teachers who reported their self-efficacy in 74 classes. The teachers and the corresponding 1,326 students indicated demographic, behavioral, and achievement characteristics of the classes. It was analysed to what extent teachers report different levels of TSE when asked in specific classes and which class characteristics predicted the level of class-specific teacher self-efficacy. In the spirit of open science, both studies were pre-registered prior to data collection (https://osf.io/ev7tf and https://osf.io/uk2xq), and the analysis code is publicly available for both studies (https://osf.io/65vfq and https://osf.io/zxcvt). For Study 1, the material and the data have been released for reuse, too.

Both studies investigate the malleability of teacher self-efficacy, but from two different perspectives: Whereas Study 1 analyzes *intentional changes* of teacher self-efficacy due to interventions and the analyses focus on differences *between* teachers, Study 2 provides information on *natural variation*[2] due do different classes and the analyses focus on differences *within teachers*. With the two levels of analyses, this dissertation supports claims that self-efficacy should be researched at the within- and between-level (Beck & Schmidt, 2012) and adds knowledge to the so far sparse research on within-processes in educational psychology (Dirk & Nett, 2022; Murayama et al., 2017). Given the popularity and assumed predictive power of teacher self-efficacy for healthy, satisfied, effective, and innovative teachers, the results of both studies are of high importance for future research on teacher self-efficacy, the development of teacher education and training programs, and the allocation of resources.

[1] Throughout this dissertation, the term *teacher education and training* is used to refer to structured learning opportunities for both pre-service and in-service teachers

[2] The term "natural variation" is used instead of "change" in Study 2 to reflect its cross-sectional design, which captures differences in teacher self-efficacy across classroom contexts without implying temporal development.

The remainder of this dissertation is organized as follows: Chapter 2 provides an overview of the definitions and domains of teacher self-efficacy. It is followed by a review of the current literature on the outcomes of teacher self-efficacy in Chapter 3, which underlines the necessity to study the malleability of teacher self-efficacy. Chapter 4 describes the central assumptions of socio-cognitive theory regarding the malleability of self-efficacy. Chapter 5 elaborates on the theoretical background of the two empirical studies, outlining the important parameters of intentional changes through interventions and natural variation across classes. Chapter 6 introduces an overview of both studies, including the central research gaps and corresponding research questions. The methodology and results from the first study, the meta-analysis on the promotion of teacher self-efficacy, can be found in Chapter 7. The methodology and results from the second study, the primary study on the variation of teacher self-efficacy across classes, are described in Chapter 8. Chapter 9 is dedicated to the discussion and interprets the studies' findings in regard to the three theoretical assumptions. Limitations, considerations for future research and implications for educational administration and practice complete Chapter 9.

2 Definitions and Domains of Teacher Self-Efficacy

2.1 Definitions of Teacher Self-Efficacy

Most authors defining teacher self-efficacy refer to Bandura's construct of self-efficacy that he introduced in his socio-cognitive theory (Bandura, 1977, 1997). As Bandura over time used slightly varying definitions (see Table 1, p. 15 in Bach, 2022), conceptualizations of teacher self-efficacy also slightly vary depending on the definition of self-efficacy they use. Bandura's earlier definitions, for example, cover two central aspects of self-efficacy, namely the expectation that a particular behavior will lead to a certain result (*outcome expectancy*) and the expectation of being able to perform a particular behavior oneself (*self-efficacy* or personal efficacy). In later works, he considered outcome expectancy to be separate from self-efficacy beliefs (Bandura, 1997, 2004, 2012). Furthermore, in the 60s, Rotter (1966) introduced the concept "locus of control" which covers people's conviction whether the cause for an outcome is rather within their control (internal locus of control) or outside their control (external locus of control). The concept of "locus of control" has since then been mixed up with outcome expectancy, leading to two almost separate strands of research. Both are talking about teacher self-efficacy (or teacher efficacy), describing related, but conceptually different constructs (so-called *jingle fallacy*; see also Henson, 2002; Pajares, 1992, 1996; Tschannen-Moran et al., 1998). The first strand is described by measures capturing how much teachers believe that they can *impact student outcomes*. The second strand examines teacher self-efficacy with measures asking teachers to rate to what extent they *believe they can execute* a given teaching-specific task. Example items are given in Table 1.

Table 1. Example Items for Two Different Strands of Teacher Self-Efficacy Definitions

Teacher Self-Efficacy Strand 1 (Locus of Control)[3]	*Teacher Self-Efficacy Strand 2 (Personal Efficacy)*
• "When a student does better than usual in science, it is often because the teacher exerted a little extra effort" (Enochs & Riggs, 1990) • "The amount a student can learn is primarily related to family background" (Gibson & Dembo, 1984)	• "How certain are you that you can explain central themes in your subjects so that even the low-achieving students understand?" (Skaalvik & Skaalvik, 2007) • "To what extent can you use a variety of assessment strategies?" (Tschannen-Moran & Woolfolk Hoy, 2001) • "I am confident in my ability to prevent disruptive behaviour in the classroom before it occurs" (Sharma et al., 2012)

[3] It must be noted that the instruments developed by Enochs and Riggs (1990) and Gibson and Dembo (1984) also contain a scale called "personal efficacy". However, these scales lack other requirements of self-efficacy measures adhering to Bandura's (2006) ideas such as capturing the capability.

This dissertation falls within the second strand and adheres to Bandura's later definition of 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 in this dissertation is therefore defined as an individual teacher's belief in being able to successfully perform teaching-specific tasks (for similar definitions, see Holzberger et al., 2013; Skaalvik & Skaalvik, 2007; Tschannen-Moran et al., 1998).

2.2 The Domain-Specificity of Teacher Self-Efficacy

Bandura (1997) pointed out that general self-efficacy beliefs, such as a person's belief in their own ability to handle challenges and persist through difficulties, have little to no predictive power for certain behaviors. Instead, self-efficacy beliefs are rather specific regarding the activity domains relevant for a person. Domains require a similar set of skills to deal with the domain's tasks successfully. Consequently, self-efficacy domains have been specified mainly in different professions such as engineering (e.g., Mamaril et al., 2016), health care (e.g., Axboe et al., 2016) or teaching (e.g., Sharma et al., 2012; Tschannen-Moran & Woolfolk Hoy, 2001).

In the field of teacher self-efficacy several domain-specific scales have been developed, which slightly differ in focus and number of their domains (e.g., six domains in D. W. Chan, 2008a; Skaalvik & Skaalvik, 2007; three domains in Sharma et al., 2012; Tschannen-Moran & Woolfolk Hoy, 2001). Many domain-specific scales show a considerable overlap with regard to the three domains of *classroom management*, *student engagement,* and *instructional strategies*:

Teacher self-efficacy for classroom management (TSE for CM) captures teachers' beliefs about their ability to establish and enforce rules and routines and to proactively deal with disruptions in the classroom (e.g., D. W. Chan, 2008a; Sharma et al., 2012; Skaalvik & Skaalvik, 2007; Tschannen-Moran & Woolfolk Hoy, 2001). *Teacher self-efficacy for student engagement* (TSE for SE) evaluates teachers' views of their ability to communicate the value of learning to students and to motivate those who are not very interested (e.g., D. W. Chan, 2008a; Sharma et al., 2012; Skaalvik & Skaalvik, 2007; Tschannen-Moran & Woolfolk Hoy, 2001). *Teacher self-efficacy for instructional strategies* (TSE for IS) measures the extent to which teachers feel able to use a variety of instructional methods and to adapt their instruction to the achievement level of individual students (e.g., D. W. Chan, 2008a; Sharma et al., 2012; Skaalvik & Skaalvik, 2007; Tschannen-Moran & Woolfolk Hoy, 2001).

These three domains of teacher self-efficacy can be viewed as central and relevant for many teachers regardless of their experience, their specific subject, their country or the school type they are teaching (Brouwers & Tomic, 2000; Dicke et al., 2014; Fackler et al., 2021; Hartl & Holzberger, 2022; Holzberger & Prestele, 2021; Klassen & Chiu, 2010; Klassen & Tze, 2014; Skaalvik & Skaalvik, 2007; Zee, Koomen, et al., 2016). They are also measured in the Teachers' Sense of Efficacy Scale (TSES) by Tschannen-Moran and Woolfolk Hoy (2001), a scientifically well-validated and common questionnaire

of teacher self-efficacy (e.g., Burgueño et al., 2019; M.-L. Chang & Engelhard, 2016; Duffin et al., 2012; Fives & Buehl, 2010; Ruan et al., 2015; Tsigilis et al., 2010), which will be used in the two studies of this dissertation. Table 2 displays example items of the TSES for each domain.

Table 2. Example Items to Measure Teacher Self-Efficacy

Domain of Teacher Self-Efficacy	*Example Item from the TSES* (Tschannen-Moran & Woolfolk Hoy, 2001)
Classroom Management	How much can you do to get children to follow classroom rules?
Student Engagement	How much can you do to motivate students who show low interest in school work?
Instructional Strategies	To what extent can you provide an alternative explanation or example when students are confused?

The three domains are interrelated, but empirically separable (e.g., Klassen et al., 2009; Pfitzner-Eden et al., 2014) and teachers can possess different levels of self-efficacy in each domain (e.g., Perera et al., 2019). In line with the specificity matching hypothesis, which posits that relations between constructs are stronger when they are thematically aligned (Ajzen & Fishbein, 1977), the three domains of TSE exhibit distinct predictive power: teacher self-efficacy for classroom management is, for example, more closely associated with actual classroom management behavior than self-efficacy for student engagement (e.g., Holzberger & Prestele, 2021; Lazarides et al., 2023).

So far, it is unknown whether the specificity matching hypothesis also applies for the malleability of the domains. There is initial evidence coming from two studies with pre-service teachers that point to differences in the domains' reactivity due to interventions (Dalioglu & Adiguzel, 2016; Pfitzner-Eden, 2016a); however, no clear pattern could be recognized as the domains, which changed, differed in the two studies. Understanding whether the domains differ in their malleability is important to design targeted and effective intervention and support programs.

3 Outcomes of a High Level of Teacher Self-Efficacy

The necessity to gain a more detailed understanding of the malleability of teacher self-efficacy becomes evident when looking at the impressive amount of primary studies, systematic reviews and meta-analyses indicating beneficial effects of a high level of teacher self-efficacy for teachers' well-being, their instructional quality and their students' outcomes (Aloe, Amo, & Shanahan, 2014; Chesnut & Burley, 2015; Duan et al., 2024; Haverback & Parault, 2008; Kim & Seo, 2018; Klassen & Tze, 2014; Klassen et al., 2011; Kleinsasser, 2014; Lee et al., 2017; Mu et al., 2024; Zee & Koomen, 2016). Zee and Koomen (2016) made a first effort to structure the abundance of dependent outcomes of teacher self-efficacy into the three categories of teacher well-being, classroom processes, and students' academic adjustment. The following paragraphs will present key findings related to these three categories of teacher self-efficacy outcomes, focusing on studies that have not gone into the review by Zee and Koomen (2016) due to their publication date. It presents a carefully selected, comprehensive, but not exhaustive, set of studies that represent the current state of research. For this reason, studies using measurement instruments that are comparable and designed in accordance with the guidelines presented by Bandura (2006) were primarily considered.

3.1 Teacher Self-Efficacy and Teacher Well-Being

Across studies, teachers consistently report more well-being when having higher levels of teacher self-efficacy: They demonstrate, for example, higher commitment to the job and higher satisfaction with the job (Chesnut & Burley, 2015; Duan et al., 2024; Klassen & Chiu, 2010; Stephanou et al., 2013; Toropova et al., 2021; H. Wang et al., 2015). They are also less likely to experience burnout, especially emotional exhaustion, and in general exhibit fewer stress symptoms (Aloe, Amo, & Shanahan, 2014; Betoret, 2006; Chiang et al., 2025; Duan et al., 2024; Fernet et al., 2012; Skaalvik & Skaalvik, 2010, 2014; H. Wang et al., 2015; Zimmermann et al., 2016). Unsurprisingly, teachers with a higher level of self-efficacy experience more positive emotions, such as joy and pride, than teachers with lower levels (Burić & Moè, 2020; Burić et al., 2020; Hascher & Hagenauer, 2016). Moreover, a recent meta-analysis confirmed that teachers with higher levels of self-efficacy are more resilient (and vice versa, Mu et al., 2024). Finally, teacher self-efficacy not only has beneficial *direct* effects on teacher well-being but also works *indirectly* as a moderator: Teacher self-efficacy can buffer stressful experiences such as classroom disturbances and thereby reduce adverse outcomes (e.g., Dicke et al., 2018; Embse et al., 2016).

Although the presented research must be interpreted with some caution as cross-sectional studies and samples with in-service teachers are over-represented (exceptions: Burić & Moè, 2020; Chesnut & Burley, 2015; Zimmermann et al., 2016) and longitudinal research points to conditions where emotional exhaustion and teacher self-efficacy are negatively related (Dicke, Parker, et al., 2015; Olivier

et al., 2024), the findings overall indicate a substantial, positive correlation between teacher self-efficacy and well-being.

3.2 Teacher Self-Efficacy and Classroom Processes

Teacher self-efficacy not only has a positive impact on teachers' well-being, but also on their teaching. Teachers with higher levels of self-efficacy report engaging more often in desirable behaviors such as emotional support (Hettinger et al., 2023b; Hußner et al., 2024; Lazarides & Schiefele, 2024), autonomy-supportive teaching practices (Lauermann & Berger, 2021), differentiated instruction (Neve et al., 2015; Suprayogi et al., 2017), or the integration of technology (Konstantinidou & Scherer, 2022). They also give more positive feedback (Woodcock et al., 2019), display more teaching enthusiasm (Burić & Moè, 2020), and have less conflictual but closer relationships with their students (Hajovsky et al., 2020). There is a plethora of studies showing positive links between teacher self-efficacy and instructional quality, regardless of whether students, external observers or the teachers themselves rated their instructional quality (Burić & Kim, 2020; Holzberger et al., 2013; Holzberger & Prestele, 2021; Klassen & Tze, 2014; Künsting et al., 2016; Lazarides et al., 2020; Leijen et al., 2024; Malinen et al., 2024; Ryan et al., 2015; but see for exceptions Lazarides et al., 2021; Thommen et al., 2022).

3.3 Teacher Self-Efficacy and Student Outcomes

Turning to outcomes on the student level, it must be noted that there is a limited number of studies evaluating direct effects on student outcomes and, at the same time, a growing number of studies pointing to a lack of systematic relations (e.g., Jerrim et al., 2023; Lauermann & Berger, 2021; Lazarides et al., 2023). Some findings point to higher levels of interest and engagement of students when taught by teachers with higher levels of self-efficacy (e.g., Fauth et al., 2019; mediated by teaching quality; Hettinger et al., 2023b; moderated by student-perceived emotional support; Zee & Koomen, 2020). In a longitudinal study by Hettinger et al. (2023a), 10th-grade students enjoyed math lessons more when they were taught by a teacher with a higher level of self-efficacy at the beginning of Grade 9. One meta-analysis also points to a small, positive relation between teacher self-efficacy and students' academic achievement (r = .10 in Kim & Seo, 2018). In sum, the conclusions on the relation between teacher self-efficacy and student outcomes are somewhat limited. It can be assumed that teacher self-efficacy is mainly indirectly related to student outcomes via classroom processes and teacher well-being.

3.4 The Need to Investigate the Malleability of Teacher Self-Efficacy

This selective review of studies evaluating outcomes related to TSE shows that, despite some shortcomings in the studies (e.g., overrepresentation of studies on instructional quality and teacher well-being compared to other outcomes), at our current state of knowledge, a high level of teacher self-efficacy has overall beneficial effects for teachers themselves, their instruction, and their students. To maintain healthy, enthusiastic and effective teachers, it is consequently necessary to investigate how a

high level of TSE can be fostered. This dissertation, therefore, considers two approaches that can impact the level of TSE. Study 1 evaluates the effects of interventions (approach: intentional change) on teacher self-efficacy and investigates how interventions must be designed to best promote TSE. Study 2 examines the role of specific class characteristics such as the gender composition or the amount of misbehavior, which naturally surround the teachers in their everyday life (approach: natural variation), for the level of teacher self-efficacy.

Understanding these two malleability processes of TSE is of further societal concern. Teacher self-efficacy is closely related to job satisfaction and commitment — two variables that have been repeatedly identified as important variables *preventing teacher turnover* and drop-out (e.g., Gundlach et al., 2024; Li & Yao, 2022). Given the high economic and educational costs of teacher drop-out (e.g., Anders, 2025; Guin, 2004; Learning Policy Institute; Ronfeldt et al., 2013; Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2025), knowing how to promote and support high levels of TSE might be, in the long term, a pathway to reduce teacher drop-out.

4 Theoretical Assumptions about the Malleability of Teacher Self-Efficacy

Bandura (1997, 2006) outlined central characteristics of self-efficacy beliefs which contain implications for the malleability that are empirically tested in this dissertation.

4.1 The Interplay of Self-Efficacy, Behavior and Environment in the Model of Triadic Reciprocal Causation

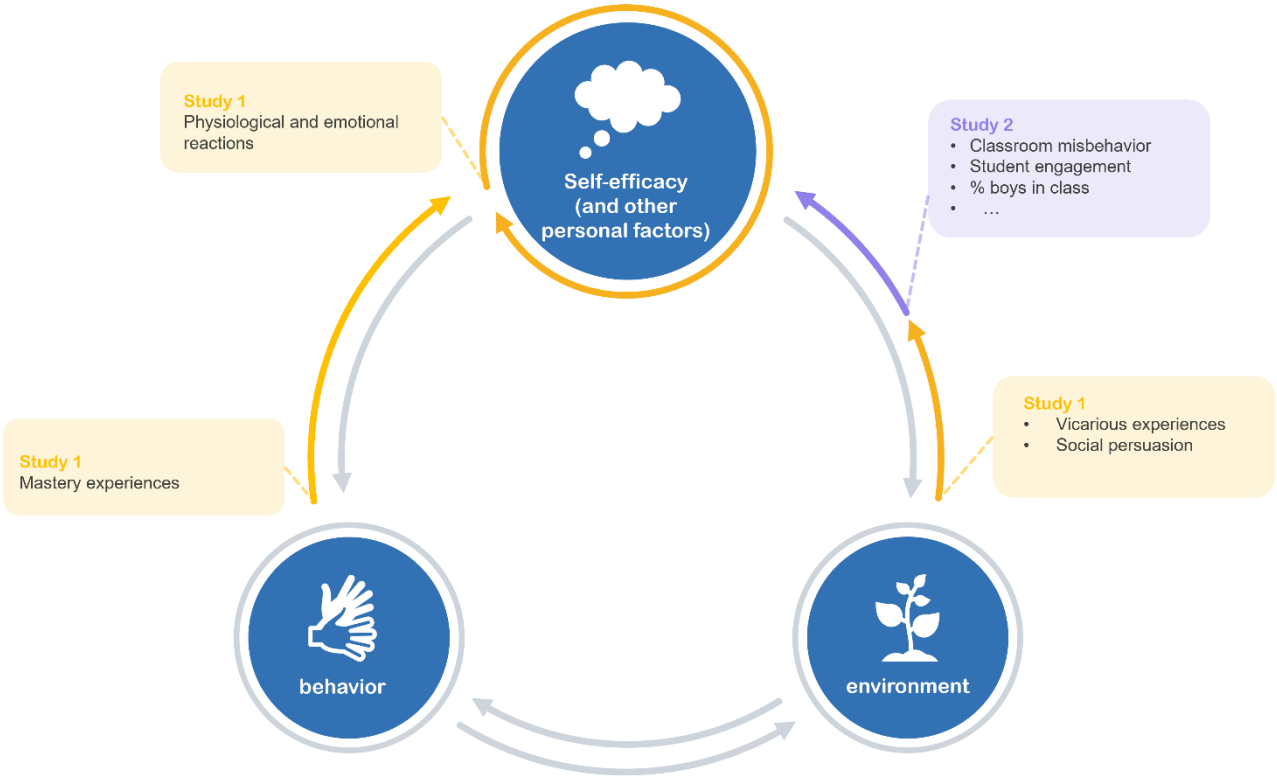
An essential part of Bandura's socio-cognitive theory is the assumption that behavior, personal factors (described as cognitive, affective, and biological factors), and environmental events influence each other and are interdependent, which is called the *model of triadic reciprocal causation* (see Figure 1). Self-efficacy beliefs are seen as central cognitive factors within the personal factors and can lead to changes in the behavior, the environment, and the affective and biological reactions. To give an example, teachers with high self-efficacy beliefs (cognitive factor) might prepare the classroom in a way to prevent disturbances (behavior), subsequently notice a quiet classroom around them (environmental event), and a steady pulse (biological reaction) and joy (affective reaction) within them. The other way around is also possible: An environmental event such as a disturbance in the classroom can lead to cognitive, affective, biological, and behavioral reactions: a teacher might yell at the student causing the disturbance (behavioral reaction), sweat (biological reaction), feel anger (affective reaction), and lower his or her self-efficacy beliefs (cognitive reaction). These two examples shall illustrate the idea of socio-cognitive theory that people (teachers) are both producers and products of their environment. Focusing on teacher self-efficacy, this means that self-efficacy beliefs are precursors and consequences of behaviors, environments, affective and biological reactions.

Concerning teacher self-efficacy, most research so far has treated teacher self-efficacy as a precursor and evaluated consequences of teacher self-efficacy for other personal factors (e.g., burnout, see Aloe, Amo, & Shanahan, 2014; Skaalvik & Skaalvik, 2010), teachers' behavior (e.g., their instructional quality, see Burić & Kim, 2020; Holzberger et al., 2013), or their environment (e.g., students' motivation, see Hettinger et al., 2023b; Lauermann & Berger, 2021). Far less research exists on the other pathways viewing teacher self-efficacy as a consequence of affective and biological reactions, behavior, and environment.

Therefore, the two studies in this dissertation focus on *teacher self-efficacy as a consequence*: Study 1 examines the effects of interventions on teacher self-efficacy and focuses on the four sources of self-efficacy that can be sorted into pathways from affective and biological reactions to self-efficacy, from behavior to self-efficacy, and from environment to self-efficacy (Figure 1). Study 2 focuses on the pathway from environment to self-efficacy, the so-called "context-specificity" and evaluates to what

extent different class characteristics — representing different environments within secondary school teachers — impact teacher self-efficacy (Figure 1).

Figure 1. The Two Empirical Studies within the Triadic Reciprocal Causation



4.2 The Four Sources of Self-Efficacy Beliefs

According to Bandura (1997), self-efficacy beliefs evolve from processing four types of experiences, the so-called sources of self-efficacy: mastery experiences, vicarious experiences, social persuasion, and physiological and emotional reactions[4]. Mastery experiences occur whenever someone successfully manages a situation or challenge due to own efforts and activities. Vicarious experiences are made when someone observes a model successfully managing a situation or challenge. Social persuasion refers to experiences of receiving feedback and verbal encouragement. Physiological and emotional reactions, such as increased palpitation or worrying thoughts, also provide information about one's capability.

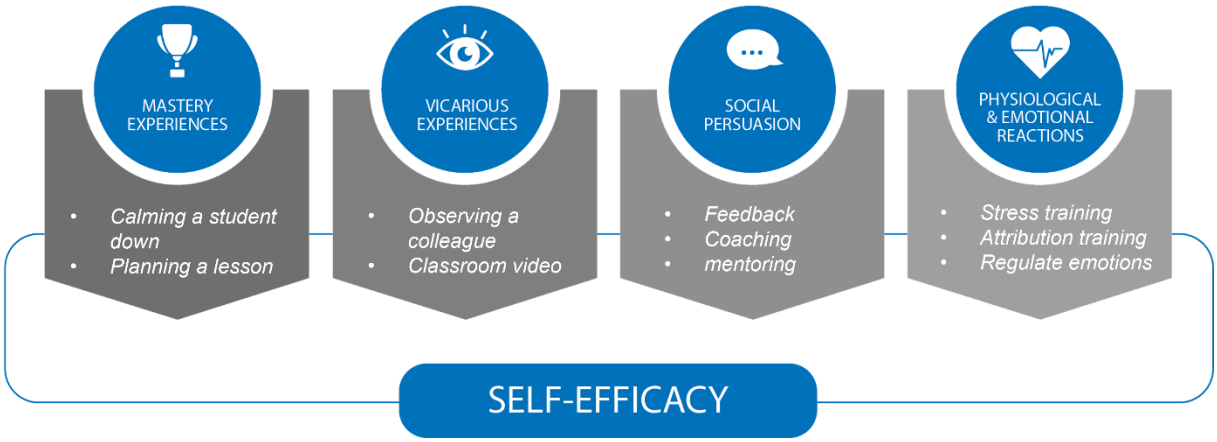
[4] Various synonyms are used in the literature to describe the four sources: Mastery experiences are also named performance accomplishments or enactive attainments. Social persuasion also appears as verbal persuasion, physiological and emotional reactions are sometimes called physiological and emotional reactions, physiological and affective states or physiological and emotional states (e.g., Bandura, 1977, 1997).

With regard to malleability processes of teacher self-efficacy, these sources can be *specifically addressed*, for example, in *interventions* (intentional change), or they can become more or less *salient* depending on the specific *environment* (natural variation). An intervention with role plays, for example, explicitly targets mastery experiences (e.g., Aasheim et al., 2020; Dicke, Elling, et al., 2015). Similarly, a class with fewer disruptions is an environment that facilitates more successful mastery experiences in classroom management than a class with more disruptions (e.g., Fackler et al., 2021; Heinschel et al., 2024). Further examples are specified in Figure 2.

Research in the field of teacher self-efficacy has fallen short in translating and capturing the sources of self-efficacy to the context of teaching (Bach, 2022; Klassen et al., 2011; D. B. Morris et al., 2017). Therefore, this dissertation aims to advance the understanding of the sources of teacher self-efficacy by a systematic analysis of sources addressed in interventions in Study 1 and proposing a variety of demographic, behavioral, and achievement characteristics of classes in secondary schools as different opportunities for mastery experiences.

Bandura (1997) also stated differential strengths of the four sources. He described mastery experiences as the most influential source of self-efficacy, but at the same time highlighted the relevance of vicarious experiences and social persuasion when little previous experience is available or new tasks are faced. Study 1 explicitly tests how different combinations of the sources influence the intervention effects and whether vicarious experiences and social persuasion are more important for pre-service teachers with less experience than for in-service teachers.

Figure 2. Examples of the Four Sources of Self-Efficacy in Interventions and the Classroom Environment



4.3 The Durability of Self-Efficacy Beliefs

There are two further important claims by Bandura (1997) to consider when thinking about the malleability of teacher self-efficacy. He stated that self-efficacy is rather stable once set and that "it takes time for a causal factor [as outlined in the triadic reciprocal model] to exert its influence" (p. 6). Consequently, self-efficacy is supposed to be rather stable after an initial malleable phase. Moreover, it is expected not to spontaneously change whenever a new environmental factor or behavior appears, but also not to be immutable to change when *enduring* environmental factors or behaviors hit. Empirical findings so far support these claims: Longitudinal studies show significant changes of teacher self-efficacy during the transition from pre-service to in-service teacher stage (Abele & Candova, 2007; Hartl & Holzberger, 2022; Mintz et al., 2020; Woolfolk Hoy & Spero, 2005) and high stability rates within in-service teachers (e.g., Bosse et al., 2016; Burić et al., 2022; Künsting et al., 2016; Praetorius et al., 2017). Regarding the second claim, school closures during COVID-19, as an example of an enduring environmental factor, also impacted teachers' level of self-efficacy (e.g., Pressley, 2021).

However, it remains unanswered what these claims mean for *intentional* change processes of interventions: does the self-efficacy of in-service teachers, which is, based on the longitudinal studies, considered as relatively stable, change as much by interventions as the self-efficacy of pre-service teachers, where TSE is considered as less set? It is also unclear whether the time of a causal factor can be further operationalized, for example, whether the duration of an intervention or the period between the measurement points impacts the change of teacher self-efficacy due to interventions. Hence, Study 1 of this dissertation explicitly tests whether the teachers' career stage (in-service vs. pre-service), the interventions' duration and the period between measurement points moderate the intervention effects on changes in teacher self-efficacy.

5 Change Processes of Teacher Self-Efficacy

The previous chapters have demonstrated the need to investigate the malleability of teacher self-efficacy (Chapter 3) and the theoretical assumptions of Bandura's (1997) socio-cognitive theory regarding the malleability of teacher self-efficacy (Chapter 4). According to the theory, the four sources of self-efficacy, and in particular mastery experiences, play a decisive role for the formation of teacher self-efficacy (Bandura, 1997). In change processes, opportunities for mastery experiences can be specifically created, for example, in interventions, or be available to a greater or lesser extent in the environment. This difference in how intentional mastery experiences and other sources are presented led to the description of the investigated change processes in Study 1 as "intentional" (focus on change through interventions) and in Study 2 as "natural" (focus on change through different classes). While data for Study 1 comes from pre- and post-measurements and aspects of time such as duration and period can be analyzed as causal factors, data for Study 2 is cross-sectional and no causal relationships over time can be investigated. Therefore, the malleability process of Study 1 is described as "change", the process of Study 2 as "variation". The following sections present the theoretical background of the two empirical studies, reviewing the research on interventions and classroom characteristics as change factors for teacher self-efficacy in more detail and highlighting the studies' relevance for educational research, teacher education and training, and administrative decision-making.

5.1 Intentional Changes of Teacher Self-Efficacy Through Interventions (Theoretical Background for Study 1)[5]

In recent years, a wide range of intervention approaches have emerged to strengthen and promote teachers' self-efficacy. Interventions in this study are defined as organized learning opportunities that differ from the everyday experiences of pre- and in-service teachers. Intervention approaches 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), or 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). This

[5] The following sections up to and including 5.1.5 are based on the published version of Study 1:
Täschner, J., Dicke, T., Reinhold, S., & Holzberger, D. (2024). "Yes, I can!" A systematic review and meta-analysis of intervention studies promoting teacher self-efficacy. *Review of Educational Research*, *95*(1), 3–52. https://doi.org/10.3102/00346543231221499
All sections have been reviewed and slightly revised for consistency and integration into the broader framework of this dissertation.

variety of approaches leaves the questions unanswered: to what extent teacher self-efficacy can be promoted on average, and which intervention characteristics constitute key factors of effective teacher self-efficacy promotion. Therefore, a meta-analysis seemed necessary.

5.1.1 The Role of the Four Sources of Self-Efficacy in Interventions

Despite the different characteristics of the interventions on the surface (e.g., content, activities), the four sources of self-efficacy (i.e., mastery experiences, vicarious experiences, social persuasion, and physiological and emotional 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).

Despite their different contents, organizational structures, involved practices, and approximation levels, interventions can offer the same sources of self-efficacy. For example, an intervention entitled to specifically foster 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 offer the same sources of self-efficacy, namely mastery experiences, vicarious experiences, and social persuasion when examined in detail (see Table 3).

Table 3. Addressed Sources of Self-Efficacy in Exemplary Interventions

	Targeted Sources of Self-Efficacy			
Intervention Title	*Mastery Experiences*	*Vicarious Experiences*	*Social Persuasion*	*Physiological and Emotional Reactions*
Classroom Management Training (Dicke, Elling, et al., 2015)	Yes	Yes	Yes	No
Student Teaching Practicum (Holdaway, 2017)	Yes	Yes	Yes	No
Coaching (Weber et al., 2019)	Yes	Yes	Yes	No

Coding the different sources of self-efficacy targeted by the interventions, is seen as a promising method to compare seemingly diverse interventions and to decompose successful strategies for fostering teacher self-efficacy. The following paragraph highlights examples for each source of self-efficacy from teacher education and training interventions derived from the systematic review of relevant studies (see also Table 4).

Mastery experiences take place whenever someone gains operational experience by doing something[6]. 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). 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*. 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 vs. excitement. For example, when teachers interpret their palpitation 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, on the contrary, also predicted (negative 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).

[6] 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, mastery experiences here are conceptualized as operational experiences. Possible biases due to this simplification are elaborated in the discussion.

Table 4. Examples of Operationalization of Sources of Self-Efficacy in Intervention Studies

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, 2012; Weber et al., 2019) • teaching mini-lessons to peers (Yough, 2019) • micro-teaching (Karalar & Altan, 2018; Sheehan & Moore, 2019; Wagler & Moseley, 2005)
Vicarious experiences	• video-taped 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 and emotional reactions	• yoga, meditation, breathing techniques or other mindfulness exercises (Ansley et al., 2021; de Carvalho et al., 2021)

5.1.2 The Central Role of Mastery Experiences 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 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. This is important to know to design effective and sustainable interventions for teacher self-efficacy.

5.1.3 Further Relevant Intervention Characteristics

Previous research on teacher self-efficacy and the professional development of teachers has pointed to further characteristics of interventions that might influence the interventions' effects and that are important to consider to derive a strong evidence base on how to design effective interventions for teacher self-efficacy.

Knowledge. It is still controversial in research how pedagogical and content knowledge impact teacher self-efficacy and whether increases in knowledge are significantly associated with increases in self-efficacy (Dicke, Parker, et al., 2015; Gold et al., 2017; Lauermann & König, 2016; D. B. Morris et al., 2017; Thomson et al., 2022). In qualitative studies, where teachers were asked what they base their self-efficacy on, various types of knowledge are repeatedly cited as a root of self-efficacy (Bautista & Boone, 2015; Palmer, 2011; Phan & Locke, 2015). Therefore, in this meta-analysis, it is investigated whether interventions that target knowledge alone differ from interventions that target at least one of the four sources of self-efficacy, according to Bandura.

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., 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. Therefore, it is tested in this meta-analysis, whether interventions that provide a moment of reflection produce larger effects than interventions without a moment of reflection. This analysis provides insights into the question whether an explicit 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) which aligns with Bandura's (1997) statement that "it takes time for a causal factor to exert its influence" (p. 6). In a study on the implementation of differentiated instruction in the classroom, the amount of professional development hours was positively related to the teachers' level of self-efficacy beliefs

(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. Accordingly, the meta-analysis presented here investigates whether there is a linear relationship between the intervention's duration in hours and its effects on teacher self-efficacy.

In addition to the duration of an intervention, this meta-analysis considered the period between the measurement points as another aspect of "time". 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). However, Bauer et al. (2020) emphasize that there is no evidence of larger effects of longer internships on pre-service teachers' competencies compared to shorter periods. Study 1 therefore examines whether longer periods of interventions are connected with larger effect sizes.

5.1.4 The Role of the Career Stage in Interventions

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 significant increases during teacher education and transition from pre-service to in-service teacher stage (e.g., Burger, 2024; Fives et al., 2007; Ma et al., 2022; Pereira Kastens et al., 2020) in-service teachers' self-efficacy tends to be rather stable over several years (Bosse et al., 2016; Burić et al., 2022; Künsting et al., 2016; Praetorius et al., 2017; Savolainen et al., 2020). However, other research shows that self-efficacy is certainly malleable over the teaching career: In comparative studies, teachers with more years of experience report higher levels of TSE (D. W. Chan, 2008b; 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 the relevant domains of classroom management, student engagement, and instructional strategies on which they can base their self-efficacy ratings. Also, studies evaluating specific interventions or major environmental changes such 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, 2021). Summed up, there is initial 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 will 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, Study 1 investigates whether pre-service and in-service teachers react similarly to interventions promoting teacher self-efficacy.

