### **REVIEW**

**Open Access** 

# The effect of professional development on in-service STEM teachers' self-efficacy: a meta-analysis of experimental studies



Xuan Zhou<sup>1,3\*</sup>, Lina Shu<sup>1</sup>, Zhihong Xu<sup>2</sup> and Yolanda Padrón<sup>1</sup>

### Abstract

This meta-analysis examined the effectiveness of teacher professional development (PD) training on improving inservice STEM teachers' self-efficacy in the K-12 teaching context. Twenty-one articles yielded 48 effect size estimates, incorporating 1412 teachers in the final analyses. Results indicated that the overall effect size of PD training on K-12 STEM teachers' self-efficacy was strong at g = 0.64 (p < 0.01). This review contributes to the STEM teacher self-efficacy literature and confirms the significant impact of PD on improving STEM teachers' self-efficacy. Results regarding the impact of substantive PD characteristics provide policymakers and practitioners seeking to improve STEM teacher self-efficacy with information for improving PD. Recommendations for researchers are also discussed in the paper.

Keywords STEM, Professional development, Teacher training, Teacher efficacy, Meta-analysis

### Introduction

Science, Technology, Engineering, and Mathematics (STEM) education has gained increasing attention globally in the past decades due to its important role in improving society's living standards, supplying economic growth, and supporting global competitiveness (Carnevale et al., 2011; Xie et al., 2015). However, research has indicated that STEM teachers continue to not be well prepared, even though STEM education is of great importance to a students' education (National Science Board, 2016, 2022). Furthermore, schools with high minority or poverty populations are more likely to lack quality K-12 STEM teachers. In addition, greater proportions of STEM teachers at these schools entered through an alternative route to certification that lack student

<sup>2</sup> Department of Agricultural Leadership, Education,

and Communications, Texas A&M University, College Station, TX 77843, USA

<sup>3</sup> Department of Research Data Service, University Libraries, Texas State University, San Marcos, TX 78640, USA



teaching and formal training experience (National Science Board, 2022; Rotermund & Burke, 2021).

Research evidence supports that STEM teachers' selfefficacy (STEM-TSE) is a fundamental factor determining their job satisfaction and willingness to stay in the profession (Kasalak & Dagyar, 2020). STEM teachers face challenges such as high attrition rates and stress levels (Fuller & Pendola, 2019). Teachers' confidence in their ability to teach is essential to teaching STEM subject areas effectively (Kelly et al., 2017; Wenner, 1995). Therefore, to improve the quality of STEM teachers' instruction, job satisfaction levels, and retention rates, it is important to increase STEM educators' self-efficacy levels.

A set of empirical studies support that providing STEM teachers with high-quality and ongoing professional development (PD) training is critical in improving STEM-TSE (Gardner et al., 2019; Kelley et al., 2020; Parker et al., 2020). PD is also associated with advancing teachers' capacities and lowering the turnover rate (Nguyen & Redding, 2018). PD activities can help teachers reflect on their professional practices, improve pedagogy, and address content knowledge (Fulton & Britton, 2011). Additionally, PD providing timing coherently

© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

<sup>\*</sup>Correspondence:

Xuan Zhou

x\_zhou@txstate.edu

<sup>&</sup>lt;sup>1</sup> Department of Educational Psychology, Texas A&M University, College Station, TX 77843, USA

embedded within teachers' daily jobs was demonstrated as effective in promoting teacher outcomes (Croft et al., 2010). As a result, effective PD activities can improve both teachers' knowledge and their confidence in teaching the content in the STEM discipline, which can have a profound impact on teacher outcomes.

To help the design and implementation of future STEM teacher PD training, determining the effectiveness of current PD training programs in improving STEM-TSE is crucial. Additionally, synthesizing the key determinants of best practices in developing a high-quality PD for K-12 STEM teachers will be beneficial for the direction of future teacher training and PD programs. The objectives of this study were to investigate the effectiveness of PD in improving K-12 STEM-TSE and examine whether learning STEM pedagogy and other substantive PD characteristics positively impacted the estimated effects of PD on STEM-TSE.

### **Conceptual framework and literature review** PD and four sources of teachers' self-efficacy

Self-efficacy is rooted in Bandura's social cognitive theory (1977, 1986). A teacher's sense of efficacy is defined as a teacher's "judgment of his or her capabilities to bring about desired outcomes of student engagement and learning, even among those students who may be difficult or unmotivated" (Tschannen-Moran & Hoy, 2001, p. 783). Effective teaching practices are influenced by teachers' beliefs about their own teaching capacities (Tschannen-Moran & Hoy, 2001). Teachers with higher efficacy beliefs are more likely to have better classroom management procedures, student engagement strategies, problem-solving skills, and adequate instructional strategies that encourage students' learning (Zimmerman, 2000).