5.1.5 Possible Interactions between Career Stage and the Sources of Self-Efficacy

Bandura (1997) further describes that the four sources of self-efficacy 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 (Grossman, Compton, et al., 2009; Grossman, Hammerness, & McDonald, 2009; Loewenberg Ball & Forzani, 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 looking 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 three 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 specifically for the development of pre-service teachers' self-efficacy (Klassen & Durksen, 2014; D. B. Morris et al., 2017; Pfitzner-Eden, 2016b; Tschannen-Moran & Johnson, 2011; Tschannen-Moran et al., 1998; 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 may gain mastery experiences in areas they are not yet comfortable with (e.g., using new technologies; dealing with heterogeneous classes), and these mastery experiences may, therefore, affect in-service teachers' self-efficacy as well.

In summary, it is still unclear whether the three sources[7]: 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.

5.1.6 Unique Contributions of Study 1 to Educational Research, Practice and Administration

Study 1 investigates the malleability of TSE through intentional changes induced by interventions. It addresses two central research gaps: the role of the four sources of self-efficacy and the influence of career stage on intervention effects (Chesnut & Burley, 2015; Klassen et al., 2011; D. B. Morris et al., 2017). The study aims to provide relevant information for self-efficacy theory, educational researchers, practitioners in teacher education and training, as well as educational administration and policy (see Table 5).

Self-efficacy theory receives some sort of falsification or verification as Study 1 tests whether theoretically proposed source combinations differ in their effectiveness, whether teacher self-efficacy is malleable in all career stages and whether interactions between career stage and sources of self-efficacy influence the intervention outcome.

Educational researchers are given a benchmark for future studies as the average effect size of interventions is calculated. Study 1 also offers them a systematic overview of which sources are targeted in existing interventions, helping them to navigate the current landscape of intervention research and to detect points for future research.

Practitioners in teacher education and training are provided with evidence-based guidance on which combinations of self-efficacy sources (and further intervention characteristics) are most effective. Further, they might benefit from insights into the specificities of the promotion of teacher self-efficacy in different career stages.

Educational administration and policy, finally, receives information about expected impacts of investments in teacher education and training. They are also informed about the current methodological quality of intervention studies, as each primary study is evaluated for quality, and can, on the one hand, ensure that policy recommendations are based on robust evidence and, on the other hand, adapt funding requirements for future research.

[7] 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 it was not possible to draw any conclusions about presumable interactions with teachers' career stages.

Table 5. Links of Study 1 to Research, Practice and Policy

Interest Groups	*Corresponding Questions That Are Answered by Study 1*
Self-Efficacy Theory	• To what extent is teacher self-efficacy malleable in different career stages? • Do different source combinations in interventions differ in their effect on teacher self-efficacy? • Are specific source combinations more effective in specific career stages?
Researchers	• What is the average effect of interventions aiming to promote teacher self-efficacy? • Which sources and source combinations are currently considered in intervention studies? • Where is a gap for future research?
Practitioners in Teacher Education and Training	• What is the average effect that can be expected from interventions aiming to promote teacher self-efficacy? • Which intervention characteristics and sources of self-efficacy influence the effects and should be considered in future interventions? • Do all teachers benefit from interventions regardless of their career stage?
Educational Administration & Policy	• Is an intervention worth the investment if TSE changes to this extent? • Does it make sense to invest in interventions in all career stages? • Where is further investment or resource allocation needed?

5.2 Natural Variation of Teacher Self-Efficacy Across Classrooms (Theoretical Background of Study 2)[8]

Based on the model of triadic reciprocal causation (Bandura, 1997) it is assumed that teacher self-efficacy is a context-specific construct that is impacted by the environment. In the context of teaching, particularly school and class characteristics constitute important environmental factors (Duan et al., 2024; Fackler et al., 2021). One presumable argument for how the environment impacts the level of teacher self-efficacy is that environmental conditions offer varying degrees of opportunities for the four sources of self-efficacy (e.g., Fackler et al., 2021). Indeed, cross-sectional studies find, for example, higher levels of teacher self-efficacy when there is more collaboration among colleagues or when their

[8] The following sections up to and including 5.2.1 are based on a working paper co-authored by Janina Täschner, Doris Holzberger, Terrence D. Jorgensen and Marjolein Zee. The manuscript has not yet been submitted for publication and has been slightly revised to ensure coherence with the overall structure of this dissertation.

principals engage in instructional leadership, which both may provide more opportunities for vicarious experiences and social persuasion (Alanoglu, 2022; Bellibaş & Liu, 2017; Çoban et al., 2023; Fackler & Malmberg, 2016; Holzberger & Prestele, 2021). However, studies also show that TSE differs more within schools than between schools (Fackler et al., 2021; van Eycken et al., 2024), pointing out the need to evaluate factors that differ within schools, and especially within teachers.

There are first studies which indicate that environmental conditions differ not only between teachers but also *within teachers* and that these contexts impact the levels of TSE within teachers: Rupp and Becker (2021) found that pre-service teachers' TSE varied within the individuals across lessons in a practicum. In a study by Zee, de Jong, & Koomen (2016), primary school teachers reported different levels of self-efficacy across individual students. However, contrary to pre-service teachers in a practicum (sample in Rupp & Becker, 2021) and primary school teachers teaching one specific class (sample in Zee, de Jong, & Koomen, 2016), secondary school teachers' work life is particularly characterized by the instruction of several *different classes.* Further support comes from Fackler et al. (2021) who found that class characteristics such as the classroom climate or the number of low achievers in a class explained more variation between teachers' TSE than school characteristics, based on a sample of more than 100,000 secondary school teachers.

Classes differ with regard to their demographic, behavioral, and achievement characteristics and therefore offer different conditions, especially for mastery experiences (see Fackler et al., 2021 for a similar explanation). In an interview study, pre-service teachers explicitly mentioned class characteristics such as students' behavior as a form of mastery experiences (Martins et al., 2015). In-service teachers have also reported experiencing more mastery experiences in classes with more heterogeneous performance and student engagement (Malmberg et al., 2014). It is known that class characteristics explain more variation between teachers' TSE than school characteristics (Fackler et al., 2021), but the question whether teachers adapt their TSE also *within* themselves according to different class characteristics and which class characteristics are particularly aligned with within-teacher TSE has seldom been explored (but see Raudenbush et al., 1992 and Ross et al., 1996).

There is some evidence of substantial variance within teachers across classes from two older studies with secondary school teachers from the USA and Canada (Raudenbush et al., 1992; Ross et al., 1996). In these studies, between 21 % and 44 % of the variance in TSE was attributable to within-teacher variations across classes. Despite these remarkable results, these studies have not been replicated so far. Therefore, Study 2 aims to replicate these studies in a European educational system and to address explicit research gaps identified in these studies (e.g., including more student characteristics as predictors, as suggested in Ross et al., 1996). It also utilizes further developments in the measurement of TSE and the class characteristics of interest. To be specific, whereas Raudenbush et al. (1992) and

Ross et al. (1996) used a single item[9] to measure TSE, Study 2 applied a *domain-specific questionnaire* of TSE that is more in line with Bandura's (2006) guidelines of self-efficacy measures. Additionally, instead of using teachers as a single source of information for class characteristics and TSE as done in the previous studies, Study 2 used *students and teachers* as sources of information for class characteristics to diminish common method bias (Podsakoff et al., 2012).

Evaluating the extent to which TSE varies within teachers across classes is important for future research to design adequate measurement instruments (see also Thommen et al., 2022). Investigating the role of class characteristics is crucial for developing future support structures that foster high levels of teacher self-efficacy.

5.2.1 Class Characteristics Associated with Teacher Self-Efficacy

Ross et al. (1996) and Raudenbush et al. (1992) did not provide categories of their class characteristics under investigation. Nevertheless, the variables in their studies can be subsumed as behavioral, demographic, and achievement characteristics (see Table 6). In continuation of these two studies, these three categories were retained in Study 2 and filled with selected variables that appeared to be highly relevant for TSE based on current research discourse.

Table 6. Investigated Class Characteristics

Category of Class Characteristics	*Investigated Variable in Raudenbush et al. (1992)*	*Investigated Variable in Ross et al. (1996)*	*Investigated Variable in Study 2*
Behavioral Characteristics	• Student engagement	• Student engagement	• Class misbehavior • Behavioral engagement • Emotional engagement
Demographic Characteristics	• Average age of students • Class size		• % of students with migration background • % of boys
Achievement Characteristics	• Track of the class		• Average grade

[9] Raudenbush et al. (1992): "To what extent do you feel successful in providing the kind of education you would like to provide for the students in this class?"
Ross et al. (1996): "If I try really hard I can get through to even the most difficult or unmotivated students in this class".

Behavioral Characteristics of Classes and Their Relation to Teacher Self-Efficacy

The two older studies on variations in TSE across classes considered student engagement as the only behavioral characteristic of classes (Raudenbush et al., 1992; Ross et al., 1996). In both studies, student engagement was a significant predictor of within-teacher variations across classes in TSE (Raudenbush et al., 1992; and the only significant predictor in Ross et al., 1996). However, more recent research conceptualizes the construct of student engagement with at least two facets, including students' *behavioral engagement*, for example, their attention or involvement, and students' *emotional engagement*, such as their interest or enjoyment (e.g., Fredricks et al., 2004; Skinner et al., 2008). Lauermann and Berger (2021) showed that classes clearly differ in their emotional and behavioral engagement. The few studies considering student engagement and teacher self-efficacy reveal considerable positive interrelations between TSE for SE and CM and students' emotional and behavioral engagement (C.-F. Chang et al., 2022; Hettinger et al., 2023b). Due to their study designs, the question remains to what extent class-specific engagement is related to class-specific TSE.

Numerous studies have demonstrated that problematic behavior in classes is connected with higher levels of exhaustion and reduced TSE (Aldrup et al., 2018; Aloe, Shisler, et al., 2014; Schwab et al., 2021; Tsouloupas et al., 2010; Zee, de Jong, & Koomen, 2016). The time teachers spend on keeping order in a classroom has been found to be negatively related to secondary school TSE for CM and SE in over 32 different countries (Fackler et al., 2021). In another study, teachers disclosing a high need for professional development in classroom management were more often in a low self-efficacy profile (Perera et al., 2019). However, all these studies treated TSE as a trait-like, personal characteristic and did not look at possible class-specific differences. Consequently, Study 2 explored whether such close interrelations between TSE and student misbehavior also hold at the within-teacher level in a sample of secondary school teachers.

Demographic Characteristics of Classes and Their Relation to Teacher Self-Efficacy

A vast body of literature suggests that gender ratio in classes and teacher behavior are related. Boys experience more referrals and suspensions from lessons (Welsh & Little, 2018) and show more externalizing behavior than girls (Demmer et al., 2017). Teachers in secondary schools have more negative interactions with boys than with girls (Jones & Dindia, 2004) and teachers' classroom management skills are perceived as better in classes with more females (Fauth et al., 2020). However, it is more likely that male students' more demanding behavior, rather than their gender, may affect TSE. Indeed, teachers have previously been found to report higher levels of TSE in classrooms with girls only as long as they did not consider additional behavioral factors (Geerlings et al., 2018). Whether differences in the gender ratio between several classes contribute to differences in teachers' TSE has not been investigated yet.

In the last years, teaching in ethnically diverse classrooms has been debated as both providing opportunities and challenges in research (e.g., Björnsson, 2020; Forghani-Arani et al., 2019). There are

first empirical results that the ethnic composition of a classroom impacts teachers' TSE (Geerlings et al., 2018). Primary school teachers reported somewhat lower levels of TSE for ethnic minority students when teaching in a classroom with fewer ethnic minority students than in classrooms with a higher proportion of ethnic minority students (Geerlings et al., 2018). Although students' ethnic background contributed to TSE, the authors point out that students' behavior had a much higher relevance for TSE (Geerlings et al., 2018). So far, it remains unknown whether such student-specific differences in primary school teachers' TSE as in Geerling et al. (2018) also appear in secondary school teachers' TSE across classes with different proportions of ethnic composition.

Achievement Characteristics of Classes and Their Relation to Teacher Self-Efficacy

There is some evidence that TSE significantly and positively correlates with students' achievement (Caprara et al., 2006; Fauth et al., 2019; Kim & Seo, 2018). Two cross-sectional studies with data from the Teaching and Learning International Survey (TALIS) from the cycles 2008 and 2013 showed that teachers feel more self-efficacious when their students perform at high levels (Fackler & Malmberg, 2016; Fackler et al., 2021). However, the association was very small, and due to the study design (teachers reported only on one class), it remains unclear whether classes with different achievement levels contribute to variations in TSE within teachers across classes.

Interrelations between the Class Characteristics

Previous studies pointed to shared variance between some class characteristics. For example, Geerlings et al. (2018) found a gender ratio effect on teacher self-efficacy only when they did not consider student behavior. Similarly, studies indicate that problematic behavior is more likely to be associated with boys than with girls (e.g., Demmer et al., 2017), and higher student engagement is related to higher achievement (e.g., Fung et al., 2018; Lei et al., 2018; Putwain et al., 2018). Therefore, Study 2 applied a methodological approach that simultaneously considers all variables together.

5.2.2 Unique Contributions of Study 2 to Educational Research and Practice

Study 2 examines whether teacher self-efficacy is truly context-specific, as suggested by theory. Its unique contribution lies in empirically testing whether this context-specificity also exists *within* teachers, specifically whether their self-efficacy beliefs vary depending on the classroom in which they teach. The study thus contributes to a validation of the assumed context-specificity and investigates whether class-level characteristics represent meaningful contexts for TSE. By its conceptual replication of Raudenbush et al. (1992) and Ross et al. (1996), the study addresses the underrepresentation of replication studies in educational psychology and strengthens the empirical foundation of TSE (Perry et al., 2022; Plucker & Makel, 2021). If TSE proves to be context-specific within individuals, this will raise important questions about the future design of self-efficacy instruments, which are currently mostly characterized by trait-like, rather than context-specific, questionnaires (Thommen et al., 2022). In addition to the theoretical validation and measurement implications, Study 2 provides future researchers

with insights into which class characteristics are most strongly related to TSE and may warrant further investigation. These class characteristics are also important for practitioners in teacher education and training, as they can be used as points for targeted intervention programs or structural support.

6 Overview of the Studies and Research Questions

Understanding the malleability of teacher self-efficacy is of societal, educational, and psychological relevance given the well-documented positive impacts of a high level of TSE on teacher well-being and instructional quality (Chapter 3). Despite its importance, key theoretical assumptions derived from Bandura's (1997) socio-cognitive theory (Chapter 4), namely the roles of the sources of self-efficacy, the career stage, and the context for the malleability of TSE, remain insufficiently explored in teacher self-efficacy research so far (Chapter 5; Fackler & Malmberg, 2016; Klassen et al., 2011; Kleinsasser, 2014; Kupers et al., 2023; Lauermann et al., 2020; D. B. Morris et al., 2017; Thommen et al., 2022; Tschannen-Moran & Woolfolk Hoy, 2001).

To address these gaps, two empirical studies were conducted (see Table 7). Both studies are deeply informed by socio-cognitive theory and investigate so far underresearched pathways of the triadic reciprocal model (see Figure 1). While both studies share a commitment to open science practices and use the TSES as a common measurement instrument of TSE, they approach the malleability of TSE from two complementary perspectives. Study 1 investigates intentional changes through interventions and focuses on differences between teachers. Study 2 examines natural variation[10] induced by different class characteristics and primarily analyzes differences within teachers. In sum, Study 1 brings clarity to the vast field of TSE interventions, while Study 2 opens the largely unexplored area of within-teacher variation.

Together, these studies offer a multifaceted contribution to self-efficacy theory, educational research, practice, and administration: By empirically testing the roles of the four sources of self-efficacy, the career stage, and class-specific contextual factors, both studies contribute to the validation and refinement of socio-cognitive theory in the field of teacher self-efficacy. For future research, Study 1 provides a benchmark for the average effectiveness of TSE interventions and evaluates the methodological quality of existing studies. Study 2 addresses the underrepresentation of replication studies in educational psychology and offers new insights into the context-specificity of TSE, with implications for the future design of measurement instruments. Furthermore, both studies offer valuable insights for practitioners: Study 1 identifies effective components of interventions, while Study 2 identifies central characteristics of classes that are relevant for shaping TSE. Both findings can inform the design of targeted support programs. Finally, educational administration and policy gain evidence-based benchmarks for evaluating and funding TSE interventions and studies. The findings also support further administrative planning as they identify factors that should be considered in future professional development initiatives.

[10] An explanation for this terminology is presented at the beginning of Chapter 5.

Table 7. Overview of the Two Empirical Studies

	Study 1	*Study 2*
Change Mechanism	• Intentional change	• Natural variation
Level of Analysis	• Between teachers, respectively between studies	• Within teachers
Theoretical Assumption under Investigation	• Malleability through the four sources of self-efficacy • Malleability in different career stages	• Malleability through different contexts
Research Questions	• RQ 1.1: Can interventions promote teachers' self-efficacy? • RQ 1.2: Are there differences in the effects of the interventions depending on the sources of self-efficacy the intervention targets? • RQ 1.3: Are there differences in the interventions' effects regarding the investigated sample's career stage? • RQ 1.4: Are there differences in the effects of the interventions based on the interaction of teachers' career stage and the source targeted by the intervention?	• RQ 2.1: To what extent does teachers' self-efficacy in the three domains of IS, SE, and CM vary across classrooms? • RQ 2.2: To what extent do behavioral, demographic, and achievement class characteristics predict teachers' class-specific self-efficacy regarding the domains of IS, SE, and CM?
Study Design	• Meta-analysis of intervention studies	• Primary study
Interpretability of Time	• Pre- and post-measures available • Duration of intervention and period between measurement points coded and analysed	• Cross-sectional study • No meaningful interpretation of time possible • Different classes align with different time points
Sample	• 115 intervention studies published between 2005 and 2022, including data from 11,284 pre- and in-service teachers	• 26 in-service secondary school teachers, 74 classes, 1326 students from the Netherlands
Measures of Open Science	• Pre-registered (https://osf.io/ev7tf/) • Data, analysis code and materials publicly available	• Pre-registered (https://osf.io/uk2xq) • Analysis code publicly available
(Main) Statistical Analyses	• Effect size estimation (Hedges' g) • Random effects models • Meta regression models	• Intraclass correlation coefficient • Multivariate multilevel regression

7 Study 1: Promoting Teacher Self-Efficacy Through Interventions. A Meta-Analysis and Systematic Review[11]

7.1 The Present Study

Given the vast amount of intervention studies on teacher self-efficacy, a meta-analysis is necessary to investigate the overall effect and to contribute evidence-based knowledge to the question of how lifelong learning opportunities in pre-service teacher education and in-service teacher training should be designed so that teachers can deal with future challenges in a self-efficacious way. The study was further driven by the repeatedly raised research gap to 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) and the theoretical and practical relevant question about the malleability in different career stages. Therefore, the research questions and expected findings are as follows:

Research Question 1.1: Can interventions promote teachers' self-efficacy? A significant positive effect is expected.

Research Question 1.2: 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, interventions that include mastery experiences should produce larger effects than interventions without mastery experiences.

Research Question 1.3: 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, larger effects for pre-service teachers than for in-service teachers are expected.

Research Question 1.4: 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, interventions targeting vicarious experiences or social persuasion are expected 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-

[11] This chapter is based on the published version of Study 1:
Täschner, J., Dicke, T., Reinhold, S., & Holzberger, D. (2024). "Yes, I can!" A systematic review and meta-analysis of intervention studies promoting teacher self-efficacy. *Review of Educational Research*, *95*(1), 3–52. https://doi.org/10.3102/00346543231221499
Minor adjustments to the structure have been made to align the chapter with the overall framework of this dissertation.

Moran et al., 1998; Tschannen-Moran & Woolfolk Hoy, 2007; Weß et al., 2020). Consequently, higher effects of this combination are expected 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. It is expected that pre-service teachers will report higher changes in their teacher self-efficacy after interventions targeting mastery experiences than in-service teachers.

Given that this is the first meta-analysis in this area, investigating additional influencing factors was of particular interest to inform future research and practice. It is hypothesized that studies with a higher study quality (indicated by a study design with randomization and a control group, a reliability coefficient higher than 0.8, and a detailed description of the intervention) will report smaller effect sizes than studies with lower study quality (A. C. K. Cheung & Slavin, 2016). All further potential influencing factors that have been mentioned in the literature, i.e., the intervention aims to solely improve knowledge, moment of reflection included in the intervention, duration and period of the intervention, measure of TSE used in the study, and the domains of teacher self-efficacy, are analyzed as robustness checks without specified hypotheses.

7.2 Method

This meta-analysis follows common guidelines for high-quality systematic reviews and meta-analysis (Alexander, 2020; Page et al., 2021; Pigott & Polanin, 2019). The meta-analysis was pre-registered on OSF (https://osf.io/ev7tf/). Changes that have been necessary after the pre-registration are reported along with explanations and assumed consequences for the meta-analysis in Table A 1 in the appendix. Within the pre-registration process, a priori power analyses were conducted with the R-package *metapower* (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 as eligible for this 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 Heim et al., 2019; Thurm & Barzel, 2020). The a-priori power analysis yielded 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 $d = 0.4$.

7.2.1 Systematic Literature Search

To find as many relevant studies as possible, a variety of search strategies was applied. First, a systematic literature search in the databases FIS Bildung, ERIC, and Web of Science[12] was conducted,

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

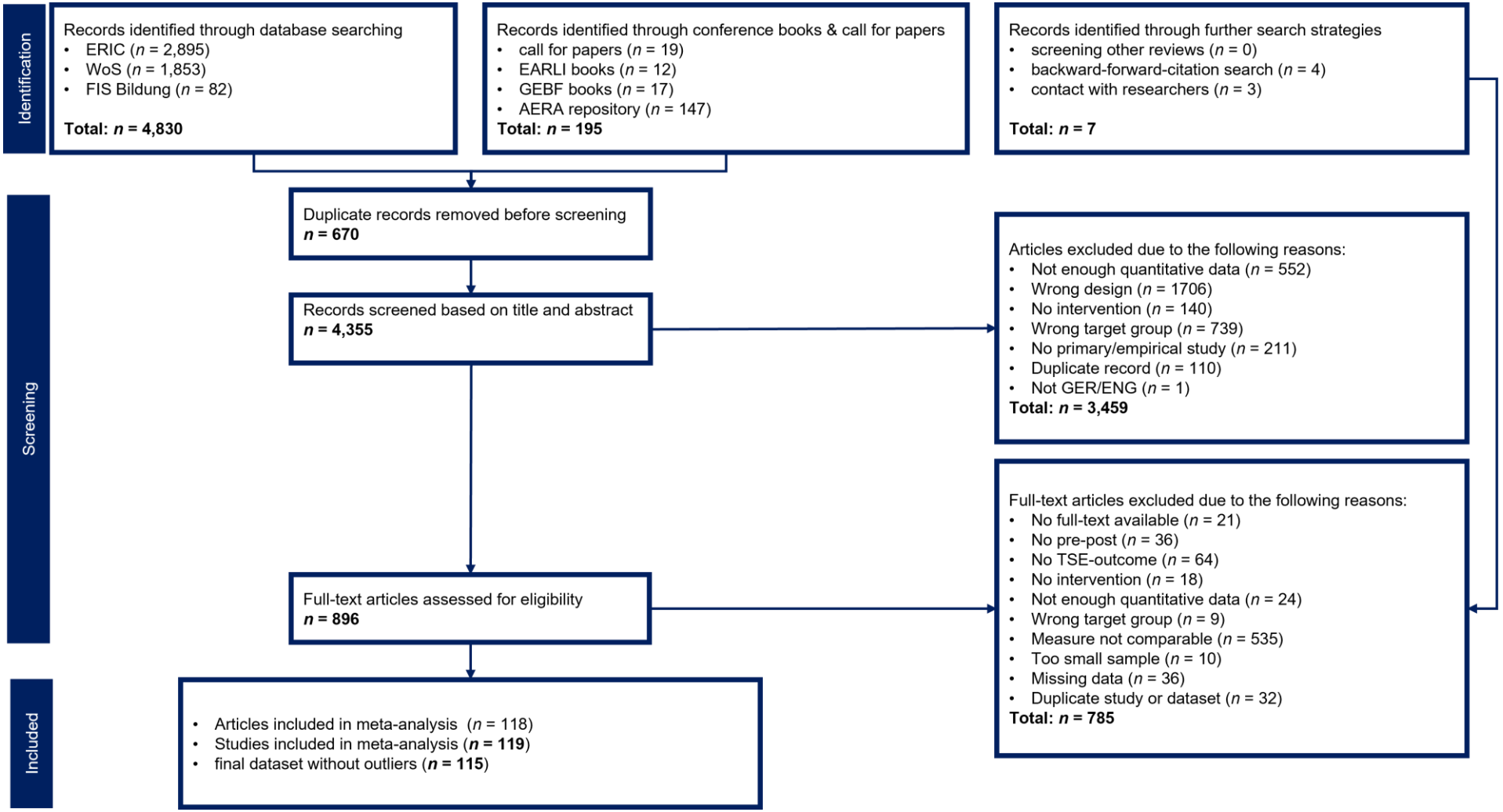
following the guidelines from Gusenbauer and Haddaway (2020) and Siddaway et al. (2019). Depending on the database, the search terms scanned either database-specific descriptors or abstract, title, and keywords (see exact search strings in Table A 2). The final search terms underwent numerous refinements following preliminary scoping searches and consultations with review team members and an information specialist. The structure of the search terms was inspired by the PICO-scheme (i.e., Population, Intervention, Comparison, Outcome; Kugley et al., 2017; Lefebvre et al., 2021): The target population is pre- and in-service teachers. All kinds of interventions, such as programs or trainings are considered. The outcome of interest is self-efficacy. A comparison was not specified within the search term. The search was conducted on December 15, 2020, and delivered 3,526 hits. The search algorithm was kept alive and search alerts until March 1, 2022 were included, 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, conference books and one online paper repository from relevant educational societies in Germany, Europe, and the United States of America, 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)* were screened, resulting in 176 possible studies.

Third, calls for intervention studies were sent out via Twitter and 12 mailing lists (five special interest groups of EARLI, six special interest groups of AERA, and one 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, a backward and forward citation search for all included studies was conducted, using the tool citationchaser (Haddaway et al., 2021), resulting in four possible studies. Finally, the reference lists from 12 previous reviews and meta-analyses related to teacher self-efficacy or intervention studies for teachers were screened with no further study identified (Aloe, Amo, & Shanahan, 2014; Chesnut & Burley, 2015; Gegenfurtner et al., 2013; Gesel et al., 2021; Iancu et al., 2018; Klassen & Kim, 2019; Klassen & Tze, 2014; Kraft et al., 2018; Mok & Staub, 2021; D. B. Morris et al., 2017; Zee & Koomen, 2016).

Figure 3. Flowchart of the Study Selection Process



7.2.2 Screening Process

The flowchart (Page et al., 2021) in Figure 3 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 the meta-analysis meet the inclusion criteria listed in Table 8. 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).

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. After the first 20 studies, interrater reliability was checked, and unclear issues were discussed. If the interrater reliability was lower than 0.8, another round of training with a further 20 studies was done. After sufficient interrater reliability ($\kappa \geq 0.8$) 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 ($\kappa = 0.81$ between coder 1 and 2, $\kappa = 0.83$ between coder 1 and 3). Disagreements were resolved through discussion. Authors for studies that met all the inclusion criteria but lacked statistical information to calculate effect sizes were contacted. Responses were included until April 2022. The final data set of 119 studies exceeds the requisites of 20 studies from the a-priori power analyses.

Table 8. Inclusion Criteria and Corresponding Reasons for Exclusion

Inclusion Criterion	*Corresponding Reason for Exclusion as Displayed in the Flowchart (Figure 3)*
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 Teachers' Sense of Efficacy Scale from Tschannen-Moran and Woolfolk Hoy (2001) to measure teacher self-efficacy a. Studies with scales with at least one subscale comparable to the three subscales of the Teachers' Sense of Efficacy Scale 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 only applied during the full-text screening (second round of screening).

7.2.3 Data Extraction

Two authors extracted the data necessary to characterize all the studies descriptively and to answer the research questions. A detailed coding manual (see https://osf.io/65vfq) provided instructions regarding the characteristics of (1) the publication, (2) the study, (3) the sample, (4) the intervention, and (5) the calculation of the effect sizes. Table 9 displays an overview of the variables extracted in each

category. More detailed information on the (2) study and (4) intervention characteristics is provided below. In cases where information was missing, the variable was coded as not reported.

Table 9. 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

Category	*Variable*	*Possible values*
	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 and emotional 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

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 0.32 (intervention description) to 0.88 (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.

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, it was coded as a second treatment group and not 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, the reliability coefficients were coded as not reported.

Intervention Characteristics

Duration of intervention was coded as stated in the study and then recalculated into hours. The hours were subsequently categorized according to the quartiles into interventions lasting 0.5–6 hours, 7–18 hours, 19–30 hours, and 31–210 hours. *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*, it was 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, the coders rated whether the studies provided *details* on the interventions (yes or no), which was used for the later study quality rating.

The code *only knowledge* was given to interventions focusing only on knowledge acquisition without mentioning any practical exercises or experiences that could be attributed to the sources (e.g., evaluating a lecture in Chao et al., 2017). The role of knowledge for self-efficacy is still debated in the literature (D. B. Morris et al., 2017). Since a main interest was to find out which characteristics of interventions best promote teacher self-efficacy, studies that focused solely on knowledge acquisition were also included. 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," it received the code "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 only watched videos transferring knowledge without modeling (e.g., watching a lecture), the code "no vicarious experiences mentioned" was applied. *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 and emotional reactions* were coded as "mentioned" whenever the intervention mentioned some practice of reducing stress levels or working with attributions (e.g., meditations, autogenic training).

Detailed information on each article's study, sample, and intervention characteristics is provided in Table A 3 and Table A 4.

7.2.4 Statistical Analyses

Computation and Aggregation of Effect Sizes

The meta-analysis combines effect sizes from two different study designs, namely studies with a pre- and post-measure from a one-group study without a control group and studies with a pre-and post-measure from the treatment and the control group. Studies without a control group were included as they may also offer valuable information on the sources targeted in their intervention. 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 (Gresko, 2013; Liaw, 2017), given t-values were transformed into the corresponding effect size measure with the formulas provided by Borenstein et al. (2021) and S. B. Morris and DeShon (2002).

S. B. Morris and DeShon (2002) present theoretical and empirical prerequisites to combine effect sizes from different study designs. Following their guidelines, raw score standardization was used for both effect sizes to avoid any bias induced by the interaction between subject and treatment. Additionally, it was 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, it could be assumed that teacher self-efficacy might not change spontaneously (Bandura, 1997). Therefore, there should be no significant change in the control group, which was also empirically analysed. These preliminary analyses were first conducted to ensure that the effect sizes coming from different study designs are comparable.

As studies reported values from up to three subscales and several treatment groups, there often were multiple correlated effect sizes per study. Following the recommendations from Viechtbauer (n.d.), therefore, two-level random effects models were computed. Additionally, cluster robust inference methods with a small sample correction were applied (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, Cook's distance and DFBETAS values were calculated (Viechtbauer, 2020; Viechtbauer & Cheung, 2010). The interpretation of Cook's distance is similar to the Mahalanobis distance. According to a general rule of thumb, every value larger than $4/n$ (here $4/119 = 0.034$) was considered an influential study. DFBETAS values show how much the overall effect size would change (in standard deviations) when the study is deleted. Any DFBETAS value larger than one was considered as an influential case (Viechtbauer & Cheung, 2010).

Moderator Analyses

Multiple meta-regression models were applied 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. To analyze RQ 1.2, whether the targeted sources of self-efficacy influence the interventions' effects, five meta-regression models were calculated. The first model tested interventions including mastery experiences against interventions targeting other sources but not mastery experiences. The second model evaluated all interventions targeting only one single source against each other ("only mastery experiences" was set as the reference category). In the third to fifth model, interventions targeting only one specific source (i.e., only mastery experiences, only vicarious experiences, or only social persuasion[13]) were tested against interventions targeting this specific source in combination with other sources. To analyze RQ 1.3, whether there are differences in the intervention effects depending on the career stage, one meta-regression model with career stage as a dichotomous moderator was computed. Concerning RQ 1.4, how specific sources interact with different career stages, four meta-regression models were conducted. The first model evaluated whether there is an interaction between interventions targeting vicarious experiences (alone or in combination) and teachers' career stage (pre-service or in-service). The second model 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. A fourth meta-regression model 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) four variables were chosen as indicators to evaluate the quality of the primary studies: The four variables were (1) randomized assignment (yes = 1, no = 0), (2) existence of a control group (yes = 1, no = 0), (3) reliability of the measure ($\geq 0.8 = 1$, $< 0.8 = 0.5$, not reported reliability = 0), and (4) details of the intervention described (yes = 1, no = 0). Following Wedderhoff and Bosnjak (2020) each variable was evaluated individually as a categorical moderator and collectively as part of a sum score (ranging from 0 to 4), which was analyzed as a continuous moderator across five meta-regression models.

Robustness Checks. Six further meta-regression models checked whether the results were robust across the domains of self-efficacy, the used scale of self-efficacy, interventions targeting only

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

knowledge, interventions involving moments of reflection, durations of interventions, and periods between measurement points.

Assessment of Publication Bias. The problem of publication bias was first addressed through the literature search by including unpublished literature, such as dissertations and reports (n = 19). Second, a funnel plot, which prints the effect sizes on the x-axis against the corresponding standard errors on the y-axis, was visually checked for asymmetry. Third, statistical tests were examined: A meta-regression on publication type and an Egger's regression test (1997) checked for indices of publication bias. 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). The additional precision effect estimate with SE (PEESE) uses the sampling variance as the predictor. Based on the conditional PET-PEESE procedure (Carter et al., 2019; Stanley & Doucouliagos, 2014), it was decided which of these two models to use for a cautious interpretation of its intercept as a publication bias-adjusted overall effect size estimate.

Exploratory Analyses

Several studies also reported follow-up measures. Therefore, first an exploratory meta-analysis was calculated that included these effect sizes, too. Secondly, a meta-regression considering the measurement point as moderator allowed to test the stability and consistency of the intervention effects over time from pre to post to follow-up measurement (pre-post effect was set as reference group).

7.3 Results

7.3.1 Checks of Requirements for Aggregation of Effect Sizes

As effect sizes stemmed from different study designs, two empirical checks were conducted to provide evidence that these effect sizes can be combined. The data showed no significant change in the control groups' teacher self-efficacy (g = 0.04, RVE SE = 0.07, CI = [-0.10, 0.17], p = 0.60) 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. There was also no significant difference (p = 0.13) 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, in the following analyses, effect sizes from both study designs were included.

7.3.2 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). These four problematic studies were deleted, and all the following results were calculated without the four influential studies.

7.3.3 Characteristics of Included Studies

One hundred and nineteen studies published between 2005 and 2022 met the inclusion criteria and reported changes in teacher self-efficacy as a consequence of interventions (reference list at the end of the appendix on p. 155). Four studies had to be excluded as they were identified as outliers. Table 10 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, career stage, and sample size). 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 Teachers' Sense of Efficacy Scale (n = 99). The n = 115 included studies reported j = 146 treatment groups, and a total of k = 318 effect size estimates.