Bandura (1977, 1997) proposed four sources of selfefficacy: mastery experiences, vicarious experiences, social persuasion, and emotional arousal. Mastery experience refers to the experiences gained when an individual takes on a new challenge and succeeds. Mastery experience provides the most powerful source of one's sense of efficacy (Bandura, 1997). Vicarious experiences refer to observing other people successfully completing a task. People are more likely to learn some positive beliefs about the self if they have positive role models in their lives. Social persuasion results when individuals receive feedback about their performance on a specific task. Emotional arousal occurs when a person experiences joy, excitement, or contentment while performing an activity.

PD is consistently used as an instrument to educate teachers. Teacher PD is defined as "activities that develop an individual's skills, knowledge, expertise, and other characteristics as a teacher" (OECD, 2009, p. 49). The research acknowledged that teacher PD training is an

essential approach to improving the quality of education (Coe et al., 2014; Darling-Hammond, 2000). Also, teacher PD programs significantly impacted the improvement of student achievement across various disciplinary foci (Blank & De Las Alas, 2009; Didion et al., 2020). In addition, a set of meta-analytic reviews of the literature have reported that teacher PD has a significant positive effect on teacher practices and outcomes, including the influence on teacher knowledge and skills, teacher socialemotional competence, and teacher well-being (Balta et al., 2015; Fischer et al., 2018; Iancu et al., 2018; Kraft et al., 2018; Oliveira et al., 2021; Thurlings & den Brok, 2017).

PD has also been found to contribute significantly to the sources of efficacy (Hill et al., 2004; Hoi et al., 2017; Ross & Bruce, 2007). For instance, PD gives teachers opportunities to complete a series of tasks related to their teaching practice and provides authentic evidence of success in mastery experiences (Tschannen-Moran & McMaster, 2009). By providing observation opportunities to see other teachers' teaching and practice, teachers gain vicarious experience by developing a sense of adequacy while contrasting with others. Teachers are socially persuaded in PD when receiving favorable or negative feedback from colleagues and trainers on how well they are doing. When teachers are enthusiastic and joyful about what they are doing, they are positively physiologically aroused, which increases their sense of efficacy (Tschannen-Moran & Hoy, 2007). Thus, PD programs can improve teachers' self-efficacy by adjusting the training environment and hierarchical difficulty levels of the training tasks.

## Core features of STEM teacher PD and its impact on teachers' self-efficacy

Desimone (2009) proposed a widely acknowledged core conceptual framework for investigating the effects of professional development on teachers' and students' outcomes. This model indicated five core features of high-quality PD that positively impact teacher attitudes and beliefs, including content focus, active learning, coherence, sustained duration, and collective and collaborative participation (Desimone, 2009). Later, Darling-Hammond et al. (2017) confirmed this framework and expanded it with an emphasis on collaboration, modeling, and support for teachers.

Regarding the core PD features for teachers in STEM disciplines, a set of studies summarized that most STEM teacher PD training was built upon Desimone's (2009) and Darling-Hammond et al.'s (2017) framework, which centered on developing teachers' content and pedagogical knowledge with various training approaches and formats (Gonzalez et al., 2022; Huang et al., 2022;

Lo, 2021). Additionally, empirical evidence has demonstrated the impact of STEM teachers' PD training on TSE as an important outcome measure of STEM teacher PD (Huang et al., 2022).

The present meta-analysis adopted the key components of Desimone's model (2009) to examine the impact of PD on STEM-TSE, since Desimone's work (2009) is one of the most widely used conceptual frameworks in studying the effectiveness of teacher PD programs. The core features of teacher PD studied in the present study include STEM-focused pedagogy training, sustained PD duration, instructional design for active learning, collaborative participation, and coherence by the timing of providing PD.

### STEM focused

The integration of STEM-focused pedagogy has been discussed as including inquiry-based teaching, engineeringdesign-based teaching, problem-based teaching, task/ project-based teaching, and the integrated-STEM (Guzey et al., 2020; Huang et al., 2022; Mohamad Hasim et al., 2022; Park et al., 2018; Thibaut et al., 2018). Teachers' knowledge of STEM pedagogy has usually been developed by two paths: a) collective learning by lectures or presentations (e.g., Atiles et al., 2013, b) scaffolding authentic experiences (Huang et al., 2022).

Empirical research has found that the development of teachers' STEM pedagogy can significantly impact STEM-TSE (Mohamad Hasim et al., 2022). Teachers who investigated inquiry-based science teaching at a science museum, for example, demonstrated significant improvement in TSE and understanding of inquiry-based instructions (Duran et al., 2009). Several other studies that investigated how inquiry-based teaching and learning pedagogy training affected STEM-TSE found similar results (Deniz & Akerson, 2013; Gosselin et al., 2010; Patrick et al., 2014). Likewise, completing tasks involving data processing, mathematical modeling, or problem-solving positively affected TSE beliefs in classroom teaching (Ertmer et al., 2014; Evans, 2011; Haney et al., 2007; Maass et al., 2022). Additionally, engaging in the engineering design process and having an authentic integrated STEM learning experience were also found to be beneficial for the development of STEM-TSE (Ferand et al., 2020; Hammack et al., 2020).

### Sustained PD duration

Literature shows that PD duration impacts teacher gains (Kennedy, 2016). Both the intensity (i.e., hours per day or per week) and the total contact time are essential components of an effective PD (Garet et al., 2001; Kowalski et al., 2020; Loucks-Horsley et al., 2009). Sustained STEM PD duration is highly correlated with other core

cant amount of hours is more likely to be of high quality (Ufnar & Shepherd, 2019). That is, PD with sustained duration gives teachers more opportunities to collaborate and engage in active learning, and spend more time on subject learning, so as to improve skills and competencies (Garet et al., 2001).

### Instructional design for active learning

Active learning in teacher PD refers to participants who are directly engaged with the training materials close to their classroom and students (Darling-Hammond et al., 2017). Active learning in STEM teacher PD varies widely in implementation, and it includes instructional approaches as diverse as self-directed theory learning and self-explored practice, observing role models, reflections, providing and receiving feedback, and learning by doing (Blank & De Las Alas, 2009; Huang et al., 2022). These active learning activities in teacher PD are built upon concrete learning theories, such as Experiential Learning Theory (Kolb, 1984), Constructivism Theory (Vygotsky, 1978), and Social Learning Theory (Bandura & Walters, 1963), which combine behavior modeling with cognitive learning.

Research supports the positive relationship between STEM teachers' active learning in PD and TSE. Richter et al. (2013) compared beginning secondary mathematics teachers with and without constructivist-oriented mentoring, which included opportunities for reflection, experimentation with various teaching strategies, and independent decision-making. Findings revealed that beginning teachers who receive constructivist mentorship experience higher levels of efficacy and job satisfaction after the training and lower levels of emotional exhaustion. Peters-Burton et al. (2015) used a cognitive apprenticeship model-based PD and found that it impacted the perceptions of scientific thinking and inquiry instruction and the self-efficacy of 19 in-service STEM teachers. The findings showed that in-service STEM teachers shifted their views of inquiry and maintained a high level of self-efficacy throughout the study. Additionally, a lecture-focused STEM integration PD with hands-on activities and teachers' self-reflection during professional learning community sessions was found to have mixed findings as to the impact on STEM-TSE (Wang & Nam, 2015).

### Collaborative participation

In STEM PD training, based on social learning theory (Bandura, 1977), collaborative participation has been increasingly seen as a core element. It refers to teacher collaborative learning facilitated by peer teachers, schools, and/or districts across small groups or disciplines (Darling-Hammond et al., 2017). STEM teachers' collaborative participation has been found to significantly impact TSE. For example, Kelley et al. (2020) investigated the effectiveness of collaborative PD training on high school STEM-TSE for integrated STEM instruction. Both science and engineering teachers worked collaboratively through an engineering design within a community of practice. The results indicated that STEM-TSE increased significantly after the PD training (Kelley et al., 2020).

### Coherence

Coherence in teacher PD refers to supporting teachers to connect learning with their knowledge and beliefs, as well as to connect their teaching with student learning that is consistent with school, district, and state policies (Desimone, 2009). The coherence feature of teacher PD is embodied by the PD providing timing. Specifically, teacher PD is usually provided with two general timing options: job-embedded versus non-job-embedded (Croft et al., 2010). Literature supports that high-guality job-embedded ongoing PD training has an effective impact on improving TSE (Althauser, 2015). The term "job-embedded" refers to teacher learning that is rooted in daily teaching practice and student learning (Darling-Hammond & McLaughlin, 1995; Hirsh, 2009). Ongoing job-embedded PD usually comes with formative assessment development, examining student work, implementing individual professional growth plans, lesson study, professional learning communities, etc. (Croft et al., 2010). Whereas for not job-embedded PD, such as a summer institute, or other separate training sessions, usually takes place in or outside the school, is removed from instruction, away from students, and is centered on issues of related to practice (Croft et al., 2010).