Most of the 318 effect sizes originate from the domain classroom management self-efficacy (k = 94), followed by the domains instructional strategies self-efficacy (k = 82), overall teacher self-efficacy (k = 75), and student engagement self-efficacy (k = 67). The interventions lasted between 0.5 and 210 hours, ranging from less than one 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, sources of self-efficacy (k = 260) could be coded. In nine studies, 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 11 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 and emotional 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).

Table 10. 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.

Table 11. Frequencies of Source Combinations in the Interventions

Individual sources mentioned (overall)	*n*	*k*
mastery experiences	78	208
vicarious experiences	59	163
social persuasion	76	200
physiological and emotional 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 and emotional reactions only	0	0
physiological and emotional 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 and emotional reactions (me/ve/sp/pr)	3	7
mastery experiences + vicarious experiences (me/ve)	4	7
vicarious experiences + social persuasion + physiological and emotional reactions (ve/sp/pr)	2	4
mastery experiences + social persuasion + physiological and emotional reactions (me/sp/pr)	2	2
vicarious experiences + physiological and emotional reactions (ve/pr)	1	1
mastery experiences + physiological and emotional reactions (me/pr)	0	0
social persuasion + physiological and emotional reactions (sp/pr)	0	0
mastery experiences + vicarious experiences + physiological and emotional 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*).

7.3.4 Overall Effect of Interventions on Teacher Self-Efficacy (RQ 1.1)

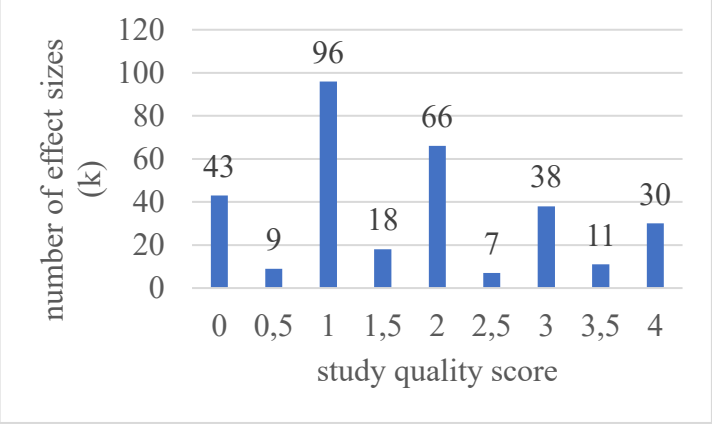
Concerning the first research question about whether interventions can promote teacher self-efficacy at all, a positive significant effect of interventions on teachers' self-efficacy was found (g = 0.47, RVE SE = 0.04, CI = [0.40, 0.54], p < .0001). The significant Q-statistic (Q [df = 317] = 3146.55, p < .0001) indicates that there is a substantial amount of heterogeneity between the studies. The exploratory analysis, which additionally included follow-up measures revealed a similar overall effect (g = 0.45, RVE SE = 0.04, CI = [0.38, 0.52], p < .0001). Teachers' self-efficacy increased from pre- to post-test (g = 0.46, RVE SE = 0.04, CI = [0.39, 0.53], p < .0001) and further from post- to follow-up test (g = 0.13, RVE SE = 0.09, CI = [-0.07, 0.33], p = 0.18), resulting in an overall increase from pre- to follow-up test from g = 0.53 (RVE SE = 0.09, CI = [0.32, 0.74], p < .0001, see Table A 5).

7.3.5 Robustness Checks of Study Quality and Further Influencing Factors

Study Quality

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 A 3 contains the study-level aggregated study quality rating. Figure 4 displays the distribution of the study quality ratings of the 318 effect sizes.

Figure 4. Distribution of Study Quality Ratings Across 318 Effect Sizes



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, see Table 12). 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 = [0.38, 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]). With regard to the single items 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 12). 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 = [0.38, 0.59]) 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]).

Further Influencing Factors

Several robustness checks across different methodological and content variables discussed theoretically as further influencing factors were applied (see Table 13). There were no significant differences between the domains of self-efficacy (p = 0.09) or the used scale of self-efficacy (p = 0.98). There was no significant difference (p = 0.52) between interventions targeting only knowledge (k = 24) as compared with interventions targeting at least one source of self-efficacy (k = 260). However, descriptively, the effect sizes differ clearly in their size. Interventions targeting at least one of the sources of self-efficacy reached considerably larger changes in teacher self-efficacy (g = 0.53, RVE SE = 0.05; CI = [0.42, 0.63]) than interventions targeting only knowledge (g = 0.27, RVE SE = 0.34, CI = [-0.55, 1.09]). Interventions that included a moment of reflection produced significantly greater changes in teacher self-efficacy (g = 0.71, RVE SE = 0.11, CI = [0.50, 0.93] than interventions without such a moment (g = 0.29, RVE SE = 0.10, CI = [0.10, 0.49]). Regarding the interventions' durations, there was no linear effect of hours of intervention on the interventions' effect size (p = 0.32). The intercept (representing the effect of the intervention with 0.5 hours) was g = 0.31 (RVE SE = 0.16, CI = [-0.02, 0.64), and the slope (representing the additional effect of every additional hour) was g = 0.01 (RVE SE = 0.01, CI = [-0.01, 0.02]). However, looking at the frequency distributions of the durations, it becomes clear that a linear regression might not be the best model to evaluate the moderating effect of intervention durations. Therefore, the durations were recategorized according to the quartiles, resulting in interventions lasting 0.5–6 hours (k = 38), 7–18 hours (k = 36), 19–30 hours (k = 29), and 31–210 hours (k = 27). Consequently, a significant difference between the different categories of duration arose (p = 0.04). The biggest changes were evident in interventions lasting 19–30 hours (g = 0.69, RVE SE = 0.15, CI = [0.37, 1.02]) and interventions lasting 7–18 hours (g = 0.61, RVE SE = 0.11, CI = [0.35, 0.86]). The smallest changes appeared in interventions lasting 0.5–6 hours (g = 0.25, RVE SE = 0.06,

CI = [0.12, 0.39]). The effects of interventions lasting 31–210 hours were $g = 0.43$ (RVE $SE = 0.1$, CI = [0.23, 0.64]). The period between measurement points captured in weeks did not significantly moderate the effects ($p = 0.06$). Periods of less than one week on average had an effect of $g = 0.67$ (RVE $SE = 0.11$, CI = [0.44, 0.88]). With every additional week, the effect decreased on average at $g = -0.01$ (RVE $SE = 0.01$, CI = [-0.02, 0.00]).

Table 12. Results of Meta-Regression Analyses Regarding Study Quality

Variable	*n*	*k*	*g*	*SE*	*95% CI*	*df*	*Q*	*p*
Study quality sum score	115	318					3018.20***	.18
intercept			0.37	0.07	[0.22, 0.52]	52.38		<.0001
slope			0.05	0.04	[-0.03, 0.13]	51.11		.18
Study quality: assignment	115	318					3092.26***	.35
not randomized	87	238	0.45	0.04	[0.37, 0.53]	84.76		<.0001
randomized	28	80	0.09	0.10	[-0.10, 0.28]	43.33		.35
Study quality: control group	115	318					3146.50***	.40
no	79	231	0.49	0.04	[0.40, 0.57]	77.10		<.0001
yes	36	87	-0.07	0.08	[-0.22, 0.09]	63.17		.40
Study quality: reliability	115[a]	318					3085.96***	.66
not reported	36	97	0.48	0.07	[0.34, 0.61]	18.45		<.0001
lower than 0.8	29	45	-0.07	0.11	[-0.31, 0.17]	14.5		.53
0.8 and more	73	176	0.01	0.08	[-0.18, 0.20]	9.04		.89
Study quality: intervention described with details	115	318					3004.39***	<.01
no	42	134	0.35	0.05	[0.25, 0.44]	40.62		<.0001
yes	73	184	0.20	0.07	[0.06, 0.34]	87.36		<.01

Note. n = number of studies; k = number of effect sizes. 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 13. Results of Meta-Regression Analyses Regarding the Robustness Checks

Moderator variables	*n*	*k*	*g*	*SE*	*95% CI*	*df*	*Q*	*p*
Used scale	115	318					3146.10***	.98
TSES	99	289	0.47	0.04	[0.39, 0.55]	96.13		< .0001
other	16	29	0.00	0.10	[-0.20, 0.21]	20.02		.98
Domain of TSE	115[a]	318					3121.42***	.09
overall	56	75	0.50	0.04	[0.42, 0.58]	94.22		<.0001
classroom management self-efficacy	74	94	-0.04	0.02	[-0.09, 0.01]	14.23		.13
instructional strategies self-efficacy	65	82	-0.02	0.03	[-0.08, 0.04]	14.20		.41
student engagement self-efficacy	56	67	-0.07	0.03	[-0.14, 0.00]	16.46		.06
Knowledge vs. Sources	103[a]	284					2249.65***	.52
only knowledge targeted in interventions	9	24	0.27	0.34	[-0.55, 1.09]	6.1		.45
at least one source targeted in interventions	97	260	0.25	0.37	[-0.68, 1.18]	5.34		.52
Moment of reflection	105[a]	294					2288.25***	<.05
no	53	152	0.29	0.09	[0.10, 0.49]	32.05		<.01
yes	56	142	0.42	0.18	[0.00, 0.84]	8.18		<.05
Intervention duration linear	48	130					924.32***	.32
intercept			0.31	0.16	[-0.02, 0.64]	25.93		.07
slope			0.01	0.01	[-0.01, 0.02]	2.92		.32
Intervention duration categorized	48[a]	130						<0.05
0.5 to 6 hours	11	38	0.25	0.06	[0.12, 0.39]	8.99		<0.01
7 to 18 hours	13	36	0.35	0.14	[0.00, 0.70]	5.53		< 0.05
19 to 30 hours	12	29	0.44	0.16	[0.10, 0.78]	19.69		<0.05
max. 210 hours	13	27	0.18	0.11	[-0.06, 0.42]	20.49		.13
Period between measurement points in weeks	96	273					2654.82***	.06
intercept			0.67	0.11	[0.44, 0.88]	45.97		<0.0001
slope			-0.01	0.01	[-0.02, 0.00]	25.75		.06

Note. n = number of studies; k = number of effect sizes. 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.

***$p < .001$.

7.3.6 The Role of the Sources for Interventions' Success (RQ 1.2)

First, it was evaluated whether mastery experiences, in general, 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 = 0.40), differ from interventions without mastery experiences (g = 0.54, RVE SE = 0.05, CI = [0.43, 0.66])[14]. In a next step, interventions targeting only one source were analyzed. Interventions targeting either only mastery experiences (k = 35), only vicarious experiences (k = 17), or only social persuasion (k = 14) did not significantly differ from each other (p = 0.16).[15] Descriptively, interventions targeting only vicarious experiences showed the largest effects (g = 0.83, RVE SE = 0.20, CI = [0.33, 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, it was investigated whether the interventions' effects differ between interventions targeting only one specific source and interventions targeting this source in combination with other sources. There was no significant difference for all three models contrasting mastery experiences, vicarious experiences, and social persuasion as single sources and in combination (p = 0.60, p = 0.53, p = 0.12). 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 = 0.34, 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 14. Average effect sizes for each source and source combination as a subgroup are displayed for further information in Table 15.

[14] In the text we report the absolute effect size, the according standard error and the confidence interval for each category. This was done by making the respective category the reference category in the meta-regression.
[15] No intervention targeted only physiological and emotional reactions (see Table 11).

Table 14. Overall Effect of Interventions and Results From Meta-Regression Analyses Regarding Career Stage and Targeted Sources (RQ 1.1-1.3)

	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								
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
Mastery experiences	97 [a]	260					2004.9472***	0.40
targeted in intervention	78	208	0.51	0.05	[0.42, 0.61]	78.87		<0.0001
not targeted in intervention	22	52	0.03	0.03	[-0.07, 0.13]	2.9		0.40
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		<0.01
in combination	70	186	0.16	0.09	[-0.06, 0.39]	5.85		.12

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 15. Average Effect Size for Each Source (Combination)

source combinations[a]	*n*	*k*	*g*	*SE*	*df*	*CI*	*p*
mastery experiences + vicarious experiences + social persuasion	40	111	0.51	0.07	38.34	[0.38, 0.65]	<.0001
mastery experiences + social persuasion	21	46	0.54	0.14	19.98	[0.25, 0.84]	<.0001
vicarious experiences + social persuasion	6	16	0.34	0.02	2.34	[0.26, 0.42]	<0.01
mastery experiences + vicarious experiences + social persuasion + physiological and emotional reactions	3	7	0.25	0.17	1.72	[-0.59, 1.09]	0.29
mastery experiences + vicarious experiences	4	7	0.43	0.07	2.34	[0.17, 0.70]	<0.05
vicarious experiences + social persuasion + physiological and emotional reactions	2	4	0.70	0.01	1	[0.51, 0.88]	<0.05
mastery experiences + social persuasion + physiological and emotional reactions	2	2	1.34	0.18	1	[-0.93, 3.72]	0.08
Alone vs. in combination							
mastery experiences only	14	35	0.39	0.08	12.19	[0.20, 0.57]	<0.001
mastery experiences in combination	64	173	0.52	0.06	65.49	[0.40, 0.64]	<0.0001
vicarious experiences only	7	17	0.98	0.26	5.92	[0.35, 1.61]	<0.01
vicarious experiences in combination	52	146	0.47	0.05	52.05	[0.37, 0.57]	<0.0001
social persuasion only	6	14	0.37	0.06	3.75	[0.20, 0.53]	<0.01
social persuasion in combination	70	186	0.51	0.06	68.43	[0.40, 0.63]	<0.0001
physiological and emotional reactions in combination	8	14	0.64	0.18	6.84	[0.22, 1.06]	<0.01

Note. These are the averaged effect sizes for each source, respectively, source combinations. Sources or source combinations that only appeared once or never are not reported here.

a. Source combinations are displayed in descending order based on the frequency of the effect sizes.

7.3.7 Promoting Teacher Self-Efficacy in Different Career Stages (RQ 1.3)

Four studies had mixed samples with pre-service *and* in-service teachers without stating separate effect sizes for each group. As it is not possible to draw any conclusions for 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 14, the results showed no significant differences between pre-service and in-service teachers in their gains in teacher self-efficacy after the interventions ($p = 0.50$). 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]).

7.3.8 The Interaction of Career Stages and Sources (RQ 1.4)

In order to examine whether different sources of self-efficacy may be differently important for teachers depending on their career stage, four interaction models were analyzed (see Table 16). 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 ($p = 0.92$). 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 $SE = 0.27$, 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 ($p = 0.01$). 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 $SE = 0.08$, CI = [0.29, 0.59]; for in-service teachers: $g = 0.55$, RVE $SE = 0.08$, CI = [0.40, 0.70]). 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, it was evaluated whether pre-service teachers benefit more from the combination of mastery experiences and social persuasion than in-service teachers. No significant moderating effect was found ($p = 0.24$). 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]).

There was no interaction effect in the fourth model evaluating the role of interventions targeting mastery experiences alone or in combination for the two different career stages (p = 0.33). 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]).

Table 16. Results of Meta-Regression Analysis Regarding the Interaction Between Career Stages and Targeted Sources (RQ 1.4)

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 14 because studies with a mixed sample of pre-service and in-service teachers were ignored in this analysis.

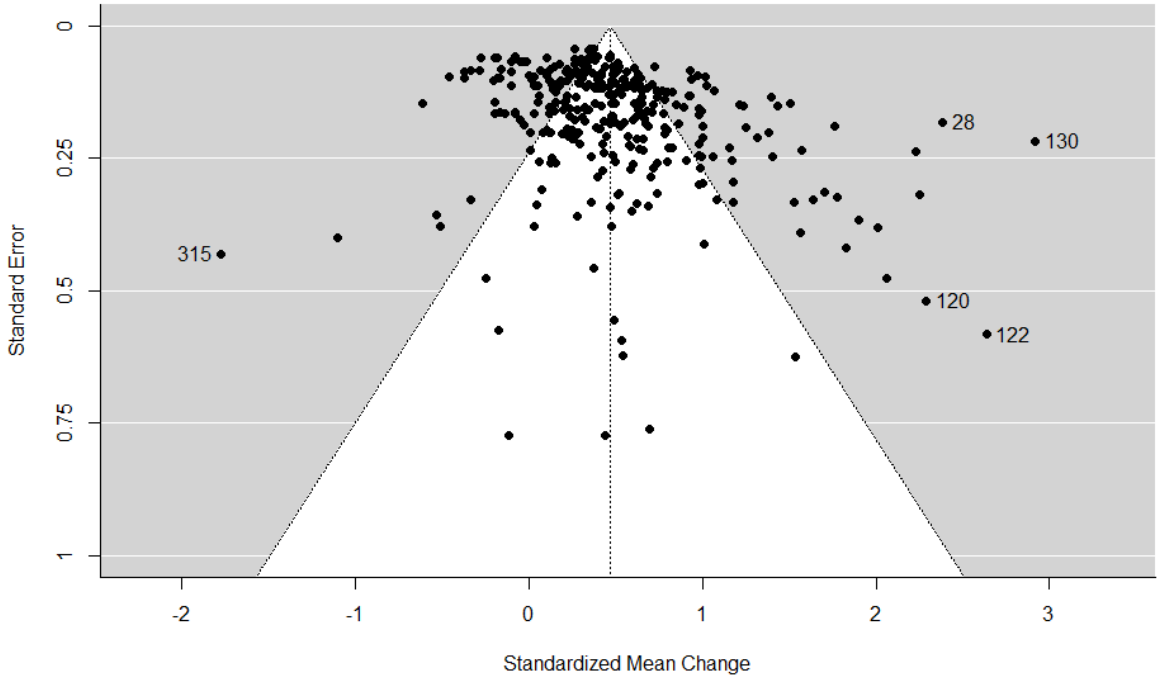
a. Because some studies are represented in more than one category of this moderator, the total sum of the categories does not add up to the number given here.

***$p < .001$.

7.3.9 Publication Bias

The funnel plot in Figure 5 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.

Figure 5. Funnel Plot of the Effect Sizes for Interventions Promoting Teacher Self-Efficacy



Note. The five most extreme effect sizes are indicated by their respective number.

To retrieve statistical information about the funnel plot's symmetry, precision-effect tests (Stanley & Doucouliagos, 2014) were applied although they could not take into account the dependency among the correlated effect size. The Egger regression test (1997) was significant ($p < 0.0001$), indicating asymmetry in the funnel plot. Following, the PEESE (precision-effect estimate with SE) was conducted. The PEESE was significant as well ($p < 0.0001$). Its limit estimate can be cautiously interpreted as the publication bias-adjusted average true effect ($b = 0.29$, CI = [0.24, 0.33]).

There was no significant difference in the effects based on the publication types ($p = 0.66$). The publication type "other" produced the biggest effect sizes ($g = 0.55$, 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]).

8 Study 2: Exploring the Context-Specificity of Teacher Self-Efficacy[16]

8.1 The Present Study

Bandura's (1997) model of the triadic reciprocal causation suggests that different contexts (= environmental factors) can impact the level of self-efficacy. In the field of teacher self-efficacy research, such contextual influences have received limited attention (Fackler & Malmberg, 2016; Klassen et al., 2011; Thommen et al., 2022). When context factors are considered, they are typically studied with research designs that analyze differences between teachers (e.g., principal's leadership style in Bellibaş & Liu, 2017). Study 2, therefore, investigates whether such context-specific variations also exist *within* teachers. Furthermore, it examines whether characteristics of the classes—such as students' behavior or achievement—serve as meaningful predictors of teacher self-efficacy at the classroom level.

The research questions under investigation are:

> **RQ 2.1: To what extent does teachers' self-efficacy in the three domains of IS, SE, and CM vary across classrooms?**
>
> **RQ 2.2: To what extent do behavioral, demographic, and achievement class characteristics predict teachers' class-specific self-efficacy regarding the domains of IS, SE, and CM?**

Since little research has been conducted in this area to date, an exploratory approach was chosen and no hypotheses were specified.

8.2 Method

The study was pre-registered on Open Science Framework (OSF) prior to data analysis (https://osf.io/uk2xq). Table A 6 reports deviations from the pre-registration, including the corresponding rationales and possible consequences. One genuine deviation from the pre-registration is that the final sample size of teachers was considerably smaller than the estimated sample size. Therefore, issues regarding statistical power and possible overestimation of significance (e.g., Hox & McNeish, 2020; McNeish & Stapleton, 2016) were considered in the interpretation of the results.

[16] This chapter is based on an unpublished working paper co-authored by Janina Täschner, Doris Holzberger, Terrence D. Jorgensen and Marjolein Zee. It has been reviewed and revised for this dissertation.

8.2.1 Participants and Procedure

Data collection has been registered at the ethics committee of the University of Amsterdam under 2021-CDE-13840. The data were collected between March and June 2022 by two master's students as part of their master's theses. Students used their personal networks and school lists to recruit teachers. All participants gave their consent to participate.

In total, 26 teachers from schools in urban areas in the central part of the Netherlands gave their consent and completed questionnaires about their TSE for three of their classes[17]. There was no missing data from the teachers. Additionally, 1,326 students in these classes answered questionnaires regarding their demographics, behavior, and perceptions of the corresponding teacher. The only missing data for students appeared within the engagement questionnaire. Forty-two students (3.2%) did not answer the questions regarding their engagement at all, so they were excluded from the calculation of the class engagement means. Accordingly, in the dataset analyzed, there were no missing values at the class level. The teachers were, on average, 41.35 years old (SD = 13.06) and had a mean teaching experience of 18.35 years (SD = 12.21). In total, 65.4% were male. Most of the teachers worked five days per week as teachers. The average number of working hours per week was 35.19 (SD = 8.4; range = 18 – 55 hours). For 24 of the 26 teachers, the three classes were in the same subject they taught; the remaining two teachers had one class that differed in subject from the other two classes (see Table A 7).

The 1,326 students belonged to 74 different classes. They were, on average, 15.27 years old (SD = 1.53). In total, 51.6% identified as male, 46.5% as female, and 1.9% as non-binary or different. 77.0% reported Dutch as their primary ethnicity, 12.8% reported a Turkish, Moroccan, or Surinamese ethnic background and 10.01% a different background. Of all students, 71.0% spoke one language at home, 29.0% reported growing up multilingually.

8.2.2 Instruments

To avoid common method bias, a multi-informant approach was used in this study, where some variables were answered by students and others by the participating teachers (Podsakoff et al., 2012).

Teachers' Classroom-Specific Sense of Self-Efficacy

Teachers' perceptions of their self-efficacy in each of the three selected classrooms were measured with the Dutch version of the Teachers' Sense of Efficacy Scale on a Likert scale ranging from 1 (*not at all*) to 7 (*absolutely well*) (TSES; Tschannen-Moran & Woolfolk Hoy, 2001; Zee, Koomen, et al., 2016). The class-specific adaption forced teachers to answer each item according to one specific class (see Table A 8). The TSES investigates TSE across the three domains of IS, SE, and CM. A confirmatory factor analysis supported the three-factor solution of the Dutch version of the TSES

[17] However, one teacher only reported on one class and two teachers only on two classes.

(https://osf.io/zxcvt/). Following Geldhof et al. (2014), level-specific reliability estimates were calculated using multilevel confirmatory factor analysis (see Table 17). The within-level alphas indicated acceptable to good internal consistency of the subscales on the within-level. The between-level alphas indicated a high reliability of the subscales on the between-level. However, the small cluster size may have led to an overestimation of the between-level reliabilities (Geldhof et al., 2014).

Table 17. Item Examples and Internal Consistencies of the Used Scales

Scale	*No. of Items*	*Item example*	*Within-level α*	*Between-level α*
TSE for instructional strategies	4	How well can you provide appropriate challenges for students who are good learners?	0.69	0.80
TSE for classroom management	4	How well do you succeed in getting students to follow class rules?	0.80	0.91
TSE for student engagement	4	To what extent do you succeed in helping students begin to find learning important?	0.76	0.91
Class misbehavior	3	How often do you experience negative interactions with students in this class?	0.90	0.85
Behavioral engagement	6	During this teacher's lesson, I listen very well.	0.84	0.96
Emotional engagement	5	When we work on something in this teacher's classes, I get bored.	0.75	0.94

Behavioral Composition

Teachers' perceived amount of class misbehavior was measured using a Dutch version of a scale originally developed by Tsouloupas et al. (2010) and adapted by Zee, de Jong, and Koomen (2016). Responses were rated on a scale from 1 (never) to 5 (almost always) (see Table A 9). Internal consistency of the scale was excellent (see Table 17).

To evaluate students' engagement in class, the students answered the short version of the Engagement Versus Disaffection with Learning Scale (Skinner et al., 2008; Zee & Koomen, 2020; see Table A 10; Likert scale: 1 (*no, that is not true*) to 5 (*yes, that is definitely true*)). The *Behavioral Engagement* subscale captures students' attention in a typical lesson taught by a specific teacher. The *Emotional Engagement* subscale measures students' attitudes during a typical lesson with the corresponding teacher. Similar to previous studies using this scale, the data from this study supported a two-factor model after removing one Emotional Engagement item with a low factor loading (see, for example, Roorda et al., 2019). Internal consistency scores were acceptable to excellent for both subscales (see Table 17).

Demographic Composition

The percentage of boys per class was calculated based on students' responses about their gender. At the time of data collection (end of the school year), teachers had good knowledge of the classes. They reported the number of students with a migration background per class.

Achievement Composition

Based on students individual grades (1 (*very poor*) to 10 (*excellent*)), a mean grade per class was calculated.

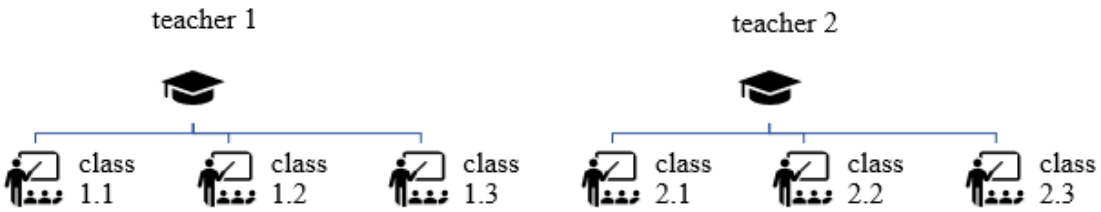
Teacher-Level Covariates

According to current research, demographic characteristics of teachers such as gender, age, and years of experience have little to no influence on the level of teacher self-efficacy (e.g., H. Y. Cheung, 2008; Fackler et al., 2021; Gkolia et al., 2016; Klassen & Chiu, 2010; Lauermann & Berger, 2021; Lazarides et al., 2020; Malinen et al., 2024; Perera et al., 2019; Raudenbush et al., 1992; Ross et al., 1996; Skaalvik & Skaalvik, 2011; Zee et al., 2018). Nevertheless, to control for possible confoundings on the between-level, teachers' gender (0 = male; 1 = female), age, and years of teaching experience were included as teacher covariates.

8.2.3 Data Analyses

The study approach resulted in a nested data structure, depicted in Figure 6, with classes (Level 1) nested within teachers (Level 2). Therefore, multilevel regression analyses were applied to answer the research questions using the software M*plus* 8.3 (Muthén & Muthén, 1998–2017). Manually constructed scale means for all outcome and predictor variables measured with a scale were used as this approach has been shown to be especially suitable for multilevel studies with small cluster sizes and a small number of clusters (Asparouhov & Muthén, 2019; Lüdtke et al., 2008).

Figure 6. Structure of the Dataset



To address the first research question, intraclass correlations (ICC) based on an unconditional random-intercept model for TSE for IS, CM, and SE were calculated to determine the variance components within and between teachers. The ICC (ρ) indicates how much of the total variance in TSE

is attributable to variance between teachers. ICC values range from 0 to 1. The closer to 1, the more different the TSE estimations are between teachers, but the more similar they are within teachers.

$$\rho = \frac{\sigma_{v_0}^2}{\sigma_{v_0}^2 + \sigma_r^2}$$

- $\sigma_{v_0}^2$ = variance between teachers
- σ_r^2 = variance within teachers

To address the second research question, one multivariate random intercept model with the three domains of TSE as outcomes was constructed. All behavioral, demographic, and achievement characteristics from the classes were included as predictors on the within-level (class-level, Level 1) and the between-level (teacher-level, Level 2). To decompose the effects of these Level-1 predictors into within- and between-teacher components, latent mean centering was applied, which has been recommended for multilevel studies with small cluster sizes (Asparouhov & Muthén, 2019; Lüdtke et al., 2011). Accordingly, each between-level coefficient is interpreted as the degree to which two teachers' average TSE for one domain (across classes they teach) is expected to differ when their class-level predictor's average differs by 1 unit. Likewise, each within-level coefficient is interpreted as the degree to which a teacher's TSE for a *particular class* is expected to differ from their own average across classes they teach, when the class-level predictor for that class differs by 1 unit from the teacher's latent average. For example, a within-level slope represents the degree to which a teacher's TSE is expected to differ between 2 classes they teach, when their class-level predictor differs by 1 unit from the teacher's latent average. Additionally, between-level covariates — teacher gender, age, and years of experience — were included. An exemplary full multilevel model for the domain TSE for CM is depicted below:

$$\begin{aligned}\mathrm{TSM(CM)}_{ij} = {} & \beta_{0j} + \beta_1\left(\mathrm{grade}_{ij} - \overline{\mathrm{grade}}_{.j}\right) + \beta_2\left(\mathrm{gen}_{ij} - \overline{\mathrm{gen}}_{.j}\right) + \beta_3\left(\mathrm{mig}_{ij} - \overline{\mathrm{mig}}_{.j}\right) \\ & + \beta_4\left(\mathrm{emo}_{ij} - \overline{\mathrm{emo}}_{.j}\right) + \beta_5\left(\mathrm{beh}_{ij} - \overline{\mathrm{beh}}_{.j}\right) + \beta_6\left(\mathrm{misb}_{ij} - \overline{\mathrm{misb}}_{.j}\right) + \beta_7\overline{\mathrm{grade}}_{.j} \\ & + \beta_8\overline{\mathrm{gen}}_{.j} + \beta_9\overline{\mathrm{mig}}_{.j} + \beta_{10}\overline{\mathrm{emo}}_{.j} + \beta_{11}\overline{\mathrm{beh}}_{.j} + \beta_{12}\overline{\mathrm{misb}}_{.j} + \beta_{13}\mathrm{tgen}_j + \beta_{14}\mathrm{age}_j \\ & + \beta_{15}\mathrm{exp}_j + \varepsilon_{ij}.\end{aligned}$$

The subscript *i* refers to individual classes (Level 1), *j* refers to teachers (Level 2). The random intercept of TSE for CM (β_{0j}) varies across teachers, accounting for similarity of TSE for CM among classes taught by the same teacher. The multivariate model includes analogous coefficients for the other TSE domains, and residual covariances among the domains are also estimated (available at https://osf.io/zxcvt, but not reported as they are not of interest).

All parameter estimates in this model were based on a maximum likelihood estimation (MLR estimator in M*plus*) with robust standard errors. The M*plus* syntax files are provided on OSF (https://osf.io/zxcvt).

8.3 Results

8.3.1 Descriptive Statistics

Table 18 shows the means, standard deviations, ICCs, and correlations of all predictor and outcome variables on the within- and the between-level.

With regard to the class characteristics, the majority of the variance was located within teachers for the variables behavioral engagement (ICC = .24), class misbehavior (ICC = .15), and percentage of boys (ICC = .16). The opposite was true for migration background (ICC = .94), where most of the variance appeared between teachers. Finally, for emotional engagement and mean grade, the variance was almost equally split between and within teachers (ICC = .50, respectively, .54).

At the within-teacher level, emotional engagement, class misbehavior, and mean grade were correlated with TSE for IS, CM, and SE in the expected directions. All these correlations were statistically significant, except for the correlation between emotional engagement and TSE for IS. None of the other class characteristics were significantly correlated with TSE. At the between-teacher level, virtually none of the class characteristics were significantly correlated with TSE. Also, none of the covariates were significantly related to TSE. As the three domains of TSE and several of the class characteristics were highly correlated with each other, a multivariate regression model was necessary to partial out potential confounders and investigate the individual influences of each class characteristic.

Figure 7 displays for each teacher their TSE means per class and per domain. The visual inspection suggests three types of TSE variations across the teachers: The first type shows a clear profile of the three domains of TSE that is visible in all three classes with clear quantitative differences between the classes (e.g., teachers 18 and 19). The second type shows no clear differences between the classes (e.g., teachers 3,7, 15, 2). For the third type no clear profile of the three domains is visible, the level of self-efficacy in each domain seems to depend more on the class (e.g., teacher 1, 14, 25).

Table 18. Descriptive Statistics and Correlations

	Variable	1	2	3	4	5	6	7	8	9	10	11	12
Behavioral Characteristics	1. Emotional Engagement[a]		.65*	−.34	−.18	.28	.29	.35	.34	.17	−.12	−.11	−.01
	2. Behavioral Engagement[a]	.64**		−.01	−.02	.21	.31	.28	.42	.24	−.32	−.29	−.16
	3. Class Misbehavior[b]	−.27	−.04		.57	−.18	−.34	.14	−.05	.14	.08	−.74	−.82
Demographic Characteristics	4. Migrant Background[b]	.19	.58**	.19		−.50	−.64*	.33	.16	.23	.01	−.41**	−.42**
	5. Percentage of Boys[a]	.20	.08	.15	−.02		.30	−.21	−.27	−.13	−.48	.17	.20
Achievement Characteristics	6. Mean Grade[a]	.13	.11	−.39**	−.07	−.09		.02	.09	−.09	−.03	.19	.30
Teacher Self-Efficacy	7. TSE for IS	.21	.00	−.47**	−.14	−.12	.32*		.80	.59	.15	.10	.04
	8. TSE for SE	.36*	.06	−.59**	−.20	−.14	.53**	.65**		.74	−.05	.23	.21
	9. TSE for CM	.38*	.18	−.67**	−.10	−.21	.39**	.56**	.66**		−.10	.19	.17
Teacher Covariates	10. Teacher Gender[c]	-	-	-	-	-	-	-	-	-		−.11	−.14
	11. Teacher Age	-	-	-	-	-	-	-	-	-	-		.97**
	12. Years of Experience	-	-	-	-	-	-	-	-	-	-	-	
Descriptive Statistics	*M*	3.77	3.67	1.87	26.78	52.09	6.67	5.47	5.17	5.75	-	41.35	18.35
	SD_w	0.29	0.29	0.69	8.05	17.34	0.49	0.40	0.60	0.60	-	-	-
	SD_b	0.29	0.16	0.29	32.28	7.36	0.54	0.54	0.54	0.55	-	12.81	11.98
	ICC	.50	.24	.15	.94	.15	.54	.64	.45	.45	-	-	-

Note. Below the diagonal, the correlations are presented on the within-level. The correlations above the diagonal are on the between-level.

[a]This variable has been reported by students. [b]This variable has been reported by teachers. [c]0 = male, 1 = female.

** two-tailed $p < .01$ * two-tailed $p < .05$, but p values may not be trustworthy because the correlations estimated in M*plus* exceeded the number of Level-2 observations (yielding a warning about standard errors being untrustworthy).

Figure 7. Means of Teacher Self-Efficacy Within Teachers Across Domains and Classes

8.3.2 Intraclass Correlation

The ICCs based on the unconditional random-intercept model were 0.40 for TSE for SE, 0.61 for TSE for IS, and 0.37 for TSE for CM. As estimates of level-2 variance based on small cluster sizes have been found to be overestimated (McNeish & Stapleton, 2016), it can be cautiously concluded that at least 60%, 39%, and 63% of the variance in TSE in these respective domains is due to differences within teachers themselves, respectively, due to differences between their classes.