### **Research questions**

Although STEM education is important in educating students, few studies have systematically analyzed the characteristics of current K-12 grade level STEM teachers' PD and examined the effectiveness of PD training on STEM teachers' outcomes. Previous reviews of STEM teacher PD (see Appendix A) have included seven systematic/content reviews (Chai, 2019; Huang et al., 2022; Lo, 2021; Margot & Kettler, 2019; Mohamad Hasim et al., 2022; Seneviratne et al., 2019; Thibaut et al., 2018); and one meta-analysis (Gonzalez et al., 2022). These review studies focused on the characteristics of general STEM Teacher PD (n=5) and features of PD for STEM integration (n=3). These reviews found that the STEM Teacher PD most often examined is by content and pedagogical knowledge training and features of STEM instructional practices.

No studies, however, have specifically synthesized existing research on the effect of in-service teacher PD on overall TSE in STEM disciplines, though literature supports that PD is an effective approach to developing TSE (Gardner et al., 2019; Kelley et al., 2020; Parker et al., 2020). More specifically, no studies have used a meta-analysis approach to examine to what extent STEM teacher PD impacts TSE, and the effectiveness of learning STEM instructional practices and other STEM PD features in enhancing TSE.

Based on Bandura's social cognitive theory (1977) as it relates to sources of efficacy, and Desimone's framework (2009) of core features of teacher PD, this study used a meta-analytic approach to examine the effect of STEM teacher PD on TSE (Fig. 1: The conceptual model of this study). This meta-analysis focuses on identifying features of PD that can have a positive effect on STEM-TSE. The findings from this study can provide great insight into how PD can contribute to STEM-TSE and, therefore, enhance the design of PD in the future. The following research questions are examined:

- 1. What is the overall effect of the PD on STEM in-service teachers' self-efficacy?
- 2. How do STEM-focused pedagogy training and PD duration moderate the effect of PD on STEM in-service teachers' self-efficacy?
- 3. How do PD substantive characteristics, including active learning type, coherence feature, collaborative participation, and contribution to the source of TE, moderate the effect of PD on STEM in-service teachers' self-efficacy?

### Methods

### Literature search

The search retrieved databases from six resources: (a) the Educational Resources Information Clearinghouse (ERIC), (b) Education Source, (c) APA PsycInfo, (d) Web of Science, (e) Engineering Village, and (f) ProQuest. The search included PD and teacher efficacy as keyword concepts. For each of these searches, thesaurus terms were searched along with keywords in the titles and abstracts to find all relevant citations. The last search was retrieved in February 2023. The following keywords were used: professional development, professional training, in-service teacher education, continuing education, teacher education, faculty development, efficacy, teacher efficacy, etc. The search also included screening the reference lists of the included articles and existing review studies on the topic of STEM teacher PD effects on TSE to identify additional studies that might not have been sourced by the database search (Badampudi et al., 2015).



Fig. 1 Conceptual model

To locate new non-peer-reviewed literature (e.g., dissertations, gray literature, and technical reports), we used these search terms to comb the ProQuest database. The full version of search words and terms appears in Additional File 1.

### Inclusion criteria

All studies were reviewed and screened by two independent reviewers. The final included studies all met the following criteria:

- a. Empirical examination of the effects of PD on teacher efficacy. Secondary data analyses, literature reviews, and conceptual papers were excluded.
- b. Available in English and appeared from January 1977, to Feb 2023 (Bandura, 1977: the onset of Social Cognitive Theory).
- c. Samples included K-12 grade level, in-service teachers, teaching a subject in a STEM discipline. Studies with samples of higher education level educators were excluded.
- d. Examined the effectiveness of PD on teacher selfefficacy in classroom teaching (i.e., classroom management, student engagement, instructional strategy, and overall teaching self-efficacy) as measured quantitatively by standardized instruments and/or researcher-designed tools. Studies that only measured collective teacher efficacy were excluded.