8.3.3 Class Characteristics as Predictors for Teacher Self-Efficacy

The multivariate random intercept model explained between 27% (TSE for IS) and 54% (TSE for SE and TSE for CM) of within-teacher variance, that is, teachers' class-specific variations in TSE (see Table 19).

From the two student-reported dimensions of engagement (emotional and behavioral engagement), only emotional engagement was significantly associated with the domain TSE for SE. When emotional engagement in a specific class was one unit higher than a teacher's latent average across their classes, the TSE for SE in that class increased by 0.80 scale points ($p = .000$). Additionally, perceived class misbehavior was significantly and negatively related to all three domains of TSE (IS: $b = -0.19$, $p = .041$, SE: $b = -0.28$, $p = .004$, CM: $b = -0.45$, $p = .000$). This means that the more a class deviated in its misbehavior from a teacher's other classes, the more negatively teachers rated their self-efficacy for that particular classroom.

The percentage of students with a migration background was not significantly related to any of the three domains of TSE. However, the percentage of boys significantly predicted TSE for CM. With 1% more boys in a class than a teacher's average, the TSE for CM dropped by .01 scale points ($p = .019$). In addition, the mean grade level of a class was significantly positively related to TSE for SE ($b = 0.42$, $p = .003$). On the between-level none of the class characteristics and none of the covariates was significantly related to TSE for SE and TSE for CM. However, teachers reported significantly higher TSE for IS when students in general reported more behavioral engagement ($b = 1.48$, $p = .047$) or when the teachers themselves were female ($b = 0.41$, $p = .024$) or older ($b = 0.15$, $p = .010$). Between-level effects were not the primary focus of our research questions, but they were necessary to estimate as a side effect of latent mean centering (Asparouhov & Muthén, 2019; Lüdtke et al., 2011).

Table 19. Multivariate Model Using Class Characteristics as Predictors for Teacher Self-Efficacy (RQ 2.2)

Predictor	Class-specific TSE for IS		Class-specific TSE for SE		Class-specific TSE for CM	
	b (SE)	*b** (SE)	*b* (SE)	*b** (SE)	*b* (SE)	*b** (SE)
Within-level						
Behavioral characteristics						
Emotional engagement[a]	0.30	0.22	0.80*	0.38*	0.50	0.24
	(0.19)	(0.15)	(0.22)	(0.11)	(0.35)	(0.16)
Behavioral engagement[a]	−0.20	−0.14	−0.36	−0.17	0.12	0.06
	(0.19)	(0.14)	(0.27)	(0.13)	(0.30)	(0.15)
Class misbehavior[b]	−0.19*	−0.33*	−0.28*	−0.32*	−0.45*	−0.51*
	(0.11)	(0.17)	(0.10)	(0.12)	(0.09)	(0.09)
Demographic characteristics						
Migrant background[b]	0.00	−0.03	−0.01	−0.09	−0.01	−0.08
	(0.01)	(0.12)	(0.01)	(0.11)	(0.01)	(0.13)
Percentage of boys[a]	0.00	−0.08	0.00	−0.13	−0.01*	−0.18*
	(0.00)	(0.13)	(0.00)	(0.12)	(0.00)	(0.08)
Achievement characteristics						
Mean grade[a]	0.14	0.17	0.42*	0.35*	0.15	0.12
	(0.12)	(0.14)	(0.15)	(0.13)	(0.11)	(0.09)
Between-level						
Behavioral characteristics						
Emotional engagement[a]	0.99	0.53	1.66	0.88	2.12	1.10
	(0.72)	(0.39)	(2.22)	(1.22)	(3.66)	(2.09)
Behavioral engagement[a]	1.48*	0.46*	1.38	0.42	0.17	0.05
	(0.88)	(0.27)	(2.39)	(0.71)	(4.22)	(1.26)
Class misbehavior[b]	0.47	0.25	2.93	1.55	5.14	2.67
	(1.69)	(0.91)	(5.37)	(2.91)	(8.65)	(4.93)
Demographic characteristics						
Migrant background[b]	0.02[e]	0.97*	−0.01	−0.41	−0.01	−0.84
	(0.01)	(0.57)	(0.03)	(1.78)	(0.05)	(2.95)
Percentage of boys[a]	0.00	0.03	−0.06	−0.84	−0.03	−0.40
	(0.03)	(0.37)	(0.08)	(1.12)	(0.11)	(1.55)
Achievement characteristics						
Mean grade[a]	0.64[e]	0.64*	−0.18	−0.17	−0.85	−0.83
	(0.39)	(0.37)	(1.09)	(1.09)	(1.75)	(1.85)
Teacher Covariates						
Teacher gender[c]	0.41*	0.37*	−0.14	−0.12	0.01	0.01
	(0.21)	(0.20)	(0.52)	(0.46)	(0.79)	(0.68)
Teacher age	0.15*	3.58*	0.04	0.87	−0.07	−1.52
	(0.07)	(1.52)	(0.18)	(4.20)	(0.29)	(6.98)
Years of experience	-0.13	-2.89	0.03	0.75	0.18	3.84
	(0.09)	(1.97)	(0.27)	(6.02)	(0.43)	(10.07)
Pseudo-R^2 statistics						
R^2_{within}		.27* (.11)		.54* (.10)		.54* (.09)
$R^2_{between}$		.97* (.26)		.88 (.57)		.74 (1.01)

Note. All predictors are centered at the latent teacher-level mean. The focus of the present study is on the within-level predictors, between-level predictors are included as a by-product of using latent-mean centering, but between-level effects are not the focus of the current study. *b* are the unstandardized slopes, *b** represent the standardized slopes.

* one-tailed $p < .05$, but p values may not be trustworthy because the parameters estimated in M*plus* exceeded the number of Level-2 observations (yielding a warning about standard errors being untrustworthy).

TSE for IS = teacher self-efficacy for instructional strategies; TSE for SE = teacher self-efficacy for student engagement; TSE for CM = teacher self-efficacy for classroom management.

[a]This variable has been reported by students.

[b]This variable has been reported by teachers.

[c]0 = male, 1 = female. [e] $p = .050$.

9 Discussion[18]

The two empirical studies presented in this dissertation set out to address the central yet underexplored question of teacher self-efficacy's malleability. Despite its conceptual prominence in educational psychology (see Chapter 3), empirical evidence on its malleability remains limited—particularly with regard to key assumptions of socio-cognitive theory concerning the roles of (1) the sources of self-efficacy, (2) teachers' career stages, and (3) the context (Chapter 5; Fackler & Malmberg, 2016; Klassen et al., 2011; Kupers et al., 2023; Lauermann et al., 2020; D. B. Morris et al., 2017; Thommen et al., 2022; Tschannen-Moran & Woolfolk Hoy, 2001).

To close this gap, two complementary empirical studies were conducted. Study 1 employed a meta-analytic approach to synthesize findings from a wide range of intervention studies, analyzing how the four sources of self-efficacy and teachers' career stages impact intervention effects. Study 2 adopted an exploratory, within-person design to examine whether teachers' self-efficacy varies across the different classes they teach, and to investigate which classroom characteristics are particularly influential for their TSE.

Together, both studies not only tested key assumptions of socio-cognitive theory but also generated practical insights for educational research, teacher education and training, and educational administration. Researchers gain a clearer picture of where intervention research still falls short. Teacher educators and professional development providers receive evidence-based guidance on how to foster TSE. Educational administrators are offered a robust empirical foundation for making informed decisions about future investments in teacher support and development.

The findings are encouraging as both studies demonstrated that teacher self-efficacy is a malleable construct. Study 1 found that interventions addressing the sources of self-efficacy produced larger intervention effects than studies without the sources; however, the intervention effects did not differ significantly between different source combinations. Though the intervention characteristics "moment of reflection" significantly impacted the intervention effects. Importantly, interventions were equally effective for pre- and in-service teachers. Study 2 revealed that teachers' self-efficacy for classroom management and student engagement varied considerably within teachers across classes. Particularly, the classes' misbehavior had a negative relation to teachers' self-efficacy beliefs.

The following sections discuss these findings in detail. First, the findings from Study 1 regarding the overall effect of the interventions (RQ 1.1) and the robustness checks, including the analyses of study quality, are discussed, as these findings serve as a foundation for the following analyses.

[18] Parts of the discussion build on findings and formulations from the published meta-analysis (Täschner et al., 2024) and the working paper on context-specific teacher self-efficacy (Täschner et al., in preparation). However, the present chapter has been substantially adapted and partially rewritten to align with the overarching structure and aims of this dissertation.

Chapter 9.2 presents the discussion of the results regarding the role of the *four sources of self-efficacy* (RQ 1.2), Chapter 9.3 contains the discussion of the role of the *career stage* (RQ 1.3) and the interactions between career stage and the four sources of self-efficacy (RQ 1.4). In Chapter 9.4 the results regarding the *context-specificity* of teacher self-efficacy (RQ 2.1 and RQ 2.2) are discussed. Chapter 9.5 synthesizes the results of both studies before Chapter 9.6 reflects the limitations of both studies. Finally, Chapters 9.7 to 9.9 address implications for educational research, administration, and teacher education and training practice.

9.1 The Effects of Interventions on Teacher Self-Efficacy (RQ 1.1)

Study 1 showed that teachers report, on average, significantly higher teacher self-efficacy beliefs after interventions. Compared to the only other meta-analysis on the promotion of teacher self-efficacy to date, the 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 Study 1 included a broader range of interventions and career stages, unlike Mok et al. (2023), who focused on specific individual support activities and examined only pre-service and beginning teachers' self-efficacy. As there are no other meta-analyses on the promotion of teacher self-efficacy, it makes sense to compare the results to related constructs, namely teacher burnout, teacher stress, and teacher well-being. Iancu et al. (2018) conducted a meta-analysis of 23 controlled intervention studies aimed at reducing teacher burnout. Zarate et al. (2019) analyzed the effects of 18 controlled mindfulness-based interventions on teacher stress and burnout. Oliveira et al. (2021) evaluated the effects of 43 quasi-experimental and experimental studies using social and emotional learning on teachers' well-being and stress. All three meta-analyses included only in-service teachers and 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; 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). Study 1's 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, which underline the promising effects of the reviewed interventions and, in general, the malleability of teacher self-efficacy through intentional changes. Finally, the publication bias-adjusted effect of $g = 0.29$ is absolutely comparable with those of the previous meta-analyses on related constructs.

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, such 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, for example, through the hours teachers spend in other classrooms for lesson study 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 a range 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 investment when compared with the financial costs caused by teacher drop-out and teacher burnout, or the societal costs associated with teaching positions that remain unfilled. Studies estimate the annual costs associated with burnout among physicians to be at least $7,600 (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, September 13). 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.

9.1.1 Robustness of the Results

To prove the robustness of the overall intervention effects, a variety of relevant characteristics of the intervention (i.e., knowledge addressed, moment of reflection included, duration of intervention, period between measurements) and the study (domain of teacher self-efficacy, used scale, study design, study quality) were tested as moderators in meta-regressions. The effects were significantly larger when the intervention was described with details (one aspect of study quality), included moments of reflection, and had a specific duration (between 6 and 30 hours). All other characteristics did not produce significant differences within the effect sizes, indicating that teacher self-efficacy can be promoted through interventions regardless of specific methodological choices within the studies. Although the meta-analysis included studies with different rigorous study designs, namely, with and without a control group, the overall effect of $g = 0.47$ is comparable in size to the effect of methodologically rigorous experimental intervention studies (e.g., Thiel et al., 2020) and can be viewed as credible.

The assumption that studies with a higher study quality score will report significantly smaller effect sizes than those with a lower score could not be confirmed. 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 (and studies) are carefully thought through and 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 ($g = 0.55$) than those without details ($g = 0.35$). However, the analyses of study quality also point to a problem in the previous and current research: Forty-three effect sizes from 16 different studies reached a study quality rating of zero. These studies were kept in the meta-analysis as there was no moderating

effect of study quality. Nevertheless, future research is encouraged to engage in more strict study designs.

The significant roles of duration of intervention and moment of reflection support certain assumptions of Bandura (1997): He viewed self-efficacy as a quite durable construct and assumed the sources of self-efficacy to exert their impact via cognitive processes (also highlighted in models on teacher change, for example, in Clarke & Hollingsworth, 2002). Although there is a need to investigate both aspects in more detail, it can be given as empirical advice for future approaches aiming to promote teacher self-efficacy, to include moments of reflection and to invest a certain amount of time. Summed up, the robustness checks showed that methodological aspects seem less relevant than content-related aspects for the promotion of TSE (similar to findings on teacher professional development in general, see Kennedy, 2016).

9.2 The Role of the Four Sources in Interventions (RQ 1.2)

The results from Study 1 revealed that although interventions that targeted sources of self-efficacy were more successful than interventions without codable sources, none of the tested combinations of the sources was significantly superior. However, the systematic and careful extraction of the sources of self-efficacy targeted in the interventions revealed some interesting facts about the distribution of the sources of self-efficacy in intervention studies. Whereas the three most often addressed combinations all included mastery experiences, the source of physiological and emotional reactions was 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). However, recent research highlights the importance of adaptive emotion regulation strategies for teacher self-efficacy, suggesting a promising direction for future research (e.g., Burić et al., 2020; Huang & Zhou, 2024).

Within the pool of interventions targeting only one specific source, interventions targeting only vicarious experiences descriptively outreached mastery experiences. This large effect must be interpreted with caution, as only seven studies contributed to this moderator, and the broad corresponding confidence interval points to a high heterogeneity among these studies. Nevertheless, 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 Grossman, Compton, et al., 2009; Grossman, Hammerness, & McDonald, 2009; Loewenberg Ball & Forzani, 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 the theoretically derived 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 the reliance on the studies' intervention descriptions, it could only be coded whether mastery experiences were offered in interventions, not whether and to what extent these experiences were perceived 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. Initial research suggests that failure experiences might hinder teacher self-efficacy, particularly among teachers for whom being a teacher is a significant aspect of their identity (e.g., Coppe et al., 2023, in a study with pre-service teachers). Future research should therefore explore how failure (unsuccessful mastery experiences) contributes to the promotion of teacher self-efficacy. Another possible explanation for the missing effect of mastery experiences is that due to the focus on the three domains of teacher self-efficacy according to the TSES, effects of mastery experiences that were more closely aligned with other domains of teacher self-efficacy were missed (see for specificity-matching principle Ajzen & Fishbein, 1977; Swann et al., 2007). Nevertheless, the only mastery experiences effect size of $g = 0.43$ can still be interpreted as a substantial effect size regarding these limitations.

Compared 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 one of a 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 mentoring (e.g., Burger, 2024; Richter et al., 2013), coaching (e.g., Hoffman et al., 2015) and feedback (e.g., feedback givers expertise, Prilop et al., 2021; Weber et al., 2018, the content delivered, Wisniewski et al., 2020, or the addressed level, Hattie & Timperley, 2007). It could be that some adverse conditions have lowered the effect size of social persuasion.

The fact that there were no significant differences between the sources supports Bandura's (1997) ideas about the four sources as a substantial foundation for self-efficacy. It also complements reflections on teacher education in general and studies on teacher self-efficacy showing that different elements of training, respectively 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; Bardach et al., 2021; Grossman, Compton, et al., 2009; Grossman, Hammerness, & McDonald, 2009;

Loewenberg Ball & Forzani, 2009; Pfitzner-Eden, 2016b). As the coding relied on the intervention descriptions in the primary studies, it only captured whether a source was mentioned. The coding could not accurately represent the order, amount, intensity, or perceived valence (for example, whether the mastery experiences were perceived as successful). 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., Bach, 2022; Tschannen-Moran & McMaster, 2009).

Another explanation for the absence of significant differences in the sources of self-efficacy may reside in different individual reactions to interventions (Holzberger et al., 2025). Recently, research has begun to investigate shared connections between personality and self-efficacy, which could also impact the effect of the different sources of self-efficacy, as persons with high levels of extraversion might, for example, profit more from vicarious experiences and social persuasion than persons with lower levels (e.g., Arens & Preckel, 2025; Perera et al., 2018). Similarly, individual differences in cognitive processes have been discussed as likely influences on teacher self-efficacy (e.g., Bach, 2022; Bandura, 1997; Kozhevnikov, 2007; D. B. Morris et al., 2017) but have not been empirically examined so far. Indeed, the significant moderation effect of "moment of reflection" as discussed previously may indicate a significant role of cognitive processes in handling experiences. Therefore, future research should examine how different personality structures and cognitive styles impact the level and change of teacher self-efficacy. Such studies would not only provide highly relevant information on how different teachers and their reactions to interventions are, but they would also contribute to further verification of the role of personal factors in the model of triadic reciprocal causation (Bandura, 1997).

9.3 The Role of the Career Stages in Interventions (RQ 1.3)

It was hypothesized that interventions would exert a smaller influence on the self-efficacy of in-service teachers, given that interventions constitute a limited amount of experience compared to the years of everyday experience. However, Study 1 showed no significant differences in the intervention effects between pre- and in-service teachers. Descriptively, the effects were even slightly larger for in-service than for pre-service teachers. Additionally, the results from the follow-up measures in Study 1 indicated further growth in in-service teachers' self-efficacy from post- to follow-up test, pointing to sustainable effects of the interventions.

It can be presumed that Bandura's (1997) theoretical assumption of a rather stable level of teacher self-efficacy, once established, is not valid for interventions. Instead, it is reasonable to hypothesize that interventions represent a *different quality of experience* than everyday experiences, especially for in-service teachers. Another explanation for the unexpectedly large effect size could be that in-service teachers possess pronounced and diversified mental schemata of classroom processes (see mental schemata of experts vs. novices: Bohle Carbonell et al., 2014; Fries et al., 2021; McKeithen

et al., 1981). These schemata help them to quickly integrate new experiences made during the interventions into new levels of self-efficacy.

Whereas the positive, sustainable effects of interventions on *in-service* teachers' self-efficacy seem well-proven, the impact of interventions on *pre-service* teachers' self-efficacy in the long term is still pending. It was striking that only one study with pre-service teachers conducted a follow-up measurement (Eder-Karavaya et al., 2021). Therefore, it could not be investigated whether there are different mechanisms responsible for promoting TSE in different career stages. There remains 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 is a task for future research to carefully examine identical or realistically very similar interventions with pre-service and in-service teachers separately.

However, it also has to be considered that the change in pre-service TSE might be underestimated. 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).

Finally, with regard to differences between the two career stages, it has to be admitted that only a global categorization between pre-service and in-service teachers was feasible as the meta-analysis had to rely on existing studies and samples. There are hints that teacher self-efficacy develops differently in different subgroups of in-service teachers. The findings from Swan et al. (2011) and Woolfolk Hoy and Spero (2005) highlight 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. 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, as also outlined by van der Scheer and Visscher, 2016).

9.3.1 No Interaction Between Career Stages and Sources of Self-Efficacy (RQ 1.4)

Although theory and previous research have pointed to specific impacts of certain sources on pre-service teachers (e.g., Bandura, 1997; Tschannen-Moran & Woolfolk Hoy, 2007; Weß et al., 2020), pre-service teachers did not specifically benefit from sources other than in-service teachers. Given that neither sources (RQ 1.2) nor career stage (RQ 1.3) significantly moderated the intervention effects when investigated separately, the lack of a significant interaction effect may not be surprising. At the same time, the descriptive interactions 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), calling for future research to focus on the under-researched

combinations (e.g., only vicarious experiences with in-service teachers) to investigate the interactions of different sources and career stages more systematically.

For both career stages, mastery experiences were part of the conditions producing the largest changes, emphasizing the central role of mastery experiences for establishing and promoting teacher self-efficacy as described by Bandura (1997). For pre-service teachers, interventions offering only mastery experiences descriptively produced the highest self-efficacy changes ($g = 0.62$), which blends in well with ideas from practice-based teacher education research that calls for early and relevant opportunities for enactive experiences throughout teacher education (e.g., developing a lesson plan at university; see Grossman, Compton, et al., 2009; Grossman, Hammerness, & McDonald, 2009; Loewenberg Ball & Forzani, 2009). For in-service teachers, the descriptively largest effect sizes appeared in interventions offering mastery experiences combined with social persuasion ($g = 0.73$). A practical implication to enhance in-service teachers' self-efficacy is enriching their daily experiences with collegial feedback sessions or coaching (Suchodoletz et al., 2018). In summary, both career stages descriptively profited the most from sources that are not easily and daily available to them.

9.4 The Context-Specificity of Teacher Self-Efficacy (RQ 2.1)

Study 2 revealed that teacher self-efficacy is in line with the theoretical assumptions indeed a context-specific construct that not only varies between teachers but also within teachers. Between 39% and 63% of the variance in teacher self-efficacy occurred across classes within teachers. The amount of variation within teachers was mostly larger than within the teachers studied by Raudenbush et al. (1992; 44%) and Ross et al. (1996; 21%), which might be due to the use of a domain-specific TSE questionnaire in Study 2. It can be assumed that this questionnaire could capture more nuances in TSE than the single-item measures used in the two previous studies.

Variations in self-efficacy beliefs within teachers have also been observed across other contexts, such as lessons (e.g., Rupp & Becker, 2021) and individual students (Zee, de Jong, & Koomen, 2016). Interestingly, the amount of variation across classes — the context considered in Study 2 — is mostly higher than that found in the study evaluating the lessons by Rupp and Becker (2021). With regard to Bandura's theory (1997), this seems to suggest that (1) pre-service teachers' TSE, the sample in Rupp and Becker (2021), fluctuates more between persons than in-service teachers TSE, the sample in Study 2 and Zee, de Jong, and Koomen, 2016, and that (2) SE is less reactive to situational, short-term contexts like lessons, compared to more long-term contexts as individual students and classes. However, whether such differences in the variability between the two career stages of pre- and in-service teachers really exist has to be shown in future studies with larger samples.

The fact that the use of a class-specific questionnaire of TSE led to differences in teachers' self-efficacy ratings raises the question of how to best measure teacher self-efficacy in future studies. So far, measures of teacher self-efficacy that don't specify a specific class as a reference dominate current

research (e.g., Dicke et al., 2014). However, initial evidence suggests that class-specific questionnaires may yield more informative results in secondary school settings (e.g., Thommen et al., 2022). It is an important quest for future research to investigate, which reference levels teachers rely on when no explicit reference (such as a specific class) is given in self-efficacy questionnaires. Qualitative methods might be a helpful add-on here to understand the internal rating processes of teachers in more detail. It should also be examined how measures with and without explicit contextual references (e.g., specific classes or teaching situations) relate to each other, and whether distinct use cases exist for these two types of measures (Cortina, 2025; inspiration for a suitable study design in Zygar-Hoffmann et al., 2024).

Moreover, the three domains of TSE varied to a different degree across classes. This finding differs from previous studies on TSE variations, which did not find such domain-specific differences across lessons or individual students (e.g., Rupp & Becker, 2021; Zee, de Jong, & Koomen, 2016). In Study 2, the majority of variance occurred within teachers in the domains of SE and CM (and thus, the class context seemed to be of high relevance for these domains). On the contrary, the majority of variance occurred between teachers in the domain of IS, and thus, the class context seemed less relevant for TSE for IS. It is presumable that TSE for IS is less influenced by the class context than by individual teacher factors such as teachers' professional knowledge or their access to instructional resources. Indeed, all four items of the subscale TSE for IS refer to challenges that probably differ more by instructional topic than by class (e.g., providing different explanations, evaluating students' understanding, adapting to individual levels of students, and providing challenges for good learners). Unfortunately, studies evaluating the role of individual teacher factors like professional knowledge for TSE have so far not considered differential relationships with separate domains of TSE (Lauermann & König, 2016; Tschannen-Moran & Woolfolk Hoy, 2007). Future studies are therefore needed to examine this assumption.

Finally, future research is advised to clarify the *meaning of within-teacher variability*: Does it represent measurement error around the mean, or is it indicative of other psychological processes? The higher variability within teachers regarding classroom management and student engagement self-efficacy in Study 2 could, for example, point to adaptive reactions to the conditions. However, other research on intraindividual variability suggests a higher amount of variability to be connected with maladaptive processes (e.g., Geukes et al., 2017; Hausen et al., 2022; Houben & Kuppens, 2020).

The exploratory visual inspection suggested different patterns of variability within teachers in Study 2. Similar differences have been found in a recent experience-sampling study (Kupers et al., 2023). Whether these differences in variability are psychologically meaningful, and what constitutes these different patterns, remains to be explored in future research.

Beyond the question of how much teacher self-efficacy varies within teachers across classrooms, Study 2 also investigated which class characteristics are related to TSE. Study 2 considered

various behavioral, demographic, and achievement characteristics of the classes. The amount of explained variance in this multivariate multilevel model—depending on the domain of TSE between 27% and 54%—mostly exceeds the amount of explained variance in previous models (Raudenbush et al., 1992; Ross et al., 1996), which might be taken as confirmation of the careful selection of the considered class characteristics. Class misbehavior was consistently negatively related to TSE in all three domains; all other class characteristics were less consistent or not at all related to TSE.

9.4.1 The Role of Class Misbehavior (RQ 2.2)

Regardless of the domain of TSE, the secondary school teachers felt less efficacious when they perceived the amount of misbehavior in class as relatively high. This finding is in line with several studies showing close connections between students' misbehavior, reduced TSE, and higher levels of exhaustion (e.g., Aldrup et al., 2018; Aloe, Shisler, et al., 2014; Duan et al., 2024; Hoogendijk et al., 2023; Tsouloupas et al., 2010; Zee, de Jong, & Koomen, 2016). With regard to the theoretical foundations of self-efficacy by Bandura (1997), two sources of self-efficacy, namely *mastery experiences* and *physiological and emotional reactions*, might explain this consistent link (Bandura, 1997). For example, Martins et al. (2015) showed that the perception of class characteristics as failure (the opposite of successful mastery experiences) was connected with lower levels of TSE. It is plausible that teachers perceive misbehavior in a class as a failure, as it hinders them from their main tasks as teachers and, therefore, lowers their self-efficacy in all domains. Research has also shown that teachers report more frustration and anger when they do not attain their goal of establishing and maintaining discipline in a class (M.-L. Chang, 2013; Frenzel et al., 2020). Negative emotions as another source of self-efficacy (i.e., physiological and emotional reactions) can lower TSE (Bach & Hagenauer, 2022; Bandura, 1997; Burić et al., 2020).

Based on the cross-sectional design of Study 2, the causality of the relation remains unclear. There is first longitudinal research indicating that TSE predicts, or better prevents, later misbehavior, but not vice versa (Dicke et al., 2014; Malinen et al., 2024). However, this preliminary evidence stems from samples of primary school teachers or pre-service teachers, and whether these findings can be transferred to secondary education is a point for future research. In the meantime, continued efforts should be made to promote high levels of teacher self-efficacy and classroom management skills.

9.4.2 The Role of Student Engagement (RQ 2.2)

Whereas students' emotional engagement was positively linked to TSE for SE, behavioral engagement was not significantly linked to any domain of TSE. Although several studies have pointed to the importance of TSE for students' emotional engagement or interest (e.g., Hettinger et al., 2023b; ten Hagen et al., 2022; Zee & Koomen, 2020), this is one of the first studies suggesting a potential reciprocal link between students' emotional engagement and TSE (see also ten Hagen et al., 2022). One explanation for the strong relation between students' emotional engagement and TSE for SE might be

teachers' interpretation of student engagement as a mastery experience (similar argumentation in Kupers et al., 2023). Another explaining mechanism could be emotional contagion, also called emotion transmission (Frenzel et al., 2018; Frenzel et al., 2009). Emotion transmission means that a student's emotional state can impact a teacher's emotional state and vice versa.

In contrast to the two previous studies on class-specific variations in TSE by Raudenbush et al. (1992) and Ross et al. (1996), students' behavioral engagement was not a significant predictor for TSE in Study 2. One possible explanation could be the way the data has been collected: In Study 2, the students themselves reported on their engagement, whereas in the two older studies, the teachers themselves reported on the students' engagement, which may have led to an overestimation of the relation due to common-method bias and social-desirability effects (Fisher & Katz, 2000; Podsakoff et al., 2012). Another explanation comes from Martin et al. (2020), who showed that students' behavioral engagement varies more between students and lessons than between days and weeks. It is, therefore, plausible that the aggregated measure at the class-level in Study 2 did not capture differences in students' behavioral engagement. Furthermore, students' behavioral engagement might be linked to teachers' actual behavior rather than their TSE. Several studies have shown close links between teacher behavior and student (behavioral) engagement (e.g., Engels et al., 2016; Hughes et al., 2011; Nguyen et al., 2018; Ruzek et al., 2016; Yang Wang, 2024). There is even one study where only teachers' behavior (i.e., autonomy-supportive practices), but not TSE, was linked to student engagement (Lauermann & Berger, 2021).

9.4.3 The Role of Demographics and Achievement (RQ 2.2)

The class characteristic of migration background was not significantly linked to any domain of TSE. One reason might be that the variation occurred on a level other than the class level. The high ICC (.94) of migration background in Study 2 indicated that migration background varied more between teachers than within teachers. It is plausible that the proportion of students with migration backgrounds varies more between schools than between classes (see Glock et al., 2019, for a school-level study on TSE). Another reason for the missing link might be that the migration background is more related to other domains of TSE, such as TSE for cultural diversity, which was not considered in Study 2 (e.g., OECD, 2019; Siwatu, 2007).

Surprisingly, teachers felt more efficacious in establishing and maintaining discipline in classes with relatively more girls. This contradicts previous studies on TSE that did not find a gender effect beyond a behavioral effect (Geerlings et al., 2018; Zee, de Jong, & Koomen, 2016). One explanation could be the different ages of the students: Whereas the previous studies investigated elementary school students, Study 2's sample was secondary school students. Similar studies with secondary school students report higher levels of discipline in secondary school classes with more females, too (Fauth et al., 2020; Heinschel et al., 2024).

Finally, the mean grade in a class was significantly related to TSE for SE. Although it may seem surprising at first glance, a closer look at the questionnaire reveals that two of the four items capturing TSE for SE are closely related to students' achievement (Tschannen-Moran & Woolfolk Hoy, 2001). It is plausible that teachers feel more confident in encouraging students to "think through the material provided" or to "allow students to discover new things for themselves" in classes with higher grades than in classes with lower grades.

Summed up, the results of Study 2 highlight teacher self-efficacy as a context-specific construct that is embedded in a dynamic interplay with class characteristics, particularly regarding behavioral aspects such as class misbehavior and students' emotional engagement.

9.5 Synthesis of the Findings

The combined findings of the two studies compellingly demonstrate that teacher self-efficacy is a malleable construct, shaped by multiple pathways outlined in Bandura's model of triadic reciprocal causation, ranging from behavioral enactment in interventions to mastery experiences in everyday classroom practice. Crucially, these findings challenge the notion of TSE as a purely fixed trait; instead, they highlight that teacher educators, principals, colleagues, and teachers themselves can actively foster TSE.

Concerning the reality of in-service teachers, Study 1 underscores the effectiveness of interventions, particularly those involving mastery experiences and social persuasion. Study 2 complements this by showing that classroom-level factors—especially student misbehavior—are consistently and negatively related to TSE, suggesting that mastery experiences are not only created but also undermined by contextual challenges. Together, these findings suggest a practical implication: principals and colleagues can enhance TSE by engaging in constructive social persuasion, such as coaching, mentoring, and feedback, especially in contexts where mastery experiences are threatened.

Mastery experiences emerge as a central mechanism across both studies. In Study 1, they were part of the most effective interventions for both pre- and in-service teachers. In Study 2, student-rated achievement and emotional engagement were positively linked to TSE, while misbehavior was negatively associated—interpretable as the presence or absence of mastery experiences. Moreover, the importance of reflection in Study 1 suggests that schools, universities and teacher training institutes should cultivate routines for teachers to reflect on and share successful experiences, not only at the end of the year but continuously. A culture of celebrating success may be key to sustaining high levels of TSE.

The two studies differed in their findings regarding a domain-specific difference. While Study 1 found interventions to be broadly effective across all domains of TSE, Study 2 revealed domain-specific contextual sensitivity — particularly that TSE for instructional strategies was less influenced by classroom-level factors than TSE for classroom management or student engagement. This

discrepancy may stem from methodological differences: Study 1 relied on study-level aggregated pre-post data, which limited visibility into intraindividual variation, while Study 2 intentionally investigated intraindividual differences using a class-specific measure of TSE. Additionally, domain-specific effects may be masked by the alignment between intervention content and measurement focus in the studies included in Study 1. Further, Study 1 focused on domain-specific differences in overall intervention effects, yet did not explore whether particular combinations of self-efficacy sources might differentially impact specific domains. Future research should investigate these nuanced interactions to better understand how targeted interventions can be optimized for distinct aspects of teaching.

In sum, the two studies converge to portray teacher self-efficacy as both changeable and context-sensitive. They advocate for interventions that foster reflection, a variety of experiences, and are oriented to the realities of classroom life to empower teachers not only in general, but in the nuanced and variable contexts in which they teach.

9.6 Limitations

Both studies of this dissertation have been designed and carried out with high rigor and transparency. This is reflected in the pre-registrations before data collection, the use of state-of-the-art statistical analyses, and the open sharing of materials and results on Open Science Framework (OSF). Furthermore, this dissertation addresses unique research gaps in self-efficacy research by combining the first meta-analytic synthesis of intervention studies on teacher self-efficacy with an exploratory, conceptual replication study (Perry et al, 2022; Plucker & Makel, 2021).

Study 1 relies on a database of more than 11,000 teachers, which is exceptional in meta-analytic research on teachers and interventions. The development of a coding manual for the sources of self-efficacy represents a novel contribution to the field, and its public availability is intended to support future research. In addition, the study considered a wide range of methodological moderators, none of which significantly influenced the results, thereby underscoring the robustness of the findings. Study 2 strengths lie in the application of a multi-informant approach (as called for by Klassen et al., 2011; ten Hagen et al., 2022) and its systematic consideration of diverse classroom characteristics. These characteristics enabled a context-sensitive and detailed understanding of teachers' everyday experiences. Nevertheless, the results from both studies must be viewed in light of several limitations described in the following.

Both studies focused on the three domains of self-efficacy that the TSES captures (Tschannen-Moran & Woolfolk Hoy, 2001). Despite theoretically representing core domains of teachers' lives, it is possible that they did not fully align with the experiences targeted by the interventions in Study 1 or with the everyday experiences of teachers in Study 2. Future research could explore whether the findings of both studies replicate in other domains of teacher self-efficacy, such as diagnostic competence or differentiation in heterogeneous classrooms (e.g., Meschede et al., 2020).

Furthermore, both studies considered a higher level of teacher self-efficacy as a positive and desirable outcome. However, first empirical studies are pointing to curvilinear relations and tipping points of TSE in its relation with other outcomes (e.g., Müller et al., 2024 for error orientation; Scherer et al., 2023 for experience; Thien & Liu, 2024 for instructional leadership). So far, the question of what constitutes a good or sufficient level of self-efficacy remains mainly theoretically discussed (e.g., Bach, 2022).

Besides, only linear relationships were considered in the analyses in both studies. While this is common practice in self-efficacy research, it might have limited the ability to detect more complex, potentially non-linear change processes (e.g., Cortina, 2025; Clarke & Hollingsworth, 2002). Exploring such patterns represents a promising direction for future studies.

Finally, although both studies were deeply informed by socio-cognitive theory and tested neglected theoretical assumptions, the theoretical assumptions were, comparable to many contemporary motivational theories, underspecified and left room for interpretation (Murayama & Keyserlingk, 2025).

9.6.1 Limitations of Study 1

Despite immense efforts to retrieve all relevant studies in line with current guidelines of systematic reviews and meta-analyses, there is a possibility that Study 1 missed relevant articles. In particular, more than 30 potentially relevant and initially included studies had to be excluded as they lacked the necessary statistics for the calculations of the meta-analysis. All authors had been contacted, but they could not provide the data. There is hope that the increased commitment to open science, more journal policies obliging their authors to share data publicly, and more incentives for data publishing and sharing will help to diminish this meta-analytic problem (Leising et al., 2022; Nosek et al., 2015; Patall, 2021).

The overall effect size was robust across a variety of methodological variables. Nevertheless, three biases might have impacted its size and should be considered in the interpretation: First, there is a chance of a selection bias as the results from the funnel plot and the PEESE indicated an overestimation of the effect due to an above-average proportion of studies with large samples and positive effect sizes and underrepresentation of smaller samples with larger standard errors and studies with negative effect sizes. Second, the effect size may also be affected by a response shift bias, which means that teachers used different reference levels for their self-efficacy ratings at the pre-test and post-test (e.g., Oort et al., 2009; Miller et al., 2023). Typically, response shift bias is associated with an underestimated effect size; however, it has not yet been tested in teacher self-efficacy research. Third, the main analyses were based on post-tests immediately after the intervention. At this point, teachers might have been overly optimistic and overrated their self-efficacy.

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

manual), coding the sources of self-efficacy in the interventions was highly inferential. Despite the moderate to substantial interrater agreement, it is possible that the coders did not capture every source combination perfectly due to the remaining subjectivity within the coding of the sources of self-efficacy. Furthermore, due to the meta-analytical approach, it was only possible to code whether the source was described in the intervention description and, therefore, probably offered to the participants. Study 1 could not consider individual cognitive processes of the teachers or specific contextual backgrounds that may have interacted with the sources of self-efficacy (e.g., cultural background in Yada et al., 2019). The intervention duration was exploratorily analysed with arbitrary categories based on the quartiles of the available data; the significant effect might be based on a statistical artefact. Future research is encouraged to investigate the role of intervention duration in a more hypothesis-driven manner.

Finally, Study 1 cannot provide information on the interplay between different intervention features, as it primarily applied single regression approaches. The remaining substantial amount of heterogeneity in each of the meta-regressions indicates that further variables or interactions could explain more between-study heterogeneity. A promising approach to back up future moderator analyses would be to use multiverse analysis in this context (e.g., Patzl et al., 2024; Voracek et al., 2019).

9.6.2 Limitations of Study 2

In line with its aims to explore the context-specificity of TSE, Study 2 was based on a convenience sampling approach. This led to participants mainly coming from urban areas in the Netherlands. Whereas teachers' demographics correspond to larger studies like TALIS, the proportion of students with a migration background is, for example, higher than in comparative studies (OECD, 2019). Further, the three classes in which the teachers reported their self-efficacy were chosen by the teachers themselves, which might have induced selection bias. Therefore, future studies with more representative samples and randomized class selections are needed to investigate whether Study 2 findings can be generalized.

Although most predictor variables were based on the aggregated data of 1,326 students, which provided a sufficient sample size, the sample size of 26 teachers scratches the lower edge of recommendations for Level-2 variables in multilevel analysis (Maas & Hox, 2004, 2005). This probably induced underestimated standard errors of the TSE variance and inflated significances. As a reaction to the small teacher sample size, manually constructed scale means were used; however, a more statistically advanced approach of modelling the measurement error, as done in latent modelling approaches, was not possible due to the sample size. Future researchers are therefore strongly encouraged to replicate the findings with larger cluster sample sizes and to statistically disentangle measurement error from true variation.

Another limitation is that the study was cross-sectional by design. Although TSE was measured in three different classes, it is impossible to draw any conclusions about the temporal interrelations

between the evaluated predictors and TSE. There are first indices that teacher self-efficacy serves more as a predictor than a consequence in classroom processes (e.g., Burić et al., 2024; Lazarides & Schiefele, 2024). Consequently, cross-lagged studies with considerate time points are needed: It would be useful to assess TSE and class characteristics in several classes right at the beginning of a school year (or whenever a teacher starts teaching a class) and again after a specific period to disentangle the directions (see Malinen et al., 2024 for a similar design).

Finally, although Study 2 considered a wide variety of class characteristics, several unaccounted variables remain that could also impact TSE. On the teacher level, future studies should consider additional teacher covariates, such as teachers' professional knowledge, their working hours, or their professional development needs (e.g., Fackler et al., 2021; Perera et al., 2019). On the class level, it seems important for future research to consider a teacher's relational history or familiarity with a class (see also de Ruiter et al., 2020, and Malinen et al., 2024 for similar ideas) as well as teachers' individual mastery experiences in these classes (e.g., D. B. Morris et al., 2017).

9.7 Implications for Future Research

Teacher self-efficacy remains a highly relevant construct for educational research. Building on the literature review on teacher self-efficacy outcomes in Chapter 2.2, the realization of the two empirical studies, and the interpretation of their findings, several promising avenues for future research have emerged. To guide future researchers, these gaps are summarized under three overarching themes: methodological advancements, individual differences, and the broader relevance and impact of teacher self-efficacy.

Methodological Advancements: A central challenge for future research lies in refining how teacher self-efficacy should be measured. In particular, rating processes without clear contextual references and the relationship between context-specific and measures without contextual references, such as a specific class, require closer examination (e.g., Thommen et al., 2022; Zygar-Hoffmann et al., 2024). Further, intervention studies should move beyond broad comparisons and systematically vary source combinations, sequences, intensities, and perceived valences to uncover nuanced mechanisms of change (e.g., Tschannen-Moran & McMaster, 2009). Moreover, follow-up assessments should become standard practice to evaluate the sustainability of intervention effects and reduce biases of self-ratings based purely on immediate post-tests. Longitudinal designs are also needed to disentangle the situational, contextual, and personal components of TSE, thereby enabling more tailored and effective interventions (see Carstensen et al., 2024).

Individual Differences: Understanding how individual cognitive processes mediate the effects of the four sources of self-efficacy is a critical next step (e.g., Bach, 2022; Bandura, 1997; D. B. Morris et al., 2017). Future studies should also explore which teacher subgroups—particularly among in-service

teachers—are most vulnerable to low self-efficacy and may benefit from targeted support (e.g., Klassen & Chiu, 2010; van der Scheer & Visscher, 2016).

Relevance and Impact of Teacher Self-Efficacy: To fully understand the role of TSE in educational processes, future research should examine its unique predictive value alongside other constructs such as professional knowledge and achievement goals. Intervention effects should be contextualized by comparing them to natural longitudinal developments and by assessing their impact on meaningful outcomes like teaching quality. Additionally, the interpretation of within-teacher variability in TSE deserves further attention, as it may reflect underlying psychological processes (e.g., Kupers et al., 2023).

In sum, teacher self-efficacy research stands at a promising crossroads. By addressing these open questions, future studies can deepen our understanding of how teachers develop confidence in their professional roles—and how educational systems can support them in doing so.

9.8 Implications for Educational Administration and Policy

Empirical Evidence for the Usefulness of Teacher Self-Efficacy Interventions: The findings of Study 1 provide educational policy-makers with compelling empirical evidence that teacher self-efficacy can be effectively promoted through well-designed interventions across all career stages (Harris, 2009; Kraft, 2020). The reported average effect size offers a valuable benchmark for future funding proposals of schools that want to implement measures to promote teacher self-efficacy or of researchers who want to evaluate a new intervention program. Strengthening teacher self-efficacy is not merely a matter of individual development; the promising average effect size justifies strategic investment in it as a means to improve teacher retention, teacher well-being, and instructional quality.

Increase Quality Standards in Funded Research and Data Availability: The large number of existing studies with limited methodological rigor highlights the need for policymakers and funding agencies to establish minimum quality standards for future funded research. These standards should include, at a minimum, the implementation of delayed treatment control groups. Study 1 also attempted to contextualize the intervention effects in terms of economic and educational costs—such as teacher attrition and burnout—but was constrained by the lack of publicly available data. To enable more precise cost-benefit analyses, governmental institutions are encouraged to improve the transparency and accessibility of such relevant data.

Necessity of Infrastructure for Knowledge Exchange and Teacher Support: Despite the existence of promising interventions, there is currently no systematic infrastructure for documenting, evaluating, and disseminating successful practices. Educational policy could play a key role in establishing platforms for exchange and upscaling, such as national and international conferences, centralized databases, or collaborative networks. These structures would facilitate the transfer of evidence-based practices across schools and regions. Given that in-service teachers benefited

substantially from interventions, educational policy should aim to create school environments in which professional development is seamlessly integrated into teachers' work routines. Such conditions should ensure that professional learning is perceived as supportive rather than burdensome. In addition, teachers should receive targeted support in managing classroom behavior, which Study 2 identified as a persistent threat to teacher self-efficacy.

9.9 Implications for Teacher Education, Training and Practice

Integrating Teacher Self-Efficacy Into Teacher Education: Teacher education programs should aim not only to convey subject-specific knowledge but also to foster pedagogical skills and thereby self-efficacy (e.g., Thiel et al., 2020). Fostering awareness about the positive outcomes of a high level of teacher self-efficacy and the possibility of promoting it might help future teachers develop their own self-efficacy and support future colleagues.

Facilitating Reflection: Formats in teacher education that aim to promote teacher self-efficacy should provide opportunities for reflection, for example, due to the use of learning portfolios or developmental journals (e.g., Michos et al., 2022). Also, in-service teachers who want to promote their self-efficacy can make use of focused reflections. A good starting point for them could be comparing different classes: why does the teacher feel more self-efficacious in one class?

Creating Opportunities for New Experiences: Pre-service teachers should be offered as many opportunities for mastery and vicarious experiences as possible. Lists with teachers' core practices can serve a good inspiration (Michigan Department of Education, n.d.; TeachingWorks, n.d.). Many topics in teacher education curricula already contain elements of these practices and can be used to design authentic learning experiences—such as developing lesson plans or participating in video-recorded role plays (e.g., Gold et al., 2017). Opportunities for non-routine experiences should be created to strengthen self-efficacy among in-service teachers. The findings suggest that combinations with forms of social persuasion, such as feedback from colleagues, mentoring, or coaching, might be beneficial (e.g., Hoogendijk et al., 2018; Richter et al., 2013).

Supporting Classroom Management Challenges: Given the prominent role of classroom behavior in shaping teachers' self-efficacy, the most challenging classes should not be assigned to novice or low-efficacy teachers (see also Keller-Schneider et al., 2020; Malinen et al., 2024). Furthermore, opportunities for targeted support (e.g., coaching from expert teachers, training for emotion regulation competencies) and further personal staff to assist teachers in challenging classes should be available.

10 Conclusion

Grounded in socio-cognitive theory (Bandura, 1997) and a broad empirical foundation linking teacher self-efficacy (TSE) to teacher well-being and classroom outcomes, this dissertation examined the malleability of TSE through two complementary lenses: intentional change via interventions and natural variation across classroom contexts. The findings affirm that teacher self-efficacy is a malleable construct, responsive to both professional development opportunities and the everyday realities of the classroom.

Study 1 demonstrated that interventions can effectively enhance TSE across all career stages, with moments of reflection emerging as a particularly influential component. While none of the four sources of self-efficacy significantly moderated the intervention effects, their consistent presence across successful interventions underscores their collective relevance. Study 2 revealed substantial intraindividual variation in TSE across classes, with classroom misbehavior standing out as a consistently negative contextual factor. These findings suggest that teacher self-efficacy is not only shaped by what teachers do and experience, but also by the environments in which they teach.

Together, the studies highlight that teacher self-efficacy should be viewed as a construct that can be cultivated by teacher educators, school leaders, and colleagues alike. Supporting pre- and in-service teachers in developing and reflecting on mastery experiences, facilitating vicarious experiences, providing constructive feedback, and fostering emotionally supportive classroom environments are promising avenues for strengthening TSE.

Future research should further investigate how teacher self-efficacy is best measured, especially in light of its context-specific nature, and explore how individual differences in cognition and personality may shape responses to interventions. Educational policy and practice are encouraged to invest in structures that promote high levels of teacher self-efficacy, not only as a means of professional development, but as a foundation for resilient, reflective, and effective teaching.

11 Declaration of AI Usage

DeepL, Grammarly, and Copilot were used to linguistically edit this dissertation (Level 3, AI-assisted editing). Copilot was used for the generation of ideas regarding the title and the structure of the implications (Level 2, AI-assisted idea generation and structuring). All suggested changes and ideas were critically reviewed by the author. This declaration is based on the Artificial Intelligence Assessment Scale (Perkins et al., 2024).

12 References

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

Abele, E. A., & Candova, A. (2007). Prädiktoren des Belastungserlebens im Lehrerberuf [Predicting teachers' stress experience: Findings from a 4-year longitudinal study]. *Zeitschrift für Pädagogische Psychologie*, *21*(2), 107–118. https://doi.org/10.1024/1010-0652.21.2.107

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*(5), 888–918. https://doi.org/10.1037/0033-2909.84.5.888

Alanoglu, M. (2022). The role of instructional leadership in increasing teacher self-efficacy: A meta-analytic review. *Asia Pacific Education Review*, *23*, 233–244. https://doi.org/10.1007/s12564-021-09726-5

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

Aldrup, K., Klusmann, U., Lüdtke, O., Göllner, R., & Trautwein, U. (2018). Student misbehavior and teacher well-being: Testing the mediating role of the teacher-student relationship. *Learning and Instruction*, *58*, 126–136. https://doi.org/10.1016/j.learninstruc.2018.05.006

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

Aloe, A. M., Shisler, S. M., Norris, B. D., Nickerson, A. B., & Rinker, T. W. (2014). A multivariate meta-analysis of student misbehavior and teacher burnout. *Educational Research Review*, *12*, 30–44. https://doi.org/10.1016/j.edurev.2014.05.003

Anders, F. (2025, August 7). *Lehrermangel im neuen Schuljahr – Länder reagieren mit Abordnungen oder Mehrarbeit [Teacher shortage in the new school year - Länder respond with delegations or extra work]*. Deutsches Schulportal. https://deutsches-schulportal.de/bildungswesen/lehrermangel-bleibt-bundesweit-ein-problem/

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

Arens, A. K., & Preckel, F. (2025). Distinct constructs with a common core: Modelling neuroticism, self-esteem, and generalized self-efficacy using bifactor exploratory structural equation modelling. *Current Psychology*, *44*, 12209–12224. https://doi.org/10.1007/s12144-025-07984-9

Asparouhov, T., & Muthén, B. (2019). Latent variable centering of predictors and mediators in multilevel and time-series models. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(1), 119–142. https://doi.org/10.1080/10705511.2018.1511375

Atay, D. (2007). Beginning teacher efficacy and the practicum in an EFL context. *Teacher Development*, *11*(2), 203–219. https://doi.org/10.1080/13664530701414720

Axboe, M. K., Christensen, K. S., Kofoed, P.-E., & Ammentorp, J. (2016). Development and validation of a self-efficacy questionnaire (SE-12) measuring the clinical communication skills of health care professionals. *BMC Medical Education*, *16*(1), 272. https://doi.org/10.1186/s12909-016-0798-7

Aykaç, N., Ulubeey, 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*, 338–351. https://doi.org/10.33200/ijcer.587566

Bach, A. (2022). *Selbstwirksamkeit im Lehrberuf: Entstehung und Veränderung sowie Effekte auf Gesundheit und Unterricht [Teacher self-efficacy: Origins, change, and effects on teacher health and classroom practice]*. Waxmann Verlag GmbH. https://doi.org/10.31244/9783830995166

Bach, A., & Hagenauer, G. (2022). Joy, anger, and anxiety during the teaching practicum: How are these emotions related to dimensions of pre-service teachers' self-efficacy? *Zeitschrift für Bildungsforschung*, *12*(2), 295–311. https://doi.org/10.1007/s35834-022-00343-9

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

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

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

Bandura, A. (2004). Health promotion by social cognitive means. *Health Education Research*, *31*(2), 143–164. https://doi.org/10.1177/1090198104263660

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Information Age Publishing.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management*, *38*(1), 9–44. https://doi.org/10.1177/0149206311410606

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*. 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

Beck, J. W., & Schmidt, A. M. (2012). Taken out of context? Cross-level effects of between-person self-efficacy and difficulty on the within-person relationship of self-efficacy with resource allocation and performance. *Organizational Behavior and Human Decision Processes*, *119*(2), 195–208. https://doi.org/10.1016/j.obhdp.2012.06.009

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

Bellibaş, M. Ş., & Liu, Y. (2017). Multilevel analysis of the relationship between principals' perceived practices of instructional leadership and teachers' self-efficacy perceptions. *Journal of Educational Administration*, *55*(1), 49–69. https://doi.org/10.1108/JEA-12-2015-0116

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

Berg, D. A. G., & Smith, L. F. (2018). The effect of school-based experience on preservice teachers' self-efficacy beliefs. *Issues in Educational Research*, *28*(3), 530–544.

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

Björnsson, J. K. (2020). Teaching culturally diverse student groups in the Nordic countries - what can the TALIS 2018 data tell us? In T. S. Frønes, A. Pettersen, J. Radišić, & N. Buchholtz (Eds.), *Equity, equality and diversity in the Nordic model of education* (pp. 75–98). Springer International Publishing.

Bohle Carbonell, K., Stalmeijer, R. E., Könings, K. D., Segers, M., & van Merriënboer, J. J. (2014). How experts deal with novel situations: A review of adaptive expertise. *Educational Research Review*, *12*, 14–29. https://doi.org/10.1016/j.edurev.2014.03.001

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

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

Bosse, S., Henke, T., Jäntsch, C., Lambrecht, J., Vock, M., & Spörer, N. (2016). Die Entwicklung der Einstellung zum inklusiven Lernen und der Selbstwirksamkeit von Grundschullehrkräften [The development of attitudes towards inclusive education and self-efficacy among elementary school teachers]. *Empirische Sonderpädagogik*, *8*(1), 103–116. https://doi.org/10.25656/01:11857

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.

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

Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education*, *16*(2), 239–253. https://doi.org/10.1016/S0742-051X(99)00057-8

Burger, J. (2024). Constructivist and transmissive mentoring: Effects on teacher self-efficacy, emotional management, and the role of novices' initial beliefs. *Journal of Teacher Education*, *75*(1), 107–121. https://doi.org/10.1177/00224871231185371

Burgueño, R., Sicilia, A., Medina-Casaubón, J., Alcaraz-Ibañez, M., & Lirola, M.-J. (2019). Psychometry of the teacher's sense of efficacy scale in Spanish teachers' education. *The Journal of Experimental Education*, *87*(1), 89–100. https://doi.org/10.1080/00220973.2018.1423542

Burić, I., Huić, A., & Sorić, I. (2024). Are student engagement and disaffection important for teacher well-being? A longitudinal examination of between- and within-person effects. *Journal of School Psychology*, *103*, 101289. https://doi.org/10.1016/j.jsp.2024.101289

Burić, I., & Kim, L. E. (2020). Teacher self-efficacy, instructional quality, and student motivational beliefs: An analysis using multilevel structural equation modeling. *Learning and Instruction*, *66*, 101302. https://doi.org/10.1016/j.learninstruc.2019.101302

Burić, I., & Moè, A. (2020). What makes teachers enthusiastic: The interplay of positive affect, self-efficacy and job satisfaction. *Teaching and Teacher Education*, *89*, 103008. https://doi.org/10.1016/j.tate.2019.103008

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

Burić, I., Zuffianò, A., & López-Pérez, B. (2022). Longitudinal relationship between teacher self-efficacy and work engagement: Testing the random-intercept cross-lagged panel model. *Contemporary Educational Psychology*, *70*, 102092. https://doi.org/10.1016/j.cedpsych.2022.102092

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

Carstensen, B., Aldrup, K., Lüdtke, O., & Klusmann, U. (2024). How stable is student teachers' emotional exhaustion? Disentangling different components of stability and change using the STARTS model. *Journal of Educational Psychology*, *116*(5), 703–718. https://doi.org/10.1037/edu0000852

Cardon, M. S., & Kirk, C. P. (2015). Entrepreneurial passion as mediator of the self–efficacy to persistence relationship. *Entrepreneurship Theory and Practice*, *39*(5), 1027–1050. https://doi.org/10.1111/etap.12089

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. (2008a). Dimensions of teacher self-efficacy among Chinese secondary school teachers in Hong Kong. *Educational Psychology*, *28*(2), 181–194. https://doi.org/10.1080/01443410701491833

Chan, D. W. (2008b). 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

Chang, C.-F., Hall, N. C., Lee, S. Y., & Wang, H. (2022). Teachers' social goals and classroom engagement: The mediating role of teachers' self-efficacy. *International Journal of Educational Research*, *113*, 101952. https://doi.org/10.1016/j.ijer.2022.101952

Chang, M.-L., & Engelhard, G. (2016). Examining the teachers' sense of efficacy scale at the item level with Rasch measurement model. *Journal of Psychoeducational Assessment*, *34*(2), 177–191. https://doi.org/10.1177/0734282915593835

Chang, M.-L. (2013). Toward a theoretical model to understand teacher emotions and teacher burnout in the context of student misbehavior: Appraisal, regulation and coping. *Motivation and Emotion*, *37*(4), 799–817. https://doi.org/10.1007/s11031-012-9335-0

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

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

Cheung, H. Y. (2008). Teacher efficacy: A comparative study of Hong Kong and Shanghai primary in-service teachers. *The Australian Educational Researcher*, *35*(1), 103–123. https://doi.org/10.1007/BF03216877

Chiang, C.-H., Hong, J.-C., Chang, K.-W., & Tai, K.-H. (2025). The role of job fit, job stress, and teaching self-efficacy in teacher well-being during the implementation of a new curriculum. *The Asia-Pacific Education Researcher*, *34*(4), 1341–1351. https://doi.org/10.1007/s40299-024-00947-x

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, *18*(8), 947–967. https://doi.org/10.1016/S0742-051X(02)00053-7

Çoban, Ö., Özdemir, N., & Bellibaş, M. Ş. (2023). Trust in principals, leaders' focus on instruction, teacher collaboration, and teacher self-efficacy: testing a multilevel mediation model. *Educational Management Administration & Leadership*, *51*(1), 95-115. https://doi.org/10.1177/1741143220968170

Collado-Mateo, D., Lavín-Pérez, A. M., Peñacoba, C., Del Coso, J., Leyton-Román, M., Luque-Casado, A., Gasque, P., Fernández-Del-Olmo, M. Á., & Amado-Alonso, D. (2021). Key factors associated with adherence to physical exercise in patients with chronic diseases and older adults: An umbrella review. *International Journal of Environmental Research and Public Health*, *18*(4), 1–24. https://doi.org/10.3390/ijerph18042023

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

Coppe, T., Devos, C., & Colognesi, S. (2023). How failure experience combined with high teacher self-aspect importance puts self-efficacy at risk among student teachers. *European Journal of Psychology of Education*, *38*(1), 25–42. https://doi.org/10.1007/s10212-021-00588-4

Cortina, K. S. (2025). Situatedness in educational research. *British Journal of Educational Psychology.* Advance online publication. https://doi.org/10.1111/bjep.70010

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., Barbara, J. M., de Castro, A. F., 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

de Ruiter, J. A., Poorthuis, A. M., Aldrup, K., & Koomen, H. M. (2020). Teachers' emotional experiences in response to daily events with individual students varying in perceived past disruptive behavior. *Journal of School Psychology*, *82*, 85–102. https://doi.org/10.1016/j.jsp.2020.08.005

Demmer, D. H., Hooley, M., Sheen, J., McGillivray, J. A., & Lum, J. A. G. (2017). Sex differences in the prevalence of oppositional defiant disorder during middle childhood: A meta-analysis. *Journal of Abnormal Child Psychology*, *45*(2), 313–325. https://doi.org/10.1007/s10802-016-0170-8

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.

Dicke, T., Parker, P. D., Marsh, H. W., Kunter, M., Schmeck, A., & Leutner, D. (2014). Self-efficacy in classroom management, classroom disturbances, and emotional exhaustion: A moderated mediation analysis of teacher candidates. *Journal of Educational Psychology*, *106*(2), 569–583. https://doi.org/10.1037/a0035504

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

Dirk, J., & Nett, U. E. (2022). Uncovering the situational impact in educational settings: Studies on motivational and emotional experiences. *Learning and Instruction*, *81*, 101661. https://doi.org/10.1016/j.learninstruc.2022.101661

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

Duan, S., Bissaker, K., & Xu, Z. (2024). Correlates of teachers' classroom management self-efficacy: A systematic review and meta-analysis. *Educational Psychology Review*, *36*(2). https://doi.org/10.1007/s10648-024-09881-2

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

Eder-Karavaya, A., Lohr, J., & Treutner, D. (2021). Förderung der berufsspezifischen Selbstwirksamkeitserwartung im Lehramtsstudium. Entwicklung und Evaluation eines Trainings für Lehramtsstudierende [Promoting professional self-efficacy in pre-service teachers. Development and evaluation of a training program]. *Psychologie in Erziehung und Unterricht*, *68*(4), 233–247. https://doi.org/10.2378/peu2021.art19d

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

Embse, N. P., Sandilos, L. E., Pendergast, L., & Mankin, A. (2016). Teacher stress, teaching-efficacy, and job satisfaction in response to test-based educational accountability policies. *Learning and Individual Differences*, *50*, 308–317. https://doi.org/10.1016/j.lindif.2016.08.001

Engels, M. C., Colpin, H., van Leeuwen, K., Bijttebier, P., van den Noortgate, W., Claes, S., Goossens, L., & Verschueren, K. (2016). Behavioral engagement, peer status, and teacher-student relationships in adolescence: A longitudinal study on reciprocal influences. *Journal of Youth and Adolescence*, *45*(6), 1192–1207. https://doi.org/10.1007/s10964-016-0414-5

Enochs, L. G., & Riggs, I. M. (1990). Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *School Science and Mathematics*, *90*(8), 694–706. https://doi.org/10.1111/j.1949-8594.1990.tb12048.x

Fackler, S., & Malmberg, L.-E. (2016). Teachers' self-efficacy in 14 OECD countries: Teacher, student group, school and leadership effects. *Teaching and Teacher Education*, *56*, 185–195. https://doi.org/10.1016/j.tate.2016.03.002

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

Fauth, B., Wagner, W., Bertram, C., Göllner, R., Roloff, J., Lüdtke, O., Polikoff, M. S., Klusmann, U., & Trautwein, U. (2020). Don't blame the teacher? The need to account for classroom characteristics in evaluations of teaching quality. *Journal of Educational Psychology*, *112*(6), 1284–1302. https://doi.org/10.1037/edu0000416

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

Fisher, R. J., & Katz, J. E. (2000). Social-desirability bias and the validity of self-reported values. *Psychology and Marketing*, *17*(2), 105–120. https://doi.org/10.1002/(SICI)1520-6793(200002)17:2<105::AID-MAR3>3.0.CO;2-9

Fives, H., & Buehl, M. M. (2010). Examining the factor structure of the Teachers' Sense of Efficacy Scale. *Journal of Experimental Education*, *78*(1), 118–134. https://doi.org/10.1080/00220970903224461

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

Forghani-Arani, N., Cerna, L., & Bannon, M. (2019). *The lives of teachers in diverse classrooms*. (OECD Education Working Paper, No. 198). https://doi.org/10.1787/8c26fee5-en

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Frenzel, A. C., Becker-Kurz, B., Pekrun, R., Goetz, T., & Lüdtke, O. (2018). Emotion transmission in the classroom revisited: A reciprocal effects model of teacher and student enjoyment. *Journal of Educational Psychology*, *110*(5), 628–639. https://doi.org/10.1037/edu0000228

Frenzel, A. C., Fiedler, D., Marx, A. K. G., Reck, C., & Pekrun, R. (2020). Who enjoys teaching, and when? Between- and within-person evidence on teachers' appraisal-emotion links. *Frontiers in Psychology*, *11*, Article 1092. https://doi.org/10.3389/fpsyg.2020.01092

Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R., & Sutton, R. E. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology*, *101*(3), 705–716. https://doi.org/10.1037/a0014695

Fries, L., Son, J. Y., Givvin, K. B., & Stigler, J. W. (2021). Practicing connections: A framework to guide instructional design for developing understanding in complex domains. *Educational Psychology Review*, *33*(2), 739–762. https://doi.org/10.1007/s10648-020-09561-x

Fung, F., Tan, C. Y., & Chen, G. (2018). Student engagement and mathematics achievement: Unraveling main and interactive effects. *Psychology in the Schools*, *55*(7), 815–831. https://doi.org/10.1002/pits.22139

Garwood, J. D., & Harris, A. H. (2020). Screencast-delivered professional development targeting teachers' self-efficacy and beginning-of-the-year classroom management practices. *Australasian Journal of Special and Inclusive Education*, *44*(1), 60–72. https://doi.org/10.1017/jsi.2020.3

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. https://doi.org/10.7202/1020976ar

Geerlings, J., Thijs, J. T., & Verkuyten, M. (2018). Teaching in ethnically diverse classrooms: Examining individual differences in teacher self-efficacy. *Journal of School Psychology*, *67*, 134–147. https://doi.org/10.1016/j.jsp.2017.12.001

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

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*(1), 72–91. https://doi.org/10.1037/a0032138

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

Geukes, K., Nestler, S., Hutteman, R., Dufner, M., Küfner, A. C. P., Egloff, B., Denissen, J. J. A., & Back, M. D. (2017). Puffed-up but shaky selves: State self-esteem level and variability in narcissists. *Journal of Personality and Social Psychology*, *112*(5), 769–786. https://doi.org/10.1037/pspp0000093

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: a construct validation. *Journal of Educational Psychology*, *76*(4), 569–582. https://doi.org/10.1037/0022-0663.76.4.569

Gkolia, A., Dimitrios, B. A., & Koustelious, A. (2016). Background characteristics as predictors of Greek teachers' self-efficacy. *International Journal of Educational Management*, *30*(3), 460–472. https://doi.org/10.1108/IJEM-03-2014-0040

Glock, S., Kovacs, C., & Pit-Ten, C. (2019). Teachers' attitudes towards ethnic minority students: Effects of schools' cultural diversity. *The British Journal of Educational Psychology*, *89*(4), 616–634. https://doi.org/10.1111/bjep.12248

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*(S1), 115–136. https://doi.org/10.1007/s11618-017-0727-5

Grammatikopoulos, V [Vasilis], Tsigilis, N., Gregoriadis, A., & Bikos, K. (2013). Evaluating an induction training program for Greek teachers using an adjusted level model approach. *Studies in Educational Evaluation*, *39*(4), 225–231. https://doi.org/10.1016/j.stueduc.2013.09.002

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. *The 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, Peter, 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

Gruber-Baldini, A. L., Velozo, C., Romero, S., & Shulman, L. M. (2017). Validation of the PROMIS® measures of self-efficacy for managing chronic conditions. *Quality of Life Research*, *26*(7), 1915–1924. https://doi.org/10.1007/s11136-017-1527-3

Guin, K. (2004). Chronic teacher turnover in urban elementary schools. *Education Policy Analysis Archives*, *12*(42), 1–30. https://doi.org/10.14507/epaa.v12n42.2004

Gundlach, H. A., Slemp, G. R., & Hattie, J. (2024). A meta-analysis of the antecedents of teacher turnover and retention. *Educational Research Review*, *44*, 100606. https://doi.org/10.1016/j.edurev.2024.100606

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.

Hajovsky, D. B., Chesnut, S. R., & Jensen, K. M. (2020). The role of teachers' self-efficacy beliefs in the development of teacher-student relationships. *Journal of School Psychology*, *82*, 141–158. https://doi.org/10.1016/j.jsp.2020.09.001

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

Hascher, T., & Hagenauer, G. (2016). Openness to theory and its importance for pre-service teachers' self-efficacy, emotions, and classroom behaviour in the teaching practicum. *International Journal of Educational Research*, *77*, 15–25. https://doi.org/10.1016/j.ijer.2016.02.003

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

Hausen, J. E., Möller, J., Greiff, S., & Niepel, C. (2022). Students' personality and state academic self-concept: Predicting differences in mean level and within-person variability in everyday school life. *Journal of Educational Psychology*, *114*(6), 1394–1411. https://doi.org/10.1037/edu0000760

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

Heinschel, A., Henschel, S., & Rjosk, C. (2024). Links between instructional quality and classroom composition in classes taught by the same teacher. *Learning and Instruction*, *94*, 102005. https://doi.org/10.1016/j.learninstruc.2024.102005

Henson, R. K. (2002). From adolescent angst to adulthood: substantive implications and measurement dilemmas in the development of teacher efficacy research. *Educational Psychologist*, *37*(3), 137–150. https://doi.org/10.1207/S15326985EP3703_1

Hettinger, K., Lazarides, R., & Schiefele, U. (2023a). Longitudinal relations between teacher self-efficacy and student motivation through matching characteristics of perceived teaching practice. *European Journal of Psychology of Education*, *39*, 1299–1325. https://doi.org/10.1007/s10212-023-00744-y

Hettinger, K., Lazarides, R., & Schiefele, U. (2023b). Motivational climate in mathematics classrooms: teacher self-efficacy for student engagement, student- and teacher-reported emotional support and student interest. *ZDM - Mathematics Education*, *55*(2), 413–426. https://doi.org/10.1007/s11858-022-01430-x

Hoffman, J. V., Wetzel, M. M., Maloch, B., Greeter, E., Taylor, L., DeJulio, S., & Vlach, S. K. (2015). What can we learn from studying the coaching interactions between cooperating teachers and preservice teachers? A literature review. *Teaching and Teacher Education*, *52*, 99–112. https://doi.org/10.1016/j.tate.2015.09.004

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., & Prestele, E. (2021). Teacher self-efficacy and self-reported cognitive activation and classroom management: A multilevel perspective on the role of school characteristics. *Learning and Instruction*, *76*, Article 101513. https://doi.org/10.1016/j.learninstruc.2021.101513

Holzberger, D., Täschner, J., & Hartl, A. (2025). Natürliche und gezielte Veränderung von Selbstwirksamkeit: Wie individuell sind die Prozesse? [Natural and intentional changes in self-efficacy: How individualized are these processes?]. In A. Bach & G. Hagenauer (Eds.), *Motivation und Emotion von Lehrpersonen* (pp. 257–274). Waxmann Verlag.

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

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

Hoogendijk, K., Tick, N. T., Hofman, A. W. H., Windig, R. J., Holland, J. G., Severiens, S. E., Vuijk, P., & van Veen, D. (2023). The impact of teachers' self-efficacy and classroom externalising problem behaviours on emotional exhaustion: Between- and within-person associations. *Current Psychology*, *42*(26), 22989–23002. https://doi.org/10.1007/s12144-022-03319-0

Houben, M., & Kuppens, P. (2020). Emotion dynamics and the association with depressive features and borderline personality disorder traits: Unique, specific, and prospective relationships. *Clinical Psychological Science*, *8*(2), 226–239. https://doi.org/10.1177/2167702619871962

Hox, J., & McNeish, D. (2020). Small samples in multilevel modeling. In R. van de Schoot (Ed.), *Small sample size solutions: A guide for applied researchers and practitioners* (pp. 215–225). Routledge.