- e. Provided the necessary quantitative information for calculating or estimating effect size.
- f. Employed randomized experimental or quasi-experimental designs and contained a control group. Single group pre/post-test design was excluded.
- g. Studies that compared the effect of the same PD on teachers' efficacy for teachers with different characteristics, such as educational level, or prior experience, were excluded.

### Screen procedures

The search in the six electronic databases resulted in an initial set of 9177 studies (Fig. 2). Covidence, a software specifically designed for systematic reviews, was adopted for the article screening. After removing duplicates, 6,365 studies were included in the title and abstract screening stage. A total of 927 studies were then identified as full-text screening eligible. Subsequently, a full-text review procedure was initiated and returned to 158 references for a further coding eligibility screening stage. Two independent raters screened during the title and abstract, and full-text screening stages, and they reached 89% and 92% agreement for the interrater reliability, respectively. The two raters then discussed and resolved the conflicts in Covidence to reach 100% agreement for the screening at each stage.



Fig. 2 PRISMA flowchart

Additionally, two raters screened the 158 studies for data coding eligibility. As this study employed the most rigorous methodology to assess the impact of PD intervention on STEM-TSE, studies were excluded if they did not fall into pre/post and control group research design trials. This data screening stage returned a rate of 100% agreement. From the 158 studies, seven studies were identified as having the pre/post and control group design but poor statistical reporting. The researchers reached out to the authors of these seven studies that did not have enough statistical information to compute the effect size. One author provided a full version of the data results via email. Overall, from the 158 studies, 137 were excluded after the coding eligibility screening stage. In summary, the multistep literature search and screening procedures resulted in a total of 21 research articles (see Additional File 2) that included 48 independent cases on in-service STEM teachers' PD effects on TSE.

### Study coding

An extensive coding scheme spreadsheet created for use with systematic reviews (Egert et al., 2018; Huang et al., 2022) was modified for use in documenting information about the following study features, including (a) study publication information, (b) teacher and PD characteristics, (c) outcome measures (TSE), and (d) statistics for calculating effect sizes. All the features of each study were coded by two independent researchers. The coders met to check the agreement on the coding. Two coders revisited the studies and decided on the most appropriate code if there was a conflict in coding. Interrater reliability was determined using Cohen's kappa, and most variables satisfied this criterion at Cohen's kappa=1. Variables of coherence and collaborative participation reached Cohen's kappa at 0.98 and 0.93, respectively. For studies where the two coders could not agree, the third author met with the two coders to resolve the disagreements, resulting in 100% agreement.

### **Outcome variables**

The outcome measure was teacher self-efficacy (TSE), which focused on teachers' overall self-efficacy in teaching in STEM disciplines, such as self-efficacy in classroom management, instructional strategy, student engagement, and STEM subject teaching. TSE that was measured in the included studies was drawn from (a) standardized instrument or its modified version, such as the Teachers' Sense of Efficacy Scale (TSES; Tschannen-Moran & Hoy, 2001); Math/Science Teaching Efficacy Belief Instrument (STEBI/MTEBI; Enochs & Riggs, 1990; Enochs et al., 2010), The Teacher Efficacy and Attitudes Toward STEM Survey (Friday Institute for Educational Innovation, 2012), and Dimensions of Attitude Toward Science (DAS; van Aalderen-Smeets & Walma van der Molen, 2013); and (b) self-designed instrument in the included studies (i.e., Nugent et al., 2018; Pinner, 2012).

### Moderator variables

Based on the conceptual framework of this study in Fig. 1 (Bandura, 1986; Desimone, 2009) and a previous review of STEM teacher PD trends (Huang et al., 2022), six moderators were examined for their contribution to improving TSE through PD examined their contribution to improving TSE through PD. The variables analyzed were:

(a) STEM-focused

Previous studies indicate that content focused on pedagogy learning is one of the most important components of high-quality PD and is dominant in current STEM teacher training (Huang et al., 2022). As indicated in the literature, the widely adopted STEM pedagogy includes inquiry-based, taskbased, problem-based, engineering design-based, and/or integrated STEM teaching (Guzey et al., 2020; Huang et al., 2022; Mohamad Hasim et al., 2022; Park et al., 2018; Thibaut et al., 2018). Therefore, according to the description of each PD program in the included articles, two categories were coded: teachers who received any type of the aforementioned STEM pedagogy training versus those who did not receive STEM pedagogy training.

(b) Sustained PD duration

As indicated by the literature, both intensity and duration of PD training are important aspects of effective PD training (Garet et al., 2001; Kowalski et al., 2020). The current meta-analysis documented the total PD contact hours for each of the studies included. The value was obtained based on the product of PD intensity (hours per dose) by PD frequency (total number of doses).

(c) Active learning

According to our proposed conceptual model, PD instructional design is one of the dominant factors determining teachers' active learning in the PD program (Blank & De Las Alas, 2009; Darling-Hammond et al., 2017; Huang et al., 2022). Three categories were identified within the included studies. First, constructivist-oriented training (e.g., van Aalderen-Smeets & Walma van der Molen, 2015), which was based on social learning theory rooted in Constructivism (Vygotsky, 1986). Constructivism emphasizes that learners construct new understandings and knowledge through active participation in the learning process and social discourse

with others (Piaget, 1973; Vygotsky, 1986, 1987). Studies were coded into this category if they mentioned that the PD instruction was based on constructivism or the self-exploratory format with an inquiry-based approach since these adhere to the constructivist learning model. The second category is cognitive apprenticeship-based design (Collins et al., 1987, 2018). This type of PD instructional design focuses on training teachers to learn by observing other exemplar teachers' best practices and fostering modeling to help with cognitive and metacognitive development, along with practice and reflection (Peters-Burton et al., 2015). Examples of included studies are Chen (2020), Long (2015), etc. The third category is a mix of the aforementioned two types (e.g., Knowles, 2017; Mintzes et al., 2013).

(d) Collaborative participation

According to collaborative learning theory, group learning aids teachers in developing their higherlevel thinking, and communication skills (Shabani et al., 2010). It is based on Vygotsky's Zone of Proximal Development (1978). We created two categories for this moderator: teachers who participated in collaborative activities versus no collaborative participation in PD. Collaborative participation was identified based on Huang et al's (2022) and Sancar et al's (2021) descriptions of collaborating with peers, such as lesson study with small group learning or discussion, co-teaching, peer coaching, peer tutoring by sharing experiences, and peer communication in the research community.

(e) Coherence

This study codified the timing of delivering the PD for the coherence feature from the included studies. According to the definition provided by Croft et al. (2010), four categories were identified: job/ classroom embedded, summer institute, separated training sessions after school or during the semester, and blended with both job-embedded and summer institute or both summer institute and followup training during the semester.

(f) PD contribution to sources of efficacy

According to the definition of sources of teacher efficacy in the literature (Bandura, 1997; Hill et al., 2004; Hoi et al., 2017; Ross & Bruce, 2007), we coded the contribution of PD to sources of efficacy based on its design and task types. Five categories were documented, including mastery experience, vicarious experiences, social persuasion, emotional arousal, and mixed if the PD contributed to more than one source.

### Effect size calculation

Morris' (2008) procedures to calculate the unconditional effect size (Hedges' *g*) for PD on STEM-TSE for each study were performed, which allowed treatment and control group comparisons and control for possible differences in the pre-training conditions (see formulas in Appendix B). Morris's procedures allow for treatment and control group comparisons and control for possible differences in the pre-training conditions. A meta-analysis for Hedge's *g* of pre-experimental scores was performed, showing that there were no baseline differences between treatment and control groups prior to the PD training ( $g_pre = -0.13$ , 95%CI [-0.27, 0.01], p = 0.06).

### Data analysis

Meta-analysis pertains to some underlying assumptions. One of the assumptions is that the integrated effect sizes are statistically independent. However, a large portion of the included studies reported multiple data cases that resulted in more than one effect size. The multiple effect sizes within a single study may not be statistically independent due to one experimental manipulation measuring more than one similar dependent variable. Therefore, a multilevel random-effects model was used to account for this dependency. The overall observed study-to-study variation was tested for heterogeneity of variance using the *Q* statistic.

To be more specific, the meta-analysis was conducted using the metafor package (version 4.0.5) from R (Viechtbauer, 2010). A meta-regression model including all moderators was performed to minimize Type I error. Robust variance estimation (RVE) of the variancecovariance matrix from the data using the *clubSandwich* package from R (version 4.0.5) was used for fitting the multilevel model accounting for dependent data. RVE was used because this approach gives way to including all dependent effect sizes in a single regression model, even when the exact dependency is unknown (Hedges et al., 2010; Pustejovsky & Tipton, 2022; Tanner-Smith & Tipton, 2014; Tanner-Smith et al., 2016). The correlation between each effect size was constrained at r=0.6(Pustejovsky & Tipton, 2022). In the meantime, to avoid multicollinearity in the meta-regression, the analyses controlling for each moderator variable were run separately to explore heterogeneity in the reported estimates to back up the results.