Huang, P., & Zhou, M. (2024). A meta-analytical investigation on the relationship between emotional labour and teacher efficacy in the Chinese context. *Educational Studies*, 1–24. https://doi.org/10.1080/03055698.2024.2369879

Hughes, J. N., Wu, W., & West, S. G. (2011). Teacher performance goal practices and elementary students' behavioral engagement: A developmental perspective. *Journal of School Psychology*, *49*, 1–23. https://doi.org/10.1016/j.jsp.2010.09.003

Hußner, I., Lazarides, R., & Symes, W. (2024). The relation between self-efficacy and teaching behaviour: A video-based analysis of student teachers. *Learning and Instruction*, *91*, 101880. https://doi.org/10.1016/j.learninstruc.2024.101880

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

James, W. (1890). *The principles of psychology* (Vol. 1). Henry Holt and Co. https://doi.org/10.1037/10538-000

Jerrim, J., Sims, S., & Oliver, M. (2023). Teacher self-efficacy and pupil achievement: much ado about nothing? International evidence from TIMSS. *Teachers and Teaching*, *29*(2), 220–240. https://doi.org/10.1080/13540602.2022.2159365

Jones, S. M., & Dindia, K. (2004). A meta-analytic perspective on sex equity in the classroom. *Review of Educational Research*, *74*(4), 443–471. https://doi.org/10.3102/00346543074004443

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*(Sup1). 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.

Karimi, M. N. (2011). The effects of professional development initiatives on EFL teachers' degree of self-efficacy. *Australian Journal of Teacher Education*, *36*(6), 50–62. https://doi.org/10.14221/ajte.2011v36n6.6

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.

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. https://doi.org/10.18844/wjet.v10i2.3456

Keller-Schneider, M., Zhong, H. F., & Yeung, A. S. (2020). Competence and challenge in professional development: Teacher perceptions at different stages of career. *Journal of Education for Teaching*, *46*(1), 36–54. https://doi.org/10.1080/02607476.2019.1708626

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

Kilday, J. E., Lenser, M. L., & Miller, A. D. (2016). Considering students in teachers' self-efficacy: Examination of a scale for student-oriented teaching. *Teaching and Teacher Education*, *56*, 61–71. https://doi.org/10.1016/j.tate.2016.01.025

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., Bong, M., Usher, E. L., Chong, W. H., Huan, V. S., Wong, I., & Georgiou, T. (2009). Exploring the validity of a teachers' self-efficacy scale in five countries. *Contemporary Educational Psychology*, *34*(1), 67–76. https://doi.org/10.1016/j.cedpsych.2008.08.001

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

Kleinsasser, R. C. (2014). Teacher efficacy in Teaching and Teacher Education. *Teaching and Teacher Education*, *44*, 168–179. https://doi.org/10.1016/j.tate.2014.07.007

Knight, D. S. (2012). Assessing the cost of instructional coaching. *Journal of Education Finance*, *38*(1), 52–80. https://doi.org/10.1353/jef.2012.0010

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

Konstantinidou, E., & Scherer, R. (2022). Teaching with technology: A large-scale, international, and multilevel study of the roles of teacher and school characteristics. *Computers & Education*, *179*, Article 104424. https://doi.org/10.1016/j.compedu.2021.104424

Kozhevnikov, M. (2007). Cognitive styles in the context of modern psychology: Toward an integrated framework of cognitive style. *Psychological Bulletin*, *133*(3), 464–481. https://doi.org/10.1037/0033-2909.133.3.464

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

Kupers, E., Loopers, J., Albers, C., Bakker, A., & Minnaert, A. (2023). An experience sampling study on the links between daily teacher self-efficacy, need-supportive teaching and student intrinsic motivation. *Frontiers in Psychology*, *14*, 1159108. https://doi.org/10.3389/fpsyg.2023.1159108

Kurt, G. (2017). Implementing the flipped classroom in teacher education: Evidence from Turkey. *Educational Technology & Society*, *20*(1), 211–221.

Lauermann, F., Benden, D., & Evers, M. (2020). Motive und Interessen [Motives and interests]. In C. Cramer, J. König, M. Rothland, & S. Blömeke (Eds.), *Handbuch Lehrerinnen- und Lehrerbildung* (pp. 791–797). Verlag Julius Klinkhardt.

Lauermann, F., & Berger, J.-L. (2021). Linking teacher self-efficacy and responsibility with teachers' self-reported and student-reported motivating styles and student engagement. *Learning and Instruction*, *76*, Article 101441. https://doi.org/10.1016/j.learninstruc.2020.101441

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

Lazarides, R., Fauth, B., Gaspard, H., & Göllner, R. (2021). Teacher self-efficacy and enthusiasm: Relations to changes in student-perceived teaching quality at the beginning of secondary education. *Learning and Instruction*, *73*, Article 101435. https://doi.org/10.1016/j.learninstruc.2020.101435

Lazarides, R., & Schiefele, U. (2024). Addressing the reciprocal nature of effects in teacher motivation research: A study on relations among teacher motivation, student-reported teaching, and student enjoyment and achievement. *Learning and Instruction*, *90*, Article 101862. https://doi.org/10.1016/j.learninstruc.2023.101862

Lazarides, R., Schiefele, U., Hettinger, K., & Frommelt, M. C. (2023). Tracing the signal from teachers to students: How teachers' motivational beliefs longitudinally relate to student interest through student-reported teaching practices. *Journal of Educational Psychology*, *115*(2), 290–308. https://doi.org/10.1037/edu0000777

Lazarides, R., & Warner, L. M. (2016). Teacher self-efficacy. In G. W. Noblit (Ed.), *Oxford research encyclopedia of education.* Oxford University Press. https://doi.org/10.1093/acrefore/9780190264093.013.890

Lazarides, R., Watt, H. M. G., & Richardson, P. W. (2020). Teachers' classroom management self-efficacy, perceived classroom management and teaching contexts from beginning until mid-career. *Learning and Instruction*, *69*, Article 101346. https://doi.org/10.1016/j.learninstruc.2020.101346

Learning Policy Institute. *2024 update: what's the cost of teacher turnover?* [Interactive tool]. https://learningpolicyinstitute.org/product/2024-whats-cost-teacher-turnover

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

Lee, W. C., Chen, V. D.-T., & Wang, L.-Y. (2017). A review of research on teacher efficacy beliefs in the learner-centred pedagogy context: Themes, trends and issues. *Asia Pacific Education Review*, *18*(4), 559–572. https://doi.org/10.1007/s12564-017-9501-x

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 September 2020). Cochrane. www.training.cochrane.org/handbook

Lei, H., Cui, Y., & Zhou, W. (2018). Relationships between student engagement and academic achievement: A meta-analysis. *Social Behavior and Personality: An International Journal*, *46*, 517–528. https://doi.org/10.2224/sbp.7054

Leijen, Ä., Pedaste, M., Baucal, A., Poom-Valickis, K., & Lepp, L. (2024). What predicts instructional quality and commitments to teaching: Self-efficacy, pedagogical knowledge or integration of the two? *Frontiers in Psychology*, *15*, Article 1287313. https://doi.org/10.3389/fpsyg.2024.1287313

Leising, D., Thielmann, I., Glöckner, A., Gärtner, A., & Schönbrodt, F. (2022). Ten steps toward a better personality science – how quality may be rewarded more in research evaluation. *Personality Science*, *3*(1), Article e6029. https://doi.org/10.5964/ps.6029

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

Li, R., & Yao, M. (2022). What promotes teachers' turnover intention? Evidence from a meta-analysis. *Educational Research Review*, *37*, Article 100477. https://doi.org/10.1016/j.edurev.2022.100477

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

Lin, L. (2018). Bias caused by sampling error in meta-analysis with small sample sizes. *PLOS ONE*, *13*(9), Article e0204056. https://doi.org/10.1371/journal.pone.0204056

Loewenberg 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

Lüdtke, O., Marsh, H. W., Robitzsch, A., & Trautwein, U. (2011). A 2 × 2 taxonomy of multilevel latent contextual models: Accuracy-bias trade-offs in full and partial error correction models. *Psychological Methods*, *16*(4), 444–467. https://doi.org/10.1037/a0024376

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. O. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, *13*(3), 203–229. https://doi.org/10.1037/a0012869

Ma, K., Chutiyami, M., & Nicoll, S. (2022). Transitioning into the first year of teaching: Changes and sources of teacher self-efficacy. *The Australian Educational Researcher*, *49*(5), 943–960. https://doi.org/10.1007/s13384-021-00481-5

Maas, C. J. M., & Hox, J. J. (2004). Robustness issues in multilevel regression analysis. *Statistica Neerlandica*, *58*(2), 127–137. https://doi.org/10.1046/j.0039-0402.2003.00252.x

Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, *1*(3), 86–92. https://doi.org/10.1027/1614-1881.1.3.86

Malinen, O.-P., Närhi, V., & Savolainen, H. (2024). The effect of teacher self-efficacy in behaviour management on classroom behavioural climate: a longitudinal multilevel cross-lagged analysis. *Educational Psychology*, *44*(4), 415–435. https://doi.org/10.1080/01443410.2024.2359672

Malmberg, L.-E., Hagger, H., & Webster, S. (2014). Teachers' situation-specific mastery experiences: teacher, student group and lesson effects. *European Journal of Psychology of Education*, *29*(3), 429–451. https://doi.org/10.1007/s10212-013-0206-1

Mamaril, N. A., Usher, E. L., Li, C. R., Economy, D. R., & Kennedy, M. S. (2016). Measuring undergraduate students' engineering self-efficacy: A validation study. *Journal of Engineering Education*, *105*(2), 366–395. https://doi.org/10.1002/jee.20121

Marlow, R., Hansford, L., Edwards, V., Ukoumunne, O. C., Norman, S., Ingarfield, S., Sharkey, S., Logan, S., & Ford, T. (2015). Teaching classroom management – a potential public health intervention? *Health Education*, *115*(3/4), 230–248. https://doi.org/10.1108/HE-03-2014-0030

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

Martin, A. J., Malmberg, L.-E., Pakarinen, E., Mason, L., & Mainhard, T. (2023). The potential of biophysiology for understanding motivation, engagement and learning experiences. *The British Journal of Educational Psychology*, *93*(51), 1–9. https://doi.org/10.1111/bjep.12584

Martin, A. J., Mansour, M., & Malmberg, L.-E. (2020). What factors influence students' real-time motivation and engagement? An experience sampling study of high school students using mobile technology. *Educational Psychology*, *40*(9), 1113–1135. https://doi.org/10.1080/01443410.2018.1545997

Martins, M., Costa, J., & Onofre, M. (2015). Practicum experiences as sources of pre-service teachers' self-efficacy. *European Journal of Teacher Education*, *38*(2), 263–279. https://doi.org/10.1080/02619768.2014.968705

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

McGee, J. E., Peterson, M., Mueller, S. L., & Sequeira, J. M. (2009). Entrepreneurial self–efficacy: Refining the measure. *Entrepreneurship Theory and Practice*, *33*(4), 965–988. https://doi.org/10.1111/j.1540-6520.2009.00304.x

McKeithen, K. B., Reitman, J. S., Rueter, H. H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology*, *13*(3), 307–325. https://doi.org/10.1016/0010-0285(81)90012-8

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review*, *28*(2), 295–314. https://doi.org/10.1007/s10648-014-9287-x

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

Meschede, N., Hardy, I. (2020). Selbstwirksamkeitserwartungen von Lehramtsstudierenden zum adaptiven Unterrichten in heterogenen Lerngruppen [Preservice teachers' self-efficacy of adaptive teaching in heterogeneous classrooms]. *Zeitschrift für Erziehungswissenschaft*, *23*(3), 565–589. https://doi.org/10.1007/s11618-020-00949-7

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

Miller, K., Espinosa, T., Araujo, I., & Gallegos, I. (2023). Response-shift bias in student self-efficacy during an actively taught physics course. *Physical Review Physics Education Research*, *19*, Article 020167. https://doi.org/10.1103/PhysRevPhysEducRes.19.020167

Mintz, J., Hick, P., Solomon, Y., Matziari, A., Ó'Murchú, F., Hall, K., Cahill, K., Curtin, C., Anders, J., & Margariti, D. (2020). The reality of reality shock for inclusion: How does teacher attitude, perceived knowledge and self-efficacy in relation to effective inclusion in the classroom change from the pre-service to novice teacher year? *Teaching and Teacher Education*, *91*, Article 103042. https://doi.org/10.1016/j.tate.2020.103042

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*, *40*, 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*, 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

Mu, G. M., Gordon, D., Liang, J., Zhao, L., Alonso, R. A., Juri, M. Z., Zhang, X., Vo, H., Gao, D., Hu, Y., & Xing, C. (2024). A meta-analysis of the correlation between teacher self-efficacy and teacher resilience: Concerted growth and contextual variance. *Educational Research Review*, *45*, 100645. https://doi.org/10.1016/j.edurev.2024.100645

Müller, J., Böhnke, A., & Thiel, F. (2024). Lehramtsstudierende im ersten Studienjahr – Die Veränderung ihrer Fehlerorientierung im Zusammenhang mit der Lehrkraftselbstwirksamkeit [First-year student teachers—The change of their error orientation in connection with the teacher self-efficacy]. *Zeitschrift für Erziehungswissenschaft*, *27*(1), 169–195. https://doi.org/10.1007/s11618-023-01203-6

Murayama, K., Goetz, T., Malmberg, L.-E., Pekrun, R., Tanaka, A., & Martin, A. J. (2017). Within-person analysis in educational psychology: Importance and illustrations. In D. W. Putwain & K. Smart (Eds.), *BJEP Monograph Series II: Part 12 The Role of Competence and Beliefs in Teaching and Learning* (pp. 71–87). British Psychological Society. https://doi.org/10.53841/bpsmono.2017.cat2023.6

Murayama, K., & Keyserlingk, L. von (2025). A critical analysis of the current motivation theories in educational psychology: Why the same theories continue to dominate. *Educational Psychologist.* Advance online publication. https://doi.org/10.1080/00461520.2025.2473894

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide [*Computer software].

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

Nguyen, T. D., Cannata, M., & Miller, J. (2018). Understanding student behavioral engagement: Importance of student interaction with peers and teachers. *The Journal of Educational Research*, *111*(2), 163–174. https://doi.org/10.1080/00220671.2016.1220359

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., . . . Yarkoni, T. (2015). Promoting an open research culture: Author guidelines for journals could help to promote transparency, openness, and reproducibility. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

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

OECD. (2019). *TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners*. https://www.oecd-ilibrary.org/content/publication/1d0bc92a-en https://doi.org/10.1787/1d0bc92a-en

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: Asystematic review with meta-analysis. *Frontiers in Psychology*, *12*, Article 677217. https://doi.org/10.3389/fpsyg.2021.677217

Olivier, E., Lazariuk, L., Archambault, I., & Morin, A. J. S. (2024). Teacher emotional exhaustion: The synergistic roles of self-efficacy and student–teacher relationships. *Social Psychology of Education*, *27*(1), 1–22. https://doi.org/10.1007/s11218-023-09826-7

Oort, F. J., Visser, M. R. M., & Sprangers, M. A. G. (2009). Formal definitions of measurement bias and explanation bias clarify measurement and conceptual perspectives on response shift. *Journal of Clinical Epidemiology*, *62*(11), 1126–1137. https://doi.org/10.1016/j.jclinepi.2009.03.013

Ortenburger, A., Gäckle, S., Menge, C., & Franz, S. (2023). *Skalenhandbuch Lehramtsstudierenden-Panel*. (LIFBi Working Paper No. 111). https://doi.org/10.5157/LIFBI:WP111:1.0

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), 89. https://doi.org/10.1186/s13643-021-01626-4

Pajares, F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, *62*(3), 307–332. https://doi.org/10.3102/00346543062003307

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, *66*(4), 543–578. https://doi.org/10.3102/00346543066004543

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

Patall, E. A. (2021). Implications of the open science era for educational psychology research syntheses. *Educational Psychologist*, *56*(2), 142–160. https://doi.org/10.1080/00461520.2021.1897009

Patzl, S., Oberleiter, S., & Pietschnig, J. (2024). Mirror, mirror on the wall: A meta-analysis on the validity of self-assessed intelligence through the lens of the multiverse. *Journal of Intelligence*, *12*(9), 81. https://doi.org/10.3390/jintelligence12090081

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*, *93*(S1), 227–238. https://doi.org/10.1111/bjep.12575

Pereira Kastens, C., Döring-Seipel, E., & Nolle, T. (2020). Selbstwirksamkeit erlangen, Belastung reduzieren? Effekte des Feedbackverhaltens der Ausbilder/innen in Unterrichtsnachbesprechungen [Acquiring self-efficacy and reducing emotional exhaustion? – Effects of teacher educators' feedback]. *Journal for Educational Research Online*, *12*(1), 67–90. https://doi.org/10.25656/01:19119

Perera, H. N., Calkins, C., & Part, R. (2019). Teacher self-efficacy profiles: Determinants, outcomes, and generalizability across teaching level. *Contemporary Educational Psychology*, *58*, 186–203. https://doi.org/10.1016/j.cedpsych.2019.02.006

Perera, H. N., Granziera, H., & McIlveen, P. (2018). Profiles of teacher personality and relations with teacher self-efficacy, work engagement, and job satisfaction. *Personality and Individual Differences*, *120*, 171–178. https://doi.org/10.1016/j.paid.2017.08.034

Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching and Learning Practice*, *21*(06). https://doi.org/10.53761/q3azde36

Perry, T., Morris, R., & Lea, R. (2022). A decade of replication study in education? A mapping review (2011–2020). *Educational Research and Evaluation*, *27*(1-2), 12–34. https://doi.org/10.1080/13803611.2021.2022315

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*, 1486. https://doi.org/10.3389/fpsyg.2016.01486

Pfitzner-Eden, F., Thiel, F., & Horsley, J. (2014). An adapted measure of teacher self-efficacy for preservice teachers: Exploring its validity across two countries. *Zeitschrift für Pädagogische Psychologie*, *28*(3), 83–92. https://doi.org/10.1024/1010-0652/a000125

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*, *90*(1), 24-46. https://doi.org/10.3102/0034654319877153

Piwowar, V., Thiel, F., & Ophardt, D. (2013). Training inservice 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

Plucker, J. A., & Makel, M. C. (2021). Replication is important for educational psychology: Recent developments and key issues. *Educational Psychologist*, *56*(2), 90–100. https://doi.org/10.1080/00461520.2021.1895796

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*, 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

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

Praetorius, A.-K., Lauermann, F., Klassen, R. M., Dickhäuser, O., Janke, S., & Dresel, M. (2017). Longitudinal relations between teaching-related motivations and student-reported teaching quality. *Teaching and Teacher Education*, *65*, 241–254. https://doi.org/10.1016/j.tate.2017.03.023

Pressley, T. (2021). Returning to teaching during COVID-19: An empirical study on elementary teachers' self-efficacy. *Psychology in the Schools*, *58*(8), 1611–1623. https://doi.org/10.1002/pits.22528

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

Putman, S. M. (2012). Examining variations in programmatic delivery on teaching candidates' sense of efficacy. *Research in the Schools*, *19*(2), 45–61.

Putwain, D. W., Symes, W., Nicholson, L. J., & Becker, S. (2018). Achievement goals, behavioural engagement, and mathematics achievement: A mediational analysis. *Learning and Individual Differences*, *68*, 12–19. https://doi.org/10.1016/j.lindif.2018.09.006

Raudenbush, S. W., Rowan, B., & Cheong, Y. F. (1992). Contextual effects on the self-perceived efficacy of high school teachers. *Sociology of Education*, *65*(2), 150–167. https://doi.org/10.2307/2112680

Reeve, J. (2018). *Understanding motivation and emotion* (7th ed.). Wiley.

Richter, D., Kunter, M., Lüdtke, O., Klusmann, U., Anders, Y., & Baumert, J. (2013). How different mentoring approaches affect beginning teachers' development in the first years of practice. *Teaching and Teacher Education,*, *36*, 166–177.

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, *50*(1), 4–36. https://doi.org/10.3102/0002831212463813

Roorda, D. L., Jorgensen, T. D., & Koomen, H. M. (2019). Different teachers, different relationships? Student-teacher relationships and engagement in secondary education. *Learning and Individual Differences*, *75*, 101761. https://doi.org/10.1016/j.lindif.2019.101761

Ross, J. A., Cousins, J. B., & Gadalla, T. (1996). Within-teacher predictors of teacher efficacy. *Teaching and Teacher Education*, *12*(4), 385–400. https://doi.org/10.1016/0742-051X(95)00046-M

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, *80*(1), 1–28. https://doi.org/10.1037/h0092976

Ruan, J., Nie, Y., Hong, J. Y., Monobe, G., Zheng, G., Kambara, H., & You, S. (2015). Cross-cultural validation of Teachers' Sense of Efficacy Scale in three Asian countries. *Journal of Psychoeducational Assessment*, *33*(8), 769–779. https://doi.org/10.1177/0734282915574021

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

Ruzek, E. A., Hafen, C. A., Allen, J. P., Gregory, A., Mikami, A. Y., & Pianta, R. C. (2016). How teacher emotional support motivates students: The mediating roles of perceived peer relatedness, autonomy support, and competence. *Learning and Instruction*, *42*, 95–103. https://doi.org/10.1016/j.learninstruc.2016.01.004

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. (2020). 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

Scherer, R., Siddiq, F., Howard, S. K., & Tondeur, J. (2023). The more experienced, the better prepared? New evidence on the relation between teachers' experience and their readiness for online teaching and learning. *Computers in Human Behavior*, *139*, Article 107530. https://doi.org/10.1016/j.chb.2022.107530

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

Schipper, T. M., Vries, S. de, 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

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

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

Schwab, S., Kulmhofer-Bommer, A., Hoffmann, L., & Goldan, J. (2021). Maths, German, and English teachers' student specific self-efficacy – is it a matter of students' characteristics? *Educational Psychology*, *41*(10), 1224–1240. https://doi.org/10.1080/01443410.2021.1934405

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

Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Ed.). (2025, February 2). *Lehrkräfteeinstellungsbedarf und -angebot in der Bundesrepublik Deutschland 2024 - 2035: Zusammengefasste Modellrechnungen der Länder [Teacher recruitment demand and supply in the Federal Republic of Germany 2024-2035: Summary of model calculations by the Länder]* (Statistische Veröffentlichungen der Bildungsministerkonferenz). www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok_2_Bericht_LEB_LEA_2024.pdf

Shanafelt, T. D., Goh, J., & Sinsky, C. A. (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

Sharma, U., Loreman, T., & Forlin, C. (2012). Measuring teacher efficacy to implement inclusive practices. *Journal of Research in Special Educational Needs*, *12*(1), 12–21. https://doi.org/10.1111/j.1471-3802.2011.01200.x

Sheehan, C. Z., & Moore, L. L. (2019). Teacher self-efficacy in SBAE methods coursework: A mixed methods study. *Journal of Agricultural Education*, *60*(3), 219–231. https://doi.org/10.5032/jae.2019.03219

Sheeran, P., Maki, A., Montanaro, E., Avishai-Yitshak, A., Bryan, A., Klein, W. M. P., Miles, E., & Rothman, A. J. (2016). The impact of changing attitudes, norms, and self-efficacy on health-related intentions and behavior: A meta-analysis. *Health Psychology*, *35*(11), 1178–1188. https://doi.org/10.1037/hea0000387

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

Siwatu, K. O. (2007). Preservice teachers' culturally responsive teaching self-efficacy and outcome expectancy beliefs. *Teaching and Teacher Education*, *23*(7), 1086–1101. https://doi.org/10.1016/j.tate.2006.07.011

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. (2011). Teacher job satisfaction and motivation to leave the teaching profession: Relations with school context, feeling of belonging, and emotional exhaustion. *Teaching and Teacher Education*, *27*(6), 1029–1038. https://doi.org/10.1016/j.tate.2011.04.001

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

Skinner, E. A., Furrer, C. J., Marchand, G., & Kindermann, T. A. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology*, *100*(4), 765–781. https://doi.org/10.1037/a0012840

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

Stevens, T., Aguirre-Munoz, Z., Harris, G., Higgins, R., & Liu, X. (2013). Middle level mathematics teachers' self-efficacy growth through professional development: Differences based on mathematical background. *Australian Journal of Teacher Education*, *38*(4), Article 9. https://doi.org/10.14221/ajte.2013v38n4.3

Suchodoletz, A. von, 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

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

Swann, W. B., Chang-Schneider, C., & Larsen McClarty, K. (2007). Do people's self-views matter? Self-concept and self-esteem in everyday life. *American Psychologist*, *62*(2), 84–94. https://doi.org/10.1037/0003-066X.62.2.84

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

Täschner, J., Dicke, T., Reinhold, S., & Holzberger, D. (2024). "Yes, I Can!" A Systematic Review and Meta-Analysis of Intervention Studies Promoting Teacher Self-Efficacy. *Review of Educational Research*, *95*(1), 3-52. https://doi.org/10.3102/00346543231221499

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

ten Hagen, I., Lauermann, F., Wigfield, A., & Eccles, J. S. (2022). Can I teach this student? A multilevel analysis of the links between teachers' perceived effectiveness, interest-supportive teaching, and student interest in math and reading. *Contemporary Educational Psychology*, *69*, Article 102059. https://doi.org/10.1016/j.cedpsych.2022.102059

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*, 1–15. https://doi.org/10.1080/19415257.2020.1763433

Thien, L. M., & Liu, P. (2024). Linear and nonlinear relationships between instructional leadership and teacher professional learning through teacher self-efficacy as a mediator: A partial least squares analysis. *Humanities and Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599-023-02500-5

Thommen, D., Grob, U., Lauermann, F., Klassen, R. M., & Praetorius, A.-K. (2022). Different levels of context-specificity of teacher self-efficacy and their relations with teaching quality. *Frontiers in Psychology*, *13*, Article 857526. https://doi.org/10.3389/fpsyg.2022.857526

Thomson, M. M., Huggins, E., Carrier, S. J., & Gray, D. (2022). Developmental trajectories for novice teachers: teaching efficacy, instructional beliefs, and domain knowledge. *International Journal of Science Education*, *44*(8), 1277–1298. https://doi.org/10.1080/09500693.2022.2075948

Thurm, D., & Barzel, B. (2020). Effects of a professional development program for teaching mathematics with technology on teachers' beliefs, self-efficacy and practices. *ZDM - Mathematics Education*, *52*, 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. (2021). Teacher job satisfaction: the importance of school working conditions and teacher characteristics. *Educational Review*, *73*(1), 1–27. https://doi.org/10.1080/00131911.2019.1705247

Tschannen-Moran, M., & Johnson, D. (2011). Exploring literacy teachers' self-efficacy beliefs: Potential sources at play. *Teaching and Teacher Education*, *27*(4), 751–761. https://doi.org/10.1016/j.tate.2010.12.005

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

Tsigilis, N., Koustelios, A., & Grammatikopoulos, V [Vasilios] (2010). Psychometric properties of the Teachers' Sense of Efficacy Scale within the Greek educational context. *Journal of Psychoeducational Assessment*, *28*(2), 153–162. https://doi.org/10.1177/0734282909342532

Tsouloupas, C. N., Carson, R. L., Matthews, R., Grawitch, M. J., & Barber, L. K. (2010). Exploring the association between teachers' perceived student misbehaviour and emotional exhaustion: the importance of teacher efficacy beliefs and emotion regulation. *Educational Psychology*, *30*(2), 173–189. https://doi.org/10.1080/01443410903494460

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

van Eycken, L., Demanet, J., & van Houtte, M. (2024). He's up to no good, is he? Teachers' self-efficacy as related to gender role attitudes and schools' sex composition. *European Journal of Psychology of Education*, *39*, 3867–3892. https://doi.org/10.1007/s10212-024-00828-3

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. (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. (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., & 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

Voracek, M., Kossmeier, M., & Tran, U. S. (2019). Which data to meta-analyze, and how? *Zeitschrift für Psychologie*, *227*(1), 64–82. https://doi.org/10.1027/2151-2604/a000357

Wagler, R., & Moseley, C. (2005). Preservice teacher efficacy: Effects of a secondary education methods course and student teaching. *Teacher Education and Practice*, *18*(4), 442–457.

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

Wang, Y [Yang] (2024). The role of teaching presence in students' behavioral engagement. *Distance Education*, 1–20. https://doi.org/10.1080/01587919.2024.2400256

Wang, Y [Yanqing], Cao, Y., Gong, S., Wang, Z., Li, N., & Ai, L. (2022). Interaction and learning engagement in online learning: The mediating roles of online learning self-efficacy and academic emotions. *Learning and Individual Differences*, *94*, 102128. https://doi.org/10.1016/j.lindif.2022.102128

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*. https://doi.org/10.1016/j.lcsi.2019.100318

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

Welsh, R. O., & Little, S. (2018). The school discipline dilemma: A comprehensive review of disparities and alternative approaches. *Review of Educational Research*, *88*(5), 752–794. https://doi.org/10.3102/0034654318791582

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

Woodcock, S., Hitches, E., & Jones, G. (2019). It's not you, it's me: Teachers' self-efficacy and attributional beliefs towards students with specific learning difficulties. *International Journal of Educational Research*, *97*, 107–118. https://doi.org/10.1016/j.ijer.2019.07.007

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

Yada, A., Tolvanen, A., Malinen, O.-P., Imai-Matsumura, K., Shimada, H., Koike, R., & Savolainen, H. (2019). Teachers' self-efficacy and the sources of efficacy: A cross-cultural investigation in Japan and Finland. *Teaching and Teacher Education*, *81*, 13–24. https://doi.org/10.1016/j.tate.2019.01.014

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.

Yoo, J. H. (2016). The effect of professional development on teacher efficacy and teachers' self-analysis of their efficacy change. *Journal of Teacher Education for Sustainability*, *18*(1), 84–94. https://doi.org/10.1515/jtes-2016-0007

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, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. http://ies.ed.gov/ncee/edlabs

Yough, M. (2019). Tapping the sources of self-efficacy: Promoting preservice teachers' sense of efficacy for instructing English language learners. *Teacher Educator*, *54*(3), 206–224. https://doi.org/10.1080/08878730.2018.1534031

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., de Jong, P. F., & Koomen, H. M. (2016). Teachers' self-efficacy in relation to individual students with a variety of social–emotional behaviors: A multilevel investigation. *Journal of Educational Psychology*, *108*(7), 1013–1027. https://doi.org/10.1037/edu0000106

Zee, M., & Koomen, H. M. (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. M. (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

Zee, M., Koomen, H. M., & de Jong, P. F. (2018). How different levels of conceptualization and measurement affect the relationship between teacher self-efficacy and students' academic achievement. *Contemporary Educational Psychology*, *55*, 189–200. https://doi.org/10.1016/j.cedpsych.2018.09.006

Zee, M., Koomen, H. M., Jellesma, F. C., Geerlings, J., & de Jong, P. F. (2016). Inter- and intra-individual differences in teachers' self-efficacy: A multilevel factor exploration. *Journal of School Psychology*, *55*, 39–56. https://doi.org/10.1016/j.jsp.2015.12.003

Zimmermann, F., Kaiser, J., Bernholt, A., Bauer, J., & Rösler, L. (2016). Veränderungsverläufe in Burnout-Dimensionen. Die Bedeutung personaler und sozialer Faktoren angehender Lehrkräfte im Vorbereitungsdienst [Developmental trajectories in dimensions of burnout: The role of personal and social factors among beginning teachers]. *Psychologie in Erziehung und Unterricht*, *63*(4), 258. https://doi.org/10.2378/peu2016.art21d

Zygar-Hoffmann, C., Hagemeyer, B., & Schönbrodt, F. D. (2024). The reflection of motive dispositions in everyday agentic motivational states. *Motivation Science*, *10*(4), 379–386. https://doi.org/10.1037/mot0000344

13 Appendix

Table A 1. Deviations from Pre-registration (Study 1)

Component of Pre-registration	*Where there deviations? (no, minor, major)*	*If yes, description of deviation(s)*	*Rationale for deviation*	*How might the results have changed due to the deviation?*
Research Questions	minor	Research Question 2d was not investigated.	During the screening process, the inclusion criterion regarding the scale had to be adjusted. It now says: "For comparability reasons, the used measure of teacher self-efficacy had to be Teachers' Sense of Efficacy Scale from Tschannen-Moran and Woolfolk Hoy (2001). Scales with at least one subscale comparable to the three subscales of the Teachers' Sense of Efficacy Scale were also included". Therefore, a comparison between the outlined characteristics of the measurement instruments was not possible anymore. However, the variable "used scale (TSES or other)" was used as a robustness check in the final analyses.	In general, the results became more comparable and precise.
Hypotheses	minor	Hypothesis 2d could not be tested because of the mentioned change in Research Question 2d.		
	minor	Hypothesis 2a and 2c changed from undirected to directed hypotheses.	While working on the meta-analysis and further exploring theoretical and empirical studies on self-efficacy, it became clear that there are assumptions about certain directions. Therefore, it was decided to test these assumptions explicitly.	The results became more informative for future research.