### **Power analysis**

A power analysis indicated that a sample size of 48 was adequate for detecting the PD effect on STEM-TSE when using the multilevel random-effects model for meta-analysis (Harrer et al., 2021). This power test was performed in R by the "dmetar" package. The power.analysis function implemented the formula by Borenstein et al. (2011) in R was used for the calculation of the power estimate. The plausible overall effect size (d) of a treatment under study compared to the control, expressed as the standardized mean difference (SMD) was set at 0.50. The expected number of studies (k) to be included in the meta-analysis was 48, and the mean sample size of the treatment group and control group in the studies to be included in the meta-analysis was set at n=14, respectively. The alpha level to be used for this power computation was 0.05 and the between-study heterogeneity was set at "moderate". The results showed that the power was 100%, which indicates that there were enough samples in this metaanalysis. We also gave the power analysis a second test using the sample size of the number of articles included (n=21). The results showed a power of 99.61%, which is higher than the usual threshold of 80% and showed that there were enough samples in this meta-analysis.

### **Publication bias**

Publication bias has the potential to skew meta-analytic results, which perhaps warrants having considerable doubts about meta-analyses that claim positive findings (Mathur & VanderWeele, 2020). Publication bias was examined by funnel plot (Duval & Tweedie, 2000) and Egger's regression test (Egger et al., 1997). Additionally, sensitivity analysis was conducted using leave-one-out analysis and Cook's distance measurements, as they essentially combine information about the leverage and fit of a study in the meta-analysis (Viechtbauer & Cheung, 2010). The *metafor* package in R was used for publication bias and sensitivity analysis checks (Viechtbauer, 2010).

### Results

### Summary of studies

In total, there were 21 articles, from which 48 effect sizes were extracted. Data from a total of 1412 teachers were included in the analysis. The sample size varied from 13 to 166 STEM educators at the end of each PD. Looking at the publication information of the included articles, about 62% were published (n=13) and 38% were unpublished (n=8; i.e., conference papers and dissertations), ranging from the year 1997 to 2022. The majority of the articles (n=14, 67%) were conducted in the United States, and the remaining articles (n=7, 33%) were from countries including China, Canada, Lebanon, Germany, Greece, and the Netherlands.

In terms of the methodological features of the included studies (n=48), approximately half of them (n=24) had a randomized controlled trial design, whereas the other half (n=24) had a quasi-experimental

design with the non-randomized assignment of participants. Regarding the grade level of the participating STEM teachers, 41% were from the primary grade level, 40% were from the purely secondary level, and 19% of the studies investigated mixed levels with both primary and secondary levels. The sample size and study characteristics are presented in Table 1. A more detailed description of each included study is included in Additional File 3.

The descriptive statistics for specific characteristics of PD training programs are presented in Table 2. In terms of STEM-focused pedagogy training, 29% of studies introduced inquiry-based learning pedagogy to teachers (n = 14), 23% of studies included integrated STEM learning pedagogy instruction (n = 11), and 19% of studies had task/problem-based learning (n=9). 13% studies covered general STEM content and pedagogy but did not include these specific learning approaches (n=6). Furthermore, 16% of studies did not focus on training teacher content and pedagogical knowledge (n=8); therefore, there was no related instruction. Regarding sustained PD duration, among these included studies, the average and median total contact hours of the PD programs were 25.86 and 18.00 h, respectively (n=31).

For the active learning features coded by its PD instructional design, 33% were coded as constructivist-oriented (n=16), 40% were cognitive apprenticeship-based design (n=19), and 27% studies adopted a mixed PD instructional design (n = 13). Moreover, 69% of studies reported that teachers had collaborative participation in the PD training (n = 33), and the other 31% of studies did not (n = 15). Additionally, for the coherence coded by PD providing timing, 13% of studies conducted PD training as job-embedded (n=6), 29% of studies offered their PD training during the summer that was called summer institutes (n = 14), 31% of studies provided separate PD training in the semester (n=15), and 19% of studies offered as blended with both job-embedded and summer institute or both summer institute and follow-up training during the semester (n=9). Four cases did not clearly indicate their coherence status by providing timing.

With regard to the PD contribution to the sources of teacher efficacy, 21% of studies demonstrated that the PD program purely contributed to the source of mastery experience for self-efficacy (n=10), 6% of studies only contributed to the source of social persuasion (n=3), 19% of studies had a PD program that only contributed to the source of vicarious experiences for self-efficacy (n=9), and 4% PD programs contributed to the emotional arousal only (n=2). However, 50% of PD programs contributed to multiple sources of self-efficacy (n=24).