Component of Pre-registration	*Where there deviations? (no, minor, major)*	*If yes, description of deviation(s)*	*Rationale for deviation*	*How might the results have changed due to the deviation?*
Study Design	no	-/-	-/-	-/-
Search Strategy	minor	Another additional search strategy, namely the forward citation search with the tool citationchaser was applied.	During the time of applying the backward citation search, the tool CitationChaser was discovered, which can easily implement backward and forward citation searches. As guides on conducting systematic reviews, for example, the Cochrane handbook also recommends this additional search strategy; it was decided to apply forward citation search additionally (Lefebvre et al., 2021).	This additional search strategy led to a more broad data base. In the final data set, two studies (= 1,6%) came from the backward-forward-citation search.
	minor	It was refrained from explicitly asking experts in the field of TSE for unpublished articles.	During several conference presentations of Study 1 there were already established connections with experts for TSE. Together with the community calls there was no need to ask further people.	No changes are expected.
Inclusion/Exclusion Criteia	minor	Three inclusion criteria were specified in more detail:		
		Inclusion criterion 1 (numbering according to the pre-registration): The quantitative data has to be sufficient to calculate the standardized mean change.	Rationale for supplement in inclusion criterion 1: To calculate the standardized mean change, it was mandatory for primary studies to report enough quantitative data (e.g., N, SD, M).	Regarding rationale 1: No changes are expected
		Inclusion criterion 2 (numbering according to the pre-registration). The sample size	Rationale for supplement in inclusion criterion 2: A simulation study showed that	Regarding rationale 2: Only one study had to be

Component of Pre-registration	*Where there deviations? (no, minor, major)*	*If yes, description of deviation(s)*	*Rationale for deviation*	*How might the results have changed due to the deviation?*
		needs to exceed 10 persons to conduct reliable parametric analyses.	study sizes below 10 impact the estimation of Hedges *g* (Lin, 2018). Therefore, only studies with a sample size larger than 10 persons were included	excluded because of this criterion. No impact is expected.
		Inclusion criterion 4 (numbering according to the pre-registration). The used measure of teacher self-efficacy had to be the Teachers' Sense of Efficacy Scale (TSES) from Tschannen-Moran and Woolfolk Hoy (2001) or a fully available scale where at least one subscale was comparable to the three subscales of the TSES.	Rationale for supplement in inclusion criterion 4: Several meta-analyses have already shown that the chosen scale of teacher self-efficacy can have a decisive influence on the results (e.g., Chesnut & Burley, 2015; Kim & Seo, 2018; Klassen et al., 2011). Therefore, it was decided to follow common criteria and to stick to the valid and often recommended TSES and comparable subscales.	Regarding rationale 3: Less studies could be included. However, the remaining number is still enormous. It is a huge benefit that the studies fit together well now.
Stopping Rule	no	-/-	-/-	-/-
Measured/Coded Variables	minor	The variables "aim of intervention" and "program leader" could not be coded. The change in the inclusion criterion regarding the scale of self-efficacy led to the omission of the variable "subject-specificity of the measure".	"Aim of intervention" could not be coded because almost no study reported it. Many studies started with "The aim of this study is" but did not make the aims of the intervention explicit. Through the thorough coding of the sources of self-efficacy and other intervention characteristics, it is still possible to describe key characteristics of each	Through the described changes more detailed information about the studies and interventions were gathered. No changes in the results are expected.
		The following new variables were added in the coding scheme: Intervention description	intervention. Because of quite short descriptions of the interventions, the studies also did not report who exactly delivered the	

Component of Pre-registration	*Where there deviations? (no, minor, major)*	*If yes, description of deviation(s)*	*Rationale for deviation*	*How might the results have changed due to the deviation?*
		Intervention description with details Period between measurement points Only knowledge targeted Moment of reflection Additional statistics if reported	intervention. In many cases it was not clear if the teacher educator is a university teacher or an experienced school teacher, for example. Therefore,the variable "program leader" had to be dropped.	
		The following variables were specified: Mean, N and SD according to pre- and posttest of each group	The new variables capture relevant aspects of the interventions that might impact the intervention effects.	
Statistical Models	minor	Studentized deleted residuals and GOSH plots could not be analysed.	Calculating studentized deleted residuals and creating GOSH plots is at the moment not possible with multivariate random effects models in metafor. However, in addition to the promised Cook's distance, DFBETAS were calculated as further influence diagnostics.	No changes are expected since other influence diagnostic functions were used.
		The I2 statistics was not used to assess heterogeneity.	As Q-statistics can be transformed to I2 statistics, it was enough to analyze the Q-statistics.	No changes are expected because equivalent heterogeneity measures were used.
		Instead of the R package "robumeta", "clubsandwich" was used.	All analyses could be performed in metafor and clubsandwich.	No echanges are expected as metafor is a reliable R package to conduct meta-analyses.

Component of Pre-registration	*Where there deviations? (no, minor, major)*	*If yes, description of deviation(s)*	*Rationale for deviation*	*How might the results have changed due to the deviation?*
Exploratory Analyses	minor	The planned exploratory analyses on the interventions' aims and the specifications of the sources of vicarious experience and social persuasion were not possible. Two further exploratory analyses were conducted: First, an overall effect size with not only pre- and post-measures but also follow-up measures included. Second, a meta-regression with the measurement period as moderator.	The interventions were not described in enough details to code the interventions' aims or the specifications of the sources. Therefore, the analyses were not possible. More studies than expected had follow up measurement points. As these provide information on the stability of the intervention effects, these additional analyses were performed.	The analyses are marked as exploratory analyses and do not affect the rest of the results.

Note. This table transparently shows all deviations from the Pre-registration (https://osf.io/ev7tf). It follows the recommendations made by Harris, Campbell, Flake, Fried, Beck, and Kline at the APS Hackathon 2019: Best Research Practices Made Easy (for more details see: https://osf.io/c8xbp/).

Table A 2. Search Strings for Web of Science, ERIC & FIS Bildung (Study 1)

Database	*Search Syntax*
Web of Science	(TS=("program effectiveness" OR "program evaluation" OR "program implementation" OR "training*" OR "intervention*" OR "professional development" OR "mentor*" OR "teaching experience" OR "student teaching" OR "internship programs" OR "practicum*" OR "professional training" OR "teacher workshop*" OR "teacher supervision" OR "practicum supervision "OR "teaching models" OR "randomized controlled trial" OR "quasiexperimental") AND TS=("teacher education" OR "teacher education programs" OR "student teachers" OR "pre-service teachers" OR "pre-service teacher education" OR "inservice teacher education" OR "teachers" OR "beginning teachers" OR "*induction*" OR "experienced teachers" OR "novices" OR "teacher interns") AND TS=("self-efficacy" OR "self efficacy") NOT TS=("case stud*"))
ERIC – search with descriptors	DE ((DE ((DE ((DE ((DE ((DE program effectiveness) OR (DE program evaluation) OR (DE program implementation) OR (DE training) OR (DE intervention) OR (DE professional development) OR (DE mentors) OR (DE teaching experience))) OR (DE randomized controlled trial) OR (DE quasiexperimental design) OR (DE student teaching) OR (DE pretests posttests) OR (DE internship programs) OR (DE practicums))) OR (DE professional training) OR (DE teacher workshops))) OR (DE teacher supervision) OR (DE practicum supervision))) OR (DE teaching models))) AND (DE ((DE ((DE teacher education) OR (DE teacher education programs) OR (DE student teachers) OR (DE preservice teachers) OR (DE preservice teacher education) OR (DE inservice teacher education) OR (DE teachers) OR (DE beginning teachers) OR (DE beginning teacher induction))) OR (DE experienced teachers) OR (DE novices) OR (DE teacher interns))) AND (DE ((DE ((DE self efficacy)))) NOT (DE case studies)
ERIC – search within abstracts	(AB ((AB ((AB ((AB ((AB ((AB program effectiveness) OR (AB program evaluation) OR (AB program implementation) OR (AB training) OR (AB intervention) OR (AB professional development) OR (AB mentors) OR (AB teaching experience))) OR (AB randomized controlled trial) OR (AB quasiexperimental design) OR (AB student teaching) OR (AB internship programs) OR (AB practicums))) OR (AB professional training) OR (AB teacher workshops))) OR (AB teacher supervision) OR (AB practicum supervision))) OR (AB teaching models))) AND (AB ((AB ((AB teacher education) OR (AB teacher education programs) OR (AB student teachers) OR (AB preservice teachers) OR (AB preservice teacher education) OR (AB inservice teacher education) OR (AB teachers) OR (AB beginning teachers) OR (AB beginning teacher induction))) OR (AB experienced teachers) OR (AB novices) OR (AB teacher interns))) AND (AB ((AB ((AB self efficacy))))) NOT AB "case stud*"
FIS Bildung	((((Schlagwörter: LEHRER oder LEHRERIN oder LEHRERBILDUNG oder LEHRAMTSSTUDIUM oder LEHRAMTSSTUDENT oder

	LEHRAMTSSTUDENTIN oder REFERENDARIAT oder REFERENDAR oder REFERENDARIN) und (Schlagwörter: SELBSTWIRKSAMKEIT)) und (Schlagwörter: EVALUATION oder EXPERIMENT oder *TRAINING oder INTERVENTION* oder PROGRAMM oder PROGRAMMEVALUATION oder QUASI-EXPERIMENT oder WIRKUNG*)) und (Sprache: deutsch oder englisch)) und (Datenquelle: "FIS Bildung" oder "Online Contents" oder "BASE (beta)")

Table A 3. Study and Sample Characteristics (Study 1)

Author	*Year*	*Publication Type*	*Study Design*	*Study Quality*	*Number of Treatment Groups*	*Used Scale*	*Career Stage*	*School Level*	*Sample Size*	*Country*
Aasheim et al.	2020	published	quasi-experimental	3	1	TSES	IST	E	207	Norway
Akiri & Dori	2021	published	one-group study	2	1	TSES	IST	999	54	Israel
Amendum & Liebfreund	2019	published	one-group study	1	1	TSES	IST	E	22	USA
Andersson et al.	2022	published	experimental	3	1	TSES	IST	E	25	Sweden
Androzzi	2011	dissertation	one-group study	1.5	1	TSES	PST	E & S	27	USA
Ansley et al.	2021a	published	experimental	4	1	TSES	PST and IST	999	51	USA
Ansley & Wander	2021b	published	quasi-experimental	3	1	TSES	PST	E & S	28	USA
Anthony et al.	2013	published	quasi-experimental	2	2	TSES	IST	E & S	48	USA
Atay	2007	published	one-group study	1	1	TSES	PST	E & S	78	Turkey
Avalos & Bascope	2014	published	one-group study	0	1	TSES	PST	E & S	54	Chile
Aydın et al.	2012	published	one-group study	1.25	1	TSES	PST	S	26	Turkey
Azukas	2018	dissertation	one-group study	1.75	1	Teacher Question-naire (self-developed)	IST	E & S	15	USA
Berg & Smith	2018	published	one-group study	1	1	TSES	PST	E	75	New Zealand
Bhatia	2012	dissertation	one-group study	1	1	TSES	PST	999	22	USA

Author	*Year*	*Publication Type*	*Study Design*	*Study Quality*	*Number of Treatment Groups*	*Used Scale*	*Career Stage*	*School Level*	*Sample Size*	*Country*
Boomgard	2013	dissertation	one-group study	1	1	TSES	IST	E & S	15	USA
Bosch & Ellis	2021	published	quasi-experimental	1	1	TSES	PST	E & S	44	USA
Bowlin et al.	2015	other	experimental	3	2	TSES	PST	E & S	153	USA
Branch	2018	dissertation	quasi-experimental	1	1	TSES	IST	E & S	40	USA
Brown et al.	2015	published	one-group study	0	1	TSES	PST	E	66	USA
Brown	2013	dissertation	experimental	3	1	TSES	IST	E & S	36	China
Bumen	2009	published	one-group study	1.83	1	TSES	IST	E & S	38	Turkey
Cabaroglu	2014	published	one-group study	1.67	1	TSES	PST	999	60	Turkey
Chan et al.	2021	published	experimental	4	1	TSES	PST	S	65	Cambodia
Chao et al.	2017	published	one-group study	1	1	TSES	IST	E & S	345	China
Charalambous et al.	2008	published	one-group study	2	1	TSES	PST	E	89	Cyprus
Conroy et al.	2019	published	experimental	4	1	TSES	IST	E	186	USA
Dalioglu & Adiguzel	2016	published	one-group study	0.25	1	TSES	PST	E & S	68	Turkey
De Carvalho et al.	2021	published	experimental	4	1	TSES	IST	E	205	Portugal
Denzer	2018	dissertation	one-group study	1	1	TSES	IST	S	14	USA
DeSantis	2018	dissertation	experimental	2	2	TSES	PST	E & S	30	USA
Devries et al.	2021	published	one-group study	0.5	1	Norwegian Teacher Self-Efficacy Scale (Skaalvik & Skaalvik, 2007)	IST	E	132	Cote d'Ivoire

Author	*Year*	*Publication Type*	*Study Design*	*Study Quality*	*Number of Treatment Groups*	*Used Scale*	*Career Stage*	*School Level*	*Sample Size*	*Country*
Dicke et al.	2015	published	quasi-experimental	3	2	TSES	IST	E & S	56	Germany
Domitrovich et al.	2016	published	experimental	4	2	Behavior Management Self-Efficacy Scale (Main & Hammond, 2008)	IST	E	319	USA
Eder-Karavaya et al.	2021	published	quasi-experimental	3	1	Multi-dimensional scale of teacher self-efficacy (Schulte, Watermann & Bögeholz, 2011)	PST	S	75	Germany
Eginli & Solhi	2021	published	one-group study	0.25	1	TSES	PST	999	32	Turkey
Erdem	2008	published	one-group study	2	1	TSES	PST	999	43	Turkey
Fernandes et al.	2019	published	quasi-experimental	1.75	2	999	IST	E & S	59	Portugal
Flynn & Tazartes	2015	other	one-group study	1.5	1	TSES	IST	S	39	USA
Garwood & Harris	2020	published	one-group study	2	1	TSES	IST	E & S	52	USA
Gold et al.	2017	published	quasi-experimental	2.5	3	scale from BIQUA-project (Kleickmann et al., 2006)	PST	E	223	Germany
Grammatikopoulos et al.	2013	published	one-group study	2	1	TSES	IST	E & S	237	Greece
Green	2021	published	experimental	3	2	TSES	IST	E & S	206	Pakistan
Gresko	2013	dissertation	one-group study	2	1	TSES	IST	E & S	29	USA
Hannemann et al.	2019	published	experimental	4	1	TSES	PST	E & S	151	Germany

Author	*Year*	*Publication Type*	*Study Design*	*Study Quality*	*Number of Treatment Groups*	*Used Scale*	*Career Stage*	*School Level*	*Sample Size*	*Country*
Hayes et al.	2020	published	experimental	3	1	TSES	IST	E	74	England
Hoffman	2009	dissertation	one-group study	0	1	TSES	IST	E	23	USA
Holdaway	2017	dissertation	experimental	3	2	TSES	PST	E	24	USA
Hoogendijk et al.	2018	published	experimental	3.5	1	TSES	IST	E	94	Nether-lands
Hui et al.	2016	published	one-group study	1	1	TSES	IST	E	11	Canada
Jackson & Miller	2020	published	quasi-experimental	0	2	TSES	PST and IST	S	72	USA
Jamil	2012	other	quasi-experimental	3	1	TSES	IST	999	335	USA
Jang & Reeve	2021	published	experimental	4	1	TSES	IST	E & S	26	Philippines
Junqueira & Matoti	2013	published	one-group study	0	1	TSES	PST	S	109	South Africa
Karalar & Altan[a]	2018	published	one-group study	1.25	1	TSES	PST	S	24	Turkey
Karimi	2011	published	quasi-experimental	2	1	TSES	IST	S	60	Iran
Kissau & Algozzine	2015	published	quasi-experimental	2	3	TSES	PST and IST	E & S	117	USA
Klassen & Durksen	2014	published	one-group study	2	1	TSES	IST	E & S	150	Canada
Klempin et al.	2020	published	quasi-experimental	2.5	1	TSES	PST	E & S	184	Germany
Klug & Seethaler	2021	published	quasi-experimental	0.88	2	TSES	PST	E	103	Austria

Author	*Year*	*Publication Type*	*Study Design*	*Study Quality*	*Number of Treatment Groups*	*Used Scale*	*Career Stage*	*School Level*	*Sample Size*	*Country*
Knoblauch & Chase	2015	published	one-group study	1	1	TSES	PST	E & S	200	USA
Knoblauch & Hoy	2008	published	one-group study	1	1	TSES	PST	E & S	102	USA
Kollmayer et al.	2019	published	quasi-experimental	3	1	Teaching Self-Efficacy Scale (Bandura, 1997)	IST	S	114	Austria
Krauskopf & Knigge	2017	published	one-group study	0.5	1	Scale for self-efficacy towards inclusive education (Bosse & Spörer, 2014)	PST	S	157	Germany
Kunz et al.	2019	published	quasi-experimental	3	1	TSES	IST	E	118	Switz-erland
Kumschick et al.	2017	published	quasi-experimental	1.83	2	TSES	PST	999	237	Germany
Kurt	2017	published	experimental	3	1	TSES	PST	999	62	Turkey
Latouche & Gascoigne	2019	published	one-group study	1	1	TSES	IST	E	113	Australia
Lentfer & Franks	2015	published	one-group study	1	1	Attitudes Towards Science Inventory (Gogolin & Swartz, 1992)	PST	S	31	USA
Liaw	2017	published	experimental	3.83	1	TSES	PST	E	57	China (Taiwan)
Ma et al.	2022	published	one-group study	0.83	1	TSES	PST	E & S	131	Australia
Mahasneh & Alwan	2018	published	experimental	2.75	1	TSES	PST	999	79	Jordan

Author	*Year*	*Publication Type*	*Study Design*	*Study Quality*	*Number of Treatment Groups*	*Used Scale*	*Career Stage*	*School Level*	*Sample Size*	*Country*
Main & Hammond	2008	published	one-group study	1	1	Teacher Inter-personal Self-Efficacy Scale (Baker, 2005)	PST	E & S	69	999
Main et al.	2016	published	one-group study	0.75	1	Teacher Efficacy for Inclusive Practice (Sharma, Loreman, & Forlin, 2012)	IST	E	37	Seychelles
Marlow et al.	2015	published	one-group study	0	1	TSES	IST	E	37	England
Marquez et al.	2016	published	one-group study	1	1	TSES	IST	E	37	USA
Marquez et al.	2016	published	experimental	3.5	1	TSES	IST	E	83	USA
Matoti et al.	2011	published	one-group study	0	1	TSES	PST	999	51	South Africa
McCullough et al.	2021	published	experimental	4	1	TSES	IST	E	26	USA
Michos et al.	2022	published	experimental	2.5	2	TSES	PST	E	81	Switz-erland
O`Neill	2015	published	one-group study	2	1	TSES	PST	E	20	Australia
Oh	2010	dissertation	one-group study	1	1	TSES	PST	E	59	USA
Ortactepe & Akyel	2015	published	one-group study	1.88	1	TSES	IST	S	50	Turkey
Ozdemir & Aydin	2021	published	one-group study	2	1	TSES	IST	S	45	Turkey
Peebles & Mendaglio[a]	2014	published	one-group study	1	1	Teacher Efficacy for Inclusive Practices (Sharma, Loreman, & Forlin, 2012)	PST	E & S	141	Canada

Author	*Year*	*Publication Type*	*Study Design*	*Study Quality*	*Number of Treatment Groups*	*Used Scale*	*Career Stage*	*School Level*	*Sample Size*	*Country*
Perkins Coppola	2019	published	one-group study	2	1	Teaching Engineering Self-Efficacy Scale (Yoon et al., 2014)	PST	E	124	USA
Putman	2012	published	quasi-experimental	0.88	2	TSES	PST	E	43	USA
Reyes et al.	2017	published	one-group study	0	1	TSES	IST	999	47	USA
Ribeiro	2009	dissertation	quasi-experimental	1	2	TSES	PST and IST	E & S	23	USA
Roberts et al.	2008	published	quasi-experimental	0	2	TSES	PST	999	138	USA
Ross & Bruce	2007	published	experimental	4	1	TSES	IST	E	106	Canada
Samuelsson et al.	2021	published	experimental	2	3	TSES	PST	E	72	Sweden
Schipper et al.	2020	published	quasi-experimental	1.67	1	TSES	IST	S	60	Netherlands
Sharma & Nuttal	2016	published	one-group study	1	1	Teacher Efficacy for Inclusive Practice (Sharma, Loreman, & Forlin, 2012)	PST	E & S	22	Australia
Sheehan & Moore	2019	published	one-group study	2	1	TSES	PST	S	20	USA
Siers et al.	2017	other	one-group study	0	1	TSES	PST	999	126	USA
Stanton et al.	2018	published	quasi-experimental	1	2	TSES	IST	E & S	70	USA
Stevens et al.	2013	published	one-group study	1	1	TSES	IST	S	58	USA

Author	*Year*	*Publication Type*	*Study Design*	*Study Quality*	*Number of Treatment Groups*	*Used Scale*	*Career Stage*	*School Level*	*Sample Size*	*Country*
Stewart	2012	published	one-group study	1	1	TSES	PST	E & S	160	USA
Stewart et al.	2011	published	quasi-experimental	1	1	TSES	PST	999	292	USA
Stripling & Roberts	2013	published	one-group study	0	1	TSES	PST	S	22	USA
Sumners	2012	dissertation	one-group study	0	1	TSES	PST	E	29	USA
Thiel et al.	2020	published	experimental	3	2	TSES	PST	E & S	114	Germany
Thomas & Mucherah	2016	published	quasi-experimental	1.75	2	TSES	PST	E	96	USA
Tschannen-Moran & McMaster	2009	published	quasi-experimental	2	4	TSES	IST	E	93	999
Van der Scheer & Visscher	2016	published	experimental	3.75	1	TSES	IST	E	62	Nether-lands
VanParys Couet	2014	dissertation	one-group study	0	1	TSES	IST	S	42	USA
Vansteelandt et al.	2020	published	experimental	3.67	2	TSES	IST	E	28	999
Vernon-Feagans et al.	2015	published	experimental	3	2	TSES	IST	E	58	USA
Wagler & Moseley	2005	published	one-group study	1.88	1	TSES	PST	S	59	USA
Weber et al.	2019	published	quasi-experimental	2	3	TSES	PST	999	121	Germany
Wess et al.	2020	published	quasi-experimental	1.75	1	self-efficacy towards planning, conducting and reflecting lessons (self-developed)	PST	E & S	47	Germany
Whitley et al.[a]	2019	published	one-group study	1	1	TSES	PST	999	22	USA

Author	*Year*	*Publication Type*	*Study Design*	*Study Quality*	*Number of Treatment Groups*	*Used Scale*	*Career Stage*	*School Level*	*Sample Size*	*Country*
Yazici & Yildirim	2017	published	one-group study	1	1	TSES	PST	S	178	Turkey
Yilmaz & Koca[a]	2017	published	one-group study	1	1	TSES	PST	999	55	Turkey
Yoo	2016	published	one-group study	0	1	TSES	IST	E & S	148	USA
Yough	2019	published	experimental	3.83	1	TSES	PST	E & S	162	USA
Yüksel H.G.	2014	published	one-group study	2	1	TSES	PST	E & S	40	Turkey
Yüksel I.	2014	published	one-group study	0.88	1	TSES	PST	999	85	Turkey
Zeuch et al.	2021	published	quasi-experimental	2.5	1	self-efficacy towards heterogeneity (Junker et al., 2020)	PST	E & S	1322	Germany

Note. Study quality is averaged per study; TSES = Teachers' Sense of Efficacy Scale from Tschannen-Moran & Woolfolk Hoy, 2001; PST = pre-service teachers; IST = in-service teachers; E = elementary schools; S = secondary schools.

a. This study is not included in the final statistical analysis because it was identified as influential study.

Table A 4. Intervention Characteristics (Study 1)

Author	*Year*	*Duration in Hours*	*Period Between Measurement Points in Weeks*	*Only Know-Ledge*	*Moment Of Reflection*	*Mastery Experiences*	*Vicarious Experiences*	*Social Persuasion*	*Physiolo-gical and Emotional Reactions*
Aasheim et al.	2020	42	36	no	yes	yes	yes	yes	no
Akiri & Dori	2021	60	40	no	no	no	no	yes	no
Amendum & Liebfreund	2019	n.a.	40	no	yes	yes	yes	yes	no
Andersson et al.	2022	17	11	no	yes	yes	yes	yes	no
Androzzi	2011	23	9	no	yes	yes	yes	yes	no
Ansley et al.	2021a	4	4	no	yes	no	yes	no	yes
Ansley & Wander	2021b	n.a.	8	no	no	yes	no	no	no
Anthony et al. (subgroup 1: attended 6 or more)	2013	n.a.	40	no	no	no	yes	yes	no
Anthony et al. (subgroup 2: attended 1 to 5)	2013	n.a.	40	no	no	no	yes	yes	no
Atay	2007	n.a.	40	no	yes	yes	no	no	no
Avalos & Bascope	2014	n.a.	16	no	no	yes	yes	yes	no
Aydın et al.	2012	145	13	no	no	yes	yes	yes	no
Azukas	2018	n.a.	36	no	yes	yes	yes	yes	no
Berg & Smith	2018	n.a.	n.a.	no	no	yes	no	no	no
Bhatia	2012	n.a.	n.a.	n.a.	n.a.	No sources codable			
Boomgard	2013	n.a.	16	no	yes	no	no	yes	no
Bosch & Ellis	2021	n.a.	n.a.	no	no	yes	yes	yes	no
Bowlin et al. (subgroup 1: video modeling condition)	2015	n.a.	18	yes	no	No sources codable			
Bowlin et al. (subgroup 2: live modeling condition)	2015	n.a.	18	no	no	no	yes	no	no

Author	*Year*	*Duration in Hours*	*Period Between Measurement Points in Weeks*	*Only Know-Ledge*	*Moment Of Reflection*	*Mastery Experiences*	*Vicarious Experiences*	*Social Persuasion*	*Physiolo-gical and Emotional Reactions*
Branch	2018	3	6	no	no	no	no	yes	no
Brown et al.	2015	n.a.	20	no	no	yes	yes	yes	no
Brown	2013	n.a.	n.a.	n.a.	n.a.	No sources codable			
Bumen	2009	n.a.	40	no	no	yes	no	yes	no
Cabaroglu	2014	n.a.	14	no	yes	yes	yes	yes	no
Chan et al.	2021	n.a.	16	no	yes	yes	no	yes	no
Chao et al.	2017	30	1	yes	no	No sources codable			
Charalambous et al.	2008	90	12	no	yes	yes	yes	yes	no
Conroy et al.	2019	7	14	no	yes	yes	yes	yes	no
Dalioglu & Adiguzel	2016	n.a.	n.a.	no	no	yes	no	no	no
De Carvalho et al.	2021	30	10	no	yes	yes	no	yes	yes
Denzer	2018	5	12	no	yes	no	no	yes	no
DeSantis (subgroup 1: group with coaching)	2018	1.75	15	no	no	yes	no	yes	no
DeSantis (subgroup 2: group without coaching)	2018	0.5	15	no	no	yes	no	no	no
Devries et al.	2021	16	0	no	yes	yes	no	yes	no
Dicke et al. (subgroup 1: treatment classroom management training)	2015	19	0	no	no	yes	yes	yes	no
Dicke et al. (subgroup 2: treatment stress management training)	2015	19	0	no	yes	yes	yes	yes	yes
Domitrovich et al. (subgroup 1: treatment PATHS to PAX)	2016	n.a.	31	no	no	yes	yes	yes	no
Domitrovich et al. (subgroup 2: treatment PAX GBG)	2016	n.a.	31	no	no	yes	yes	yes	no

Author	*Year*	*Duration in Hours*	*Period Between Measurement Points in Weeks*	*Only Know-Ledge*	*Moment Of Reflection*	*Mastery Experiences*	*Vicarious Experiences*	*Social Persuasion*	*Physiolo-gical and Emotional Reactions*
Eder-Karavaya et al.	2021	n.a.	4	no	yes	no	yes	yes	yes
Eginli & Solhi	2021	n.a.	12	no	no	yes	no	no	no
Erdem	2008	84	20	no	yes	yes	no	yes	no
Fernandes et al. (subgroup 1)	2019	18	9	no	yes	yes	no	yes	no
Fernandes et al. (subgroup 2)	2019	18	9	no	yes	yes	no	yes	no
Flynn & Tazartes	2015	64	40	no	no	yes	no	yes	no
Garwood & Harris	2020	n.a.	n.a.	yes	no	No sources codable			
Gold et al. (subgroup 1)	2017	n.a.	20	no	yes	no	yes	yes	no
Gold et al. (subgroup 2)	2017	n.a.	20	no	yes	yes	no	yes	no
Gold et al. (subgroup 3)	2017	n.a.	20	no	yes	yes	yes	yes	no
Grammatikopoulos et al.	2013	60	8	yes	no	No sources codable			
Green (subgroup 1)	2021	30	5	no	yes	yes	no	yes	no
Green (subgroup 2)	2021	30	5	yes	no	No sources codable			
Gresko	2013	19	15	no	yes	yes	yes	yes	no
Hannemann et al.	2019	15	4	no	yes	yes	yes	yes	no
Hayes et al.	2020	48	24	no	yes	no	yes	yes	no
Hoffman	2009	n.a.	n.a.	no	no	yes	no	yes	no
Holdaway (subgroup 1: internship as usual)	2017	n.a.	n.a.	no	no	yes	yes	yes	no
Holdaway	2017	n.a.	n.a.	no	yes	yes	yes	yes	no

Author	*Year*	*Duration in Hours*	*Period Between Measurement Points in Weeks*	*Only Know-Ledge*	*Moment Of Reflection*	*Mastery Experiences*	*Vicarious Experiences*	*Social Persuasion*	*Physiolo-gical and Emotional Reactions*
(subgroup 2: internship plus intensive coaching)									
Hoogendijk et al.	2018	9	n.a.	no	yes	yes	yes	yes	no
Hui et al.	2016	17	12	no	yes	yes	no	yes	yes
Jackson & Miller (subgroup 1: traditional teacher education)	2020	n.a.	20	no	no	no	yes	no	no
Jackson & Miller (subgroup 2: alternative teacher education)	2020	n.a.	20	yes	no	No sources codable			
Jamil	2012	n.a.	40	no	no	yes	yes	yes	yes
Jang & Reeve	2021	n.a.	n.a.	no	yes	no	yes	no	no
Junqueira & Matoti	2013	n.a.	24	n.a.	n.a.	No sources codable			
Karalar & Altan[a]	2018	56	7	no	yes	yes	no	yes	no
Karimi	2011	24	n.a.	no	no	yes	yes	yes	no
Kissau & Algozzine (subgroup 1: face-to-face)	2015	n.a.	15	no	yes	yes	yes	yes	no
Kissau & Algozzine (subgroup 2: online)	2015	n.a.	15	no	yes	yes	yes	yes	no
Kissau & Algozzine (subgroup 3: hybrid)	2015	n.a.	15	no	yes	yes	yes	yes	no
Klassen & Durksen	2014	n.a.	7	no	no	yes	yes	yes	no
Klempin	2020	n.a.	20	no	yes	yes	yes	yes	no
Klug & Seethaler (subgroup 1: face-to-face)	2021	4	20	no	no	yes	yes	no	no
Klug & Seethaler (subgroup 2: online)	2021	4	20	no	no	no	yes	no	no
Knoblauch & Chase	2015	n.a.	16	no	no	yes	no	no	no
Knoblauch & Hoy	2008	n.a.	16	no	no	yes	yes	no	no

Author	*Year*	*Duration in Hours*	*Period Between Measurement Points in Weeks*	*Only Know-Ledge*	*Moment Of Reflection*	*Mastery Experiences*	*Vicarious Experiences*	*Social Persuasion*	*Physiolo-gical and Emotional Reactions*
Kollmayer et al.	2019	75	20	no	yes	yes	no	yes	no
Krauskopf & Knigge	2017	n.a.	20	n.a.	n.a.	No sources codable			
Kumschick et al. (subgroup 1)	2017	6	4	no	no	no	yes	no	no
Kumschick et al. (subgroup 2)	2017	6	4	no	no	no	yes	no	no
Kunz et al.	2019	14	4	no	yes	yes	yes	yes	no
Kurt	2017	21	14	no	no	yes	yes	yes	no
Latouche & Gascoigne	2019	2	0	no	no	no	yes	yes	no
Lentfer & Franks	2015	30	n.a.	no	yes	yes	no	no	no
Liaw	2017	n.a.	20	no	yes	yes	yes	yes	no
Ma et al.	2022	n.a.	4	no	no	yes	no	no	no
Mahasneh & Alwan	2018	n.a.	n.a.	yes	yes	No sources codable			
Main & Hammond	2008	n.a.	n.a.	n.a.	n.a.	No sources codable			
Main et al.	2016	n.a.	20	no	yes	yes	no	no	no
Marlow et al.	2015	48	24	no	yes	yes	yes	yes	no
Marquez et al. (study 1)	2016	n.a.	2	no	no	no	yes	no	no
Marquez et al. (study 3)	2016	n.a.	15	no	no	yes	yes	no	no
Matoti et al.	2011	n.a.	24	n.a.	n.a.	No sources codable			
McCullough et al.	2021	n.a.	14	no	yes	yes	yes	yes	no
Michos et al. (subgroup 1: with mentoring)	2022	n.a.	4	no	yes	yes	no	yes	no
Michos et al. (subgroup 2: without mentoring)	2022	n.a.	4	no	yes	yes	no	yes	no
O`Neill	2015	26	13	no	yes	yes	yes	yes	no

Author	*Year*	*Duration in Hours*	*Period Between Measurement Points in Weeks*	*Only Know-Ledge*	*Moment Of Reflection*	*Mastery Experiences*	*Vicarious Experiences*	*Social Persuasion*	*Physiolo-gical and Emotional Reactions*
Oh	2010	n.a.	16	no	no	yes	yes	yes	no
Ortactepe & Akyel	2015	n.a.	32	no	yes	no	no	yes	no
Ozdemir & Aydin	2021	8	0	no	yes	yes	no	yes	no
Peebles & Mendaglio[a]	2014	n.a.	13	no	no	yes	no	no	no
Perkins Coppola	2019	n.a.	14	no	yes	yes	no	yes	no
Putman (subgroup 1: traditional)	2012	n.a.	40	no	yes	yes	yes	yes	no
Putman (subgroup 2: blocked)	2012	n.a.	20	no	yes	yes	yes	yes	no
Reyes et al.	2017	n.a.	n.a.	no	no	yes	no	no	no
Ribeiro (subgroup 1)	2009	n.a.	11	no	no	No sources codable			
Ribeiro (subgroup 2)	2009	n.a.	11	no	no				
Roberts et al. (subgroup 1: pairs)	2008	n.a.	11	no	no	yes	no	yes	no
Roberts et al. (subgroup 2: alone)	2008	n.a.	11	no	no	yes	no	no	no
Ross & Bruce	2007	14	16	no	no	yes	yes	yes	no
Samuelsson et al. (subgroup 1: fieldwork)	2021	12	3	no	no	yes	no	yes	no
Samuelsson et al. (subgroup 2: simulation)	2021	3	n.a.	no	no	yes	yes	yes	no
Samuelsson et al. (subgroup 3: peer teaching)	2021	3	n.a.	no	no	yes	yes	yes	no
Schipper et al.	2020	n.a.	40	no	no	yes	yes	yes	no
Sharma & Nuttal	2016	18	9	yes	no	No sources codable			

Author	*Year*	*Duration in Hours*	*Period Between Measurement Points in Weeks*	*Only Know-Ledge*	*Moment Of Reflection*	*Mastery Experiences*	*Vicarious Experiences*	*Social Persuasion*	*Physiolo-gical and Emotional Reactions*
Sheehan & Moore	2019	n.a.	20	no	yes	yes	no	no	no
Siers et al.	2017	n.a.	8	n.a.	n.a.	No sources codable			
Stanton et al. (subgroup 1: cadre teachers)	2018	n.a.	40	no	no	yes	no	yes	no
Stanton et al. (subgroup 2: non-cadre teachers)	2018	n.a.	40	no	no	No sources codable			
Stevens et al.	2013	n.a.	6	yes	no	No sources codable			
Stewart	2012	n.a.	n.a.	no	yes	No sources codable			
Stewart et al.	2011	15	20	no	yes	yes	no	yes	no
Stripling & Roberts	2013	n.a.	n.a.	no	no	yes	yes	no	no
Sumners	2012	n.a.	14	no	no	no	no	yes	no
Thiel et al. (subgroup 1: functional video)	2020	6	6	no	no	no	yes	yes	no
Thiel et al. (subgroup 2: dysfunctional video)	2020	6	6	no	no	no	yes	yes	no
Thomas & Mucherah (subgroup 1: immersive practicum)	2016	90	12	no	yes	yes	yes	yes	no
Thomas & Mucherah (subgroup 2: normal practicum)	2016	33	9	no	no	yes	yes	yes	no
Tschannen-Moran & McMaster (subgroup 1)	2009	3	4	no	no	yes	no	no	no
Tschannen-Moran & McMaster (subgroup 2)	2009	3	4	no	no	no	yes	no	no
Tschannen-Moran & McMaster (subgroup 3)	2009	5	4	no	no	yes	yes	yes	no
Tschannen-Moran & McMaster (subgroup 4)	2009	6	4	no	no	yes	yes	yes	no
Van der Scheer & Visscher	2016	39	40	no	yes	yes	yes	yes	no

Author	*Year*	*Duration in Hours*	*Period Between Measurement Points in Weeks*	*Only Know-Ledge*	*Moment Of Reflection*	*Mastery Experiences*	*Vicarious Experiences*	*Social Persuasion*	*Physiolo-gical and Emotional Reactions*
VanParys Couet	2014	n.a.	n.a.	n.a.	n.a.	No sources codable			
Vansteelandt et al. (subgroup 1: group professional development)	2020	24	40	no	yes	yes	yes	yes	no
Vansteelandt et al. (subgroup 2: individual professional development)	2020	24	40	no	yes	yes	yes	yes	no
Vernon-Feagans et al. (subgroup 1: face-to-face)	2015	n.a.	40	no	no	yes	no	yes	no
Vernon-Feagans et al. (subgroup 2: remote)	2015	n.a.	40	no	no	yes	no	yes	no
Wagler & Moseley	2005	n.a.	40	no	yes	yes	no	yes	no
Weber et al. (subgroup 1: live coaching)	2019	n.a.	4	no	yes	yes	yes	yes	no
Weber et al. (subgroup 2: virtual coaching with peers)	2019	n.a.	4	no	yes	yes	yes	yes	no
Weber et al. (subgroup 3: virtual coaching with peers + experts)	2019	n.a.	4	no	yes	yes	yes	yes	no
Wess et al.	2020	n.a.	n.a.	no	yes	yes	no	yes	no
Whitley et al.[a]	2019	n.a.	20	no	no	yes	no	no	no
Yazici & Yildirim	2017	n.a.	28	n.a.	n.a.	No sources codable			
Yilmaz & Koca [a]	2017	84	14	n.a.	n.a.	No sources codable			
Yoo	2016	n.a.	5	no	no	no	yes	yes	yes
Yough	2019	2	0	no	yes	yes	yes	yes	yes
Yüksel H.G.	2014	210	28	no	yes	yes	yes	yes	no
Yüksel I.	2014	24	12	n.a.	n.a.	No sources codable			
Zeuch et al.	2021	n.a.	20	no	yes	no	yes	yes	no

Note. One row represents one intervention group

a. This study is not included in the final statistical analysis because it was identified as influential study.

Table A 5. Results of Exploratory Meta-Regression on Measurement Point (Study 1)

Moderator variables	*n*	*k*	*g*	*SE*	*95% CI*	*df*	*Q*	*p*
Measurement period	115[a]	352					3617.9751***	.03
Pre-post	115	318	0.46	0.04	[0.39, 0.53]	112.11		<0.0001
Post-follow up	11	17	-0.33	0.09	[-0.57, -0.10]	4.89		.01
Pre-follow up	11	17	0.06	0.08	[-0.16, 0.28]	4.79		.49

Note. n = number of studies; *k* = number of effect sizes. 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 A 6. Deviations from Pre-registration (Study 2)

Component of Preregistration	*Were there deviations? (no, minor, major)*	*If yes, description of deviation(s)*	*Rationale for deviation*	*How might the results have been changed due to the deviation?*
Hypotheses	minor	No hypotheses were investigated.	During the preparation of the theoretical background, it became clear that there aren't strong indices supporting hypotheses.	The study was conducted in an exploratory manner.
Data Collection Procedures	no	-/-	-/-	-/-
Sample Size	major	The sample size of 26 teachers is far below the planned sample size of 60 teachers.	There were no more capacities to recruit more teachers.	The study is underpowered, and the standard errors might have been underestimated, which led to an increased type I-error rate.
Measured Variables	no	-/-	-/-	-/-
Statistical Models	minor	Classroom characteristics were added all at once and not sequentially into the model.	There were no hypotheses regarding the order of the predictors.	No changes expected. The model considering all interrelations might be even provide better results.

Note. This table transparently shows all deviations from the pre-registration (https://osf.io/ev7tf). It follows the recommendations made by Harris, Campbell, Flake, Fried, Beck, and Kline at the APS Hackathon 2019: Best Research Practices Made Easy (for more details see: https://osf.io/c8xbp/).

Table A 7. Distribution of Subjects Across Teachers and Classes (Study 2)

Subject	*Number of Teachers*	*Number of Classes*
Mathematics	6	16
Economics	5	14
Dutch	3	9
English	3	9
Biology	2	6
Chemistry	2	5
Natural Sciences	3	5
Economy & Physical Education	1	3
Physical Education	1	3
Political/Social Sciences	1	3
Religion/Ethics	1	3

Note. The subjects are ordered by frequency of classes. The number of teachers exceeds $n = 26$ because two teachers taught two subjects. One teacher reported only on one class, therefore, the number of classes in total is $n = 76$. The teachers taught 12 different subjects; mathematics ($n = 6$ teachers) and economics ($n = 5$ teachers) were the most frequently taught subjects.

Table A 8. Teachers' Sense of Efficacy Scale (Study 2)

Number of item	*NL*	EN
	Hieronder staan vragen over zaken die moeilijkheden voor docenten kunnen veroorzaken in de klas. Hoe ervaart u dit in deze klas? *Wilt u hieronder aangeven in hoeverre u het met elke stelling eens bent door het betreffende hokje aan te klikken?*	Below are questions about issues that may cause difficulties for teachers in the classroom. How do you experience this in this class? Would you please indicate below the extent to which you agree with each statement by checking the appropriate box?
1	*Hoe goed lukt het u om een andere uitleg of voorbeeld te geven wanneer leerlingen iets niet begrijpen?*	How well do you manage to give a different explanation or example when students do not understand?
2	*Hoe goed lukt het u om storend gedrag in de klas te beperken?*	How well do you manage to limit disruptive behavior in the classroom?
3	*Hoe goed lukt het u om leerlingen het vertrouwen te geven dat zij hun schoolwerk goed kunnen doen?*	How well do you succeed in giving students confidence that they can do their schoolwork well?
4	*In hoeverre kunt u vaststellen wat leerlingen hebben begrepen van uw uitleg?*	To what extent can you determine what students have understood from your explanations?
5	*Hoe goed lukt het u om leerlingen te stimuleren door te denken op de geboden stof?*	How well do you succeed in encouraging students to think through on the material provided?
6	*Hoe goed lukt het u om leerlingen de klassenregels te laten volgen?*	How well do you succeed in getting students to follow class rules?
7	*Hoe goed lukt het u om uw lessen af te stemmen op het juiste niveau van individuele leerlingen?*	How well do you succeed in tailoring your lessons to the appropriate level of individual students?
8	*In hoeverre lukt het u om leerlingen te helpen leren belangrijk te gaan vinden?*	To what extent do you succeed in helping students begin to find learning important?
9	*Hoe goed kunt u voorkomen dat een paar leerlingen de sfeer in de klas negatief beinvloeden?*	How well can you prevent a few students from negatively affecting the classroom atmosphere?
10	*Hoe goed kunt u passende uitdagingen bieden aan leerlingen die goed kunnen leren?*	How well can you provide appropriate challenges for students who are good learners?
11	*In hoeverre lukt het u om leerlingen zelf nieuwe dingen te laten ontdekken?*	To what extent do you succeed in allowing students to discover new things for themselves?
12	*In hoeverre kunt u aan leerlingen duidelijk maken wat voor gedrag u van hen verwacht?*	To what extent can you make clear to students what kind of behavior you expect from them?

Reference: Tschannen-Moran & Woolfolk Hoy, 2001; Zee, Koomen, et al., 2016

Answer categories: 7-point Likert scale with 1 “Helemaal niet (Not at all)”, 2 “Niet (No)”, 3 “Meer niet dan wel (More not than well)”, 4 “Gemiddeld (Average)”, 5 “Meer wel dan niet (More than not)”, 6 “Wel (Well)” and 7 “Helemaal wel (Completely well)”

Items 1, 4, 7 and 10 belong to the subscale instructional strategies self-efficacy.

Items 2, 6, 9 and 12 belong to the subscale classroom management self-efficacy.

Items 3, 5, 8, and 11 belong to the subscale student engagement self-efficacy.

To evaluate the factorial structure, a confirmatory factor analysis using the COMPLEX function of Mplus 8.3 (Muthén & Muthén, 1998-2017) that accounts for the dependence of clustered data was conducted. This analysis uses robust maximum likelihood estimation (MLR), producing robust standard errors and a mean-adjusted chi-square test statistic. The three-factor-solution barely yielded an acceptable fit to the data $\chi^2(51) = 79.231$, $p < 0.01$, RMSEA = 0.09, 90% CI [0.05, 0.12], CFI = 0.919, SRMR = 0.064 which might be the result of having only 26 clusters. Factor loadings ranged between 0.60 and 0.86 (see https://osf.io/zxcvt).

Table A 9. Scale for Teacher Perceived Classroom Misbehavior (Study 2)

Number of item	*NL*	EN
1	*Hoe vaak ervaart u negatieve interacties met leerlingen in deze klas?**	How often do you experience negative interactions with students in this class?*
2	*Hoe vaak heeft u te maken met leerlingen met gedragsproblemen in deze klas?**	How often do you deal with students with behavior problems in this class?*
3	*Hoe belastend ervaart u het omgaan met probleemgedrag in deze klas?***	How stressful do you experience dealing with problem behavior in this class?**

Reference: Tsouloupas et al., 2010; Zee, de Jong, & Koomen, 2016

*Answer categories: Five-point Likert scale with 1 "Nooit (Never)", 2 "Soms (Sometimes)", 3 "Regelmatig (Regularly)", 4 "Vaak (Often)" and 5 "(Bijna) altijd (Almost) Always"

**Answer categories: Five-point Likert scale with 1 "Helemaal niet belastend (Not stressful at all)", 2 "Nauwelijks belastend (Barely stressful)", 3 "In redelijke mate belastend (Fairly stressful)", 4 "In hoge mate belastend (Highly stressful)" and 5 "In zeer hoge mate belastend (Very highly stressful)"

A confirmatory factor analysis accounting for the clustered data produced an excellent fit to the data ($\chi^2(0) = 0.00$, $p < 0.01$, RMSEA = 0.00, 90% CI [0.00, 0.00], CFI = 1.00, SRMR = 0.00). Factor loadings ranged from 0.81 to 0.91 (see https://osf.io/zxcvt).

Table A 10. Engagement Versus Disaffection With Learning Scale (Study 2)

Number of item	*NL*	EN
	Vragen over je klas: Deze vragen gaan over jouw klas bij de leraar voor wie je de vragen hebt ingevuld en hoe leuk jij de lessen van deze leraar vindt. Het is de bedoeling dat je een vinkje zet in het hokje van het antwoord dat jij het best vindt passen. Bij sommige vragen kan et lastig zijn om een antwoord te geven. Maar het is de bedoeling dat je altijd een antwoord geeft. Kies dan wat jij het best vindt passen.	Questions about your class: These questions are about your class with the teacher for whom you completed the questions and how much you like this teacher's classes. You should check the box of the answer you think fits best. Some questions can be difficult to answer. But you should always give an answer. Then choose what you think fits best.
1	*Ik doe mijn best om het goed te doen tijdens dit vak.*	I try my best to do well during this subject.
2	*Tijdens de lessen van deze leraar denk ik aan andere dingen. (R)*	During this teacher's lessons I think about other things. (R)
3	*Tijdens de lessen van deze leraar luister ik heel goed.*	During this teacher's lessons, I listen very carefully.
4	*In de lessen van deze leraar doe ik alleen maar alsof ik aan het werk ben. (R)*	During this teacher's lessons I only pretend to be working. (R)
5	*Ik let op in de lessen van deze leraar.*	I pay attention in this teacher's lessons.
6	*In de lessen van deze leraar werk ik zo hard als ik kan.*	In this teacher's lessons, I work as hard as I can.
7	*In de lessen van deze leraar voel ik me fijn.*	In this teacher's lessons I feel comfortable.
8	*Als we in de lessen van deze leraar aan het werk zijn, ben ik geinteresseerd.*	When we are working in this teacher's lessons, I am interested.
9	*Wanneer we in de lessen van deze leraar aan iets nieuws beginnen, dan voel ik me zenuwachtig. (R)**	When we are working on something new in this teacher's lessons, I feel nervous. (R)*
10	*Als we in de lessen van deze leraar aan iets werken, verveel ik me. (R)*	When we work on something in this teacher's classes, I get bored. (R)
11	*Ik vind het leuk om nieuwe dingen te leren in de klas.*	I like learning new things in class.
12	*Ik vind de lessen van deze leraar best wel vervelend. (R)*	I find this teacher's classes quite boring. (R)

Reference: Skinner et al., 2008; Zee & Koomen, 2020

Answer categories: Five-point Likert scale with 1 “Nee, dat is niet zo (No, that is not true)”, 2 “Dat is meestal niet zo (It is not usually so)”, 3 “Soms (Sometimes)”, 4 “Dat is meestal wel zo (It is usually so)” and 5 “Ja, dat is zeker zo (Yes, this is definitely so)”

Items 1 to 6 belong to subscale behavioral engagement;

Items 7 to 12 belong to subscale emotional engagement

Reverse coded items are indicated with (R)

* this item had to be removed in the analyses due to low factor loading

The assumed two-factor structure was evaluated with a confirmatory factor analysis considering the dependent structure (students clustered in classes) by using the TYPE = COMPLEX function in M*plus* 8.3 (see https://osf.io/zxcvt/). After removing one emotional engagement item with a low factor loading ($\lambda = 0.18$), the two-factor model showed sufficient fit ($\chi^2(43) = 434.70$, $p < 0.01$, RMSEA = 0.08 [0.08; 0.09], CFI = 0.90, TLI = 0.88, SRMR = 0.06).

Reference List of Studies Included in the Meta-Analysis (Study 1)

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

Akiri, E., & Dori, Y. J. (2021). Professional growth of novice and experienced STEM teachers. *Journal of Science Education and Technology*, *31*, 129–142. https://doi.org/10.1007/s10956-021-09936-x

Amendum, S. J., & Liebfreund, M. D. (2019). Situated learning, professional development, and early reading intervention: A mixed methods study. *The Journal of Educational Research*, *112*(3), 342–356. https://doi.org/10.1080/00220671.2018.1523782

Andersson, K., Sandgren, O., Rosqvist, I., Ahlander, V. L., Hansson, K., & Sahlen, B. (2022). Enhancing teachers' classroom communication skills - Measuring the effect of a continued professional development programme for mainstream school teachers. *Child Language Teaching and Therapy*, *38*(2), 166–179. https://doi.org/10.1177/02656590211070997

Androzzi, J. (2011). *The effects of early clinical teaching experiences on pre-service teachers' self-efficacy* [Doctoral dissertation, The University of North Carolina]. NC DOCKS (North Carolina Digital Online Collection of Knowledge and Scholarship). https://libres.uncg.edu/ir/uncg/f/Androzzi_uncg_0154D_10775.pdf

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

Ansley, B. M., & Wander, M. A. (2021). Self-care options for resilient educators (SCORE) teaches aspiring teachers how to manage stress in light of COVID-related disruptions. *OBM Integrative and Complementary Medicine*, *6*(4). https://doi.org/10.21926/obm.icm.2104039

Anthony, A., Gimbert, B., & Fultz, D. (2013). The effect of e-coaching attendance on alternatively certified teachers' sense of self-efficacy. *Journal of Technology and Teacher Education*, *21*(3), 277–299.

Atay, D. (2007). Beginning teacher efficacy and the practicum in an EFL context. *Teacher Development*, *11*(2), 203–219. https://doi.org/10.1080/13664530701414720

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

Aydın, S., Demırdöğen, B., & Tarkin, A. (2012). Are they efficacious? Exploring pre-service teachers' teaching efficacy beliefs during the practicum. *The Asia-Pacific Education Researcher*, *21*(1), 203–213.

Azukas, M. E. (2018). *A personal professional learning cohort cultivating a community of practice to lead school district change* (Publication No. 10977233) [Doctoral dissertation, Arizona State University]. ProQuest Dissertations and Theses Global.

Berg, D. A. G., & Smith, L. F. (2018). The effect of school-based experience on preservice teachers' self-efficacy beliefs. *Issues in Educational Research*, *28*(3), 530–544.

Bhatia, P. (2012). *Mind over matter: Contributing factors to self-efficacy in Montessori teachers* (Publication No. 3547952) [Doctoral dissertation, University of Colorado Denver]. ProQuest Dissertations and Theses Global.

Boomgard, M. C. (2013). *Changes in perceived teacher self-efficacy and burnout as a result of facilitated discussion and self-reflection in an online course to prepare teachers to work with students with autism* [Doctoral dissertation, University of San Francisco]. The USF Scholarship Repository. https://repository.usfca.edu/diss/78

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.

Branch, Eugene, Jr. (2018). *An evaluation of the proficiency of strategic e-mentoring in improving the personal self-efficacy of music teachers in an urban school district* (Publication No. 10751641) [Doctoral dissertation, Northcentral University]. ProQuest Dissertations and Theses Global.

Brown, A. L., Lee, J., & Collins, D. (2015). Does student teaching matter? Investigating pre-service teachers' sense of efficacy and preparedness. *Teaching Education*, *26*(1), 77–93. https://doi.org/10.1080/10476210.2014.957666

Brown, J. K. W. (2013). *Relationship between teacher efficacy and professional development in international schools* (Publication No. 3567834) [Doctoral dissertation, Capella University]. ProQuest Dissertations and Theses Global.

Bumen, N. T. (2009). Possible effects of professional development on Turkish teachers' self-efficacy and classroom practice. *Professional Development in Education*, *35*(2), 261–278. https://doi.org/10.1080/13674580802568385

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

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

Charalambous, C. Y., Philippou, G. N., & Kyriakides, L. (2008). Tracing the development of preservice teachers' efficacy beliefs in teaching mathematics during fieldwork. *Educational Studies in Mathematics*, *67*(2), 125–142. https://doi.org/10.1007/s10649-007-9084-2

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

Denzer, A. Q. (2018). *Evaluation of a psychoeducationally-based program addressing bullying among students with disabilities through teacher training* (Publication No. 10871577) [Doctoral dissertation, Kent State University]. ProQuest Dissertations and Theses Global.

DeSantis, W. J. (2018). *The effect of data-driven feedback and coaching using a mixed-reality simulation in a preservice teacher education program* [Doctoral dissertation, Western Connecticut State University]. WestCollections Repository. https://westcollections.wcsu.edu/

Devries, K., Balliet, M., Thornhill, K., Knight, L., Procureur, F., N'Djore, Y. A., N'Guessan, D. G., Merrill, KG, Dally, M., Allen, E., Hossain, M., Cislaghi, B., Tanton, C., & Quintero, L. (2021). Can the 'Learn in peace, educate without violence' intervention in Cote d'Ivoire reduce teacher violence? Development of a theory of change and formative evaluation results. *BMJ OPEN*, *11*(11). https://doi.org/10.1136/bmjopen-2020-044645

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

Domitrovich, C. E., Bradshaw, C. P., Berg, J. K., Pas, E. T., Becker, K. D., Musci, R., Embry, D. D., & Ialongo, N. (2016). How do school-based prevention programs impact teachers? Findings from a randomized trial of an integrated classroom management and social-emotional program. *Prevention Science*, *17*(3), 325–337. https://doi.org/10.1007/s11121-015-0618-z

Eder-Karavaya, A., Lohr, J., & Treutner, D. (2021). Förderung der berufsspezifischen Selbstwirksamkeitserwartung im Lehramtsstudium. Entwicklung und Evaluation eines Trainings für Lehramtsstudierende [Promoting professional self-efficacy in pre-service teachers. Development and evaluation of a training program]. *Psychologie in Erziehung und Unterricht*, *68*(4), 233–247. https://doi.org/10.2378/peu2021.art19d

Eginli, I., & Solhi, M. (2021). The impact of practicum on pre-service EFL teachers' self-efficacy beliefs: First step into professionalism. *Shanlax International Journal of Education*, *9*(4), 223–235. https://doi.org/10.34293/education.v9i4.4135

Erdem, M. (2008). The effects of the blended teaching practice process on prospective teachers' teaching self - efficacy and epistemological beliefs. *Eurasian Journal of Educational Research*, *8*(30), 81–98.

Fernandes, L., Peixoto, F., Gouveia, M. J., Silva, J. C., & Wosnitza, M. (2019). Fostering teachers' resilience and well-being through professional learning: Effects from a training programme. *Australian Educational Researcher*, *46*(4), 681–698. https://doi.org/10.1007/s13384-019-00344-0

Flynn, R., & Tazartes, L. (2015). *Professional development for teachers in urban schools: Enhancing behavior management skills* [Roundtable presentation]. Annual Meeting of the American Educational Research Association, Chicago, Illinois, United States.

Garwood, J. D., & Harris, A. H. (2020). Screencast-delivered professional development targeting teachers' self-efficacy and beginning-of-the-year classroom management practices. *Australasian Journal of Special and Inclusive Education*, *44*(1), 60–72. https://doi.org/10.1017/jsi.2020.3

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

Grammatikopoulos, V., Tsigilis, N., Gregoriadis, A., & Bikos, K. (2013). Evaluating an induction training program for Greek teachers using an adjusted level model approach. *Studies in Educational Evaluation*, *39*(4), 225–231. https://doi.org/10.1016/j.stueduc.2013.09.002

Green, Z. A. (2021). Strengthening career adaptation among school teachers in Pakistan: Test of strengths-based career intervention imparted through emotionalized learning experiences. *International Journal for Educational and Vocational Guidance.* Advance online publication. https://doi.org/10.1007/s10775-021-09502-2

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.

Hannemann, L., Uhde, G., & Thies, B. (2019). Evaluation eines Classroom-Management-Trainings für Lehramtsstudierende [Evaluation of a classroom management training for pre-service teachers]. *Zeitschrift für Bildungsforschung*, *9*(3), 309–327. https://doi.org/10.1007/s35834-019-00255-1

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. *The British Journal of Educational Psychology*, *90*(2), 330–348. https://doi.org/10.1111/bjep.12284

Hoffman, E. A. (2009). *The impact of the literacy coach on teachers' sense of self-efficacy during the practice of the America's choice model of literacy workshop in an urban elementary school* (Publication No. 3367300) [Doctoral dissertation, Saint Joseph's University]. ProQuest Dissertations and Theses Global.

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

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

Hui, C., Snider, L., & Couture, M. (2016). Self-regulation workshop and occupational performance coaching with teachers: A pilot study. *Canadian Journal of Occupational Therapy*, *83*(2), 115–125. https://doi.org/10.1177/0008417415627665

Jackson, N., & Miller, R. (2020). Teacher candidates' sense of self-efficacy toward classroom management. *Journal of Education*, *200*(3), 153–163. https://doi.org/10.1177/0022057419881169

Jamil, F. M. (2012). *MyTeachingPartner: A professional development intervention for teacher self-efficacy* [Paper presentation]. Annual meeting of the Society for Research on Educational Effectiveness (SREE), Washington D.C., United States.

Jang, H. R., & Reeve, J. (2021). Intrinsic instructional goal adoption increases autonomy-supportive teaching: A randomized control trial and intervention. *Learning and Instruction*, *73*, Article 101415. https://doi.org/10.1016/j.learninstruc.2020.101415

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*(Sup1). 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.

Karimi, M. N. (2011). The effects of professional development initiatives on EFL teachers' degree of self-efficacy. *Australian Journal of Teacher Education*, *36*(6), 50–62. https://doi.org/10.14221/ajte.2011v36n6.6

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., & 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

Klempin, C., Rehfeldt, D., Seibert, D., Brämer, M., Köster, H., Lücke, M., Nordmeier, V., & Sambanis, M. (2020). Stabilisierung der Selbstwirksamkeitserwartung über Komplexitätsreduktion [Stabilization of self-efficacy expectation via complexity reduction]. *Unterrichtswiss*, *48*, 151–177. https://doi.org/10.1007/s42010-019-00058-3

Klug, J., & Seethaler, E. (2021). Wie Lehramtsstudierende in unterschiedlichen Lernsettings Lerngewinn, Motivation und emotionales Befinden einschätzen [How pre-service teachers estimate their learning gains, motivation, and emotional well-being in different learning settings]. *Ludwigsburger Beiträge zur Medienpädagogik*, *21*, 1–16. https://doi.org/10.21240/lbzm/21/05

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

Knoblauch, D., & Hoy, A. W. (2008). "Maybe I can teach those kids." The influence of contextual factors on student teachers' efficacy beliefs. *Teaching and Teacher Education*, *24*(1), 166–179. https://doi.org/10.1016/j.tate.2007.05.005

Kollmayer, M., Pfaffel, A., Schultes, M.-T., Lueftenegger, M., Finsterwald, M., Popper, V., Joestl, G., Spiel, C., & Schober, B. (2019). REFLECT – ein Interventionsprogramm zum Aufbau von Lehrkräftekompetenzen für Reflexive Koedukation [REFLECT - an intervention program for building teacher competencies for reflective coeducation]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *51*(2, SI), 110–122. https://doi.org/10.1026/0049-8637/a000210

Krauskopf, K., & Knigge, M. (2017). Die Entwicklung von Selbstwirksamkeitserwartungen bei Lehramtsstudierenden bezogen auf die Gestaltung inklusiven Unterrichts - eine längsschnittliche Untersuchung zum Einfluss emotionaler und sozialer Kompetenzfacetten [The development of self-efficacy expectations in pre-service teachers regarding inclusive teaching - a longitudinal study on the impact of emotional and social competence facets]. *Empirische Pädagogik*, *31*(3), 5–20. https://doi.org/10.3726/978-3-653-02450-0/4

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

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

Kurt, G. (2017). Implementing the flipped classroom in teacher education: Evidence from Turkey. *Educational Technology & Society*, *20*(1), 211–221.

Latouche, A. P., & Gascoigne, M. (2019). In-service training for increasing teachers' ADHD knowledge and self-efficacy. *Journal of Attention Disorders*, *23*(3, SI), 270–281. https://doi.org/10.1177/1087054717707045

Lentfer, V. S., & Franks, B. (2015). The redirect behavior model and the effects on pre-service teachers' self-efficacy. *Journal of Education and Practice*, *6*(35), 79–87.

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

Ma, K., McMaugh, A., & Cavanagh, M. (2022). Changes in pre-service teacher self-efficacy for teaching in relation to professional experience placements. *Australian Journal of Education*, *66*(1), 57–72. https://doi.org/10.1177/00049441211060474

Mahasneh, A. M., & Alwan, A. F. (2018). The effect of project-based learning on student teacher self-efficacy and achievement. *International Journal of Instruction*, *11*(3), 511–524. https://doi.org/10.12973/iji.2018.11335a

Main, S., Chambers, D. J., & Sarah, P. (2016). Supporting the transition to inclusive education: Teachers' attitudes to inclusion in the Seychelles. *International Journal of Inclusive Education*, *20*(12), 1270–1285. https://doi.org/10.1080/13603116.2016.1168873

Main, S., & Hammond, L. (2008). Best practice or most practiced? Pre-service teachers' beliefs about effective behaviour management strategies and reported self-efficacy. *Australian Journal of Teacher Education*, *33*(4).

Marlow, R., Hansford, L., Edwards, V., Ukoumunne, O. C., Norman, S., Ingarfield, S., Sharkey, S., Logan, S., & Ford, T. (2015). Teaching classroom management – a potential public health intervention? *Health Education*, *115*(3/4), 230–248. https://doi.org/10.1108/HE-03-2014-0030

Marquez, B., Vincent, C., Marquez, J., Pennefather, J., Smolkowski, K., & Sprague, J. (2016). Opportunities and challenges in training elementary school teachers in classroom management: Initial results from classroom management in action, an online professional development program. *Journal of Information Technology for Teacher Education*, *24*(1), 87–109.

Matoti, S. N., Junqueira, K. E., & Odora, R. J. (2011). A comparative study of pre-service teachers' self-efficacy beliefs before and after work-integrated learning. *South African Journal of Higher Education*, *25*(6), 1140–1154.

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

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

Oh, S. (2010). *The sources that influence student teachers' sense of efficacy* (Publication No. 3438806) [Doctoral dissertation, Iowa State University]. ProQuest Dissertations and Theses Global.

O'Neill, S. C. (2015). Preparing preservice teachers for inclusive classrooms: Does completing coursework on managing challenging behaviours increase their classroom management sense of efficacy? *Australasian Journal of Special Education*, *40*(2), 117–140. https://doi.org/10.1017/jse.2015.10

Ortaçtepe, D., & Akyel, A. S. (2015). The effects of a professional development program on English as a foreign language teachers' efficacy and classroom practice. *TESOL Journal*, *6*(4), 680–706. https://doi.org/10.1002/tesj.185

Ozdemir, N. K., & Aydin, G. (2021). Brief motivational interviewing training for teachers providing career services in Turkey: A pilot study. *International Journal for Educational and Vocational Guidance*, *22*, 263–280. https://doi.org/10.1007/s10775-021-09484-1

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

Perkins Coppola, M. (2019). Preparing preservice elementary teachers to teach engineering: Impact on self-efficacy and outcome expectancy. *School Science and Mathematics*, *119*(3), 161–170. https://doi.org/10.1111/ssm.12327

Putman, S. M. (2012). Examining variations in programmatic delivery on teaching candidates' sense of efficacy. *Research in the Schools*, *19*(2), 45–61.

Reyes, M. E., Hutchinson, C. J., & Little, M. (2017). Preparing educators to teach effectively in inclusive settings. *SRATE Journal*, *26*(1), 21–29.

Ribeiro, J. J. (2009). *How does a co-learner delivery model in professional development affect teachers' self-efficacy in teaching mathematics and specialized mathematics knowledge for teaching?* (Publication No. 3387318) [Doctoral dissertation, Johnson & Wales University]. ProQuest Dissertations and Theses Global.

Roberts, T. G., Harlin, J. F., & Briers, G. E. (2008). Peer modeling and teaching efficacy: The influence of two student teachers at the same time. *Journal of Agricultural Education*, *49*(2), 13–26. https://doi.org/10.5032/jae.2008.02013

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.

Samuelsson, M., Samuelsson, J., & Thorsten, A. (2021). Simulation training is as effective as teaching pupils: Development of efficacy beliefs among pre-service teachers. *Journal of Technology and Teacher Education*, *29*(2), 225–251.

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

Sharma, U., & Nuttal, A. (2016). The impact of training on pre-service teacher attitudes, concerns, and efficacy towards inclusion. *Asia-Pacific Journal of Teacher Education*, *44*(2), 142–155.

Sheehan, C. Z., & Moore, L. L. (2019). Teacher self-efficacy in SBAE methods coursework: A mixed methods study. *Journal of Agricultural Education*, *60*(3), 219–231. https://doi.org/10.5032/jae.2019.03219

Siers, R. R., Fox, J. T., McCormick, K. M., & Paprzycki, P. (2017). *Adapting and adjusting in professional development schools to positively impact preservice efficacy* [Paper presentation]. Annual meeting of the American Educational Research Association, San Antonio, Texas, United States.

Stanton, K., Cawthon, S., & Dawson, K. (2018). Self-efficacy, teacher concerns, and levels of implementation among teachers participating in drama-based instruction professional development. *Teacher Development*, *22*(1), 51–77.

Stevens, T., Aguirre-Munoz, Z., Harris, G., Higgins, R., & Liu, X. (2013). Middle level mathematics teachers' self-efficacy growth through professional development: Differences based on mathematical background. *Australian Journal of Teacher Education*, *38*(4). https://doi.org/10.14221/ajte.2013v38n4.3

Stewart, T. (2012). Classroom teacher leadership: Service-learning for teacher sense of efficacy and servant leadership development. *School Leadership & Management*, *32*(3), 233–259. https://doi.org/10.1080/13632434.2012.688741

Stewart, T., Allen, K. W., & Bai, H. (2011). The effects of service-learning participation on pre-internship educators' teachers' sense of efficacy. *Alberta Journal of Educational Research*, *57*(3), 298–316.

Stripling, C. T., & Roberts, T. G. (2013). Investigating the effects of a math-enhanced agricultural teaching methods course. *Journal of Agricultural Education*, *54*(1), 124–138. https://doi.org/10.5032/jae.2013.01124

Sumners, S. E. (2012). *Civic education in the preservice classroom: A study of gaming* [Doctoral dissertation, Mississippi State University]. Mississippi State University Scholars Junction. https://hdl.handle.net/11668/18482

Thiel, F., Boehnke, 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

Thomas, K. E., & Mucherah, W. M. (2016). The contextual difference: Developing preservice teacher efficacy through immersive learning experiences. *Education and Urban Society*, *48*(4), 364–383. https://doi.org/10.1177/0013124514533795

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. *The Elementary School Journal*, *110*(2), 228–245. https://doi.org/10.1086/605771

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

VanParys Couet, J. M. (2014). *The effects of positive behavior intervention and supports on teacher self-efficacy and teaching anxiety* (Publication No. 3614517) [Doctoral dissertation, Walden University]. ProQuest Dissertations and Theses Global.

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

Vernon-Feagans, L., Bratsch-Hines, M., Varghese, C., Bean, A., & Hedrick, A. (2015). The targeted reading intervention: Face-to-face vs. Webcam literacy coaching of classroom teachers. *Learning Disabilities Research & Practice*, *30*(3), 135–147. https://doi.org/10.1111/ldrp.12062

Wagler, R., & Moseley, C. (2005). Preservice teacher efficacy: Effects of a secondary education methods course and student teaching. *Teacher Education and Practice*, *18*(4), 442–457.

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

Weß, R., Priemer, B., Weusmann, B., Ludwig, T., Sorge, S., & Neumann, I. (2020). Der Verlauf von lehrbezogenen Selbstwirksamkeitserwartungen angehender MINT-Lehrkräfte im Studium [The trajectory of teaching-related self-efficacy expectations of undergraduate STEM teachers]. *Zeitschrift für Pädagogische Psychologie*, *34*(3-4), 1–18. https://doi.org/10.1024/1010-0652/a000272

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

Yazici, F., & Yildirim, T. (2017). Effect of pedagogical formation program on pre-service history teachers' perceived self-efficacy. *European Journal of Educational Research*, *6*(3), 357–367. https://doi.org/10.12973/eu-jer.6.3.357

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.

Yoo, J. H. (2016). The effect of professional development on teacher efficacy and teachers' self-analysis of their efficacy change. *Journal of Teacher Education for Sustainability*, *18*(1), 84–94. https://doi.org/10.1515/jtes-2016-0007

Yough, M. (2019). Tapping the sources of self-efficacy: Promoting preservice teachers' sense of efficacy for instructing English language learners. *Teacher Educator*, *54*(3), 206–224. https://doi.org/10.1080/08878730.2018.1534031

Yüksel, H. G. (2014). Becoming a teacher: Tracing changes in pre-service English as a foreign language teachers' sense of efficacy. *South African Journal of Education*, *34*(3), Article 958. https://doi.org/10.15700/201409161104

Yüksel, İ. (2014). Investigating the impact of classroom management course on self-efficacy levels: An experimental study on pre-service teachers. *Education and Science*, *39*(171), 259–269.

Zeuch, N., Souvignier, E., Kutner, R., & Dutke, S. (2021). Empirische Arbeit: Heterogenitätsbezogene Einstellungen und Selbstwirksamkeit von Lehramtsstudierenden fördern: Das Projekt „Dealing with Diversity" an der WWU Münster [Empirical work: Promoting heterogeneity-related attitudes and self-efficacy of pre-service student teachers: The project "Dealing with diversity" at the WWU Münster]. *Psychologie in Erziehung und Unterricht*, *68*(3), 199–212. https://doi.org/10.2378/peu2021.art14d