#### Table 1 Sample sizes and study characteristics

Sample size			
Number of included articles (p	pre-post and treatment–control)		21
Number of independent studi	es (TSE as outcome measure effect sizes)		48
Article publication informat	ion ( <i>n</i> =21)	# of articles	% of articles (%)
Publish status	Journal articles	13	62
	Unpublished manuscripts	8	38
Country	United States	14	67
	Other countries	6	33
Methodological features of	the included studies (n=48)	# of studies	% of studies (%)
Research design	Randomized controlled trial	24	50
	Quasi-experimental	24	50
Grade level	Primary level	20	41
	Secondary level	19	40
	Mixed K-12	9	19

## Research question 1: main effect of overall unconditional estimate

To address our first research question, the effects of PD on K-12 STEM teachers' self-efficacy were investigated by the unconditional model without any moderator variables (Table 3). The overall summary estimate for the 48 standardized mean difference (SMD) effect sizes was 0.64 (95% CI [0.20, 1.08]) with a *p*-value of 0.0045. The data demonstrated a statistically significant high amount of heterogeneity ( $I^2$ =94.99%, *Q* (47)=362.381, *p*<0.0001; Higgins et al., 2003). This result showed that, compared with the control group, on average, the STEM teachers' PD training yielded a statistically significant impact on STEM teachers' self-efficacy with a medium SMD effect size (Cohen, 1977).

### Sensitivity analysis and publication bias

The results effectively demonstrated that the current meta-analysis does not have significant publication bias. A funnel plot in Fig. 3 demonstrated an approximately symmetric plot for the effect sizes, which suggests free of publication bias. Egger's regression test (Egger et al., 1997) was further inspected as a supplement test (t=0.72, p=0.47). Its non-significant result supports no statistically significant publication bias. As it was replicated in all leave-one-out sensitivity analyses, the positive effect of PD on TSE was not driven by specific studies ([0.4922, 0.6517]; ps<0.0001). Cook's distance measure detected one study to be a potential sensitive case (Cook's d=0.22). The random-effects model was therefore twice

run, once with and once without that study. The difference between the  $I^2$  with and without that study was 1.17%, indicating a minimally decreased between-study variation by removing this study. Given these results, the study was kept in the analyses.

## Research question 2: moderation effect of STEM-focused pedagogy learning

A meta-regression model was performed to examine the moderation effect of STEM-focused pedagogy learning (see Model 1 in Table 4). First, we hypothesized that teachers who received STEM-focused pedagogical knowledge, such as inquiry-based, problem-based, and/ or task-based learning strategies, would demonstrate a higher effect size on improving teacher self-efficacy than teachers who did not receive STEM pedagogical knowledge from the PD training. However, the adjusted moderation effects of STEM pedagogy learning were statistically and negatively significant at p < 0.01 ( $\beta = -1.46$ , 95%CI=[-2.55, -0.37]), indicating that STEM-focused pedagogy learning was less effective in promoting TSE than those not learning.

### Subsequent analysis

The results of the effect of STEM-focused pedagogy learning were negative compared with teachers who did not learn STEM pedagogy. This finding raised a flag reminding us to look into whether the impact of STEMfocused pedagogy learning interacted with the impact of PD duration or dosage. Literature shows that PD

PD characteristics	# of studies	% or mean (median) (%)
STEM-focused		
Inquiry-based	14	29
Integrated STEM	11	23
Task/problem-based	9	19
No specific pedagogy learning	14	29
Active learning		
Constructivist-oriented	16	33
Cognitive apprenticeship- based	19	40
Mixed	13	27
Collaborative participation		
Participated	33	69
Not participated	15	31
Coherence		
Job-embedded	6	13
Summer institutes only	14	29
Separate training in semester	15	31
Mixed	9	19
Not specified	4	8
Contribution to sources of efficacy		
Mastery experience	10	21
Social persuasion	3	6
Emotional arousal	2	4
Vicarious experiences	9	19
Multiple sources of efficacy	24	50
PD duration	31	25.86 h (18 h)

Table 3 Unconditional model

Unconditional main effect	Effect size	SE	р	95% CI	l <sup>2</sup>
	0.64	0.23	0.0045**	[0.20, 1.08]	94.99%
**p<0.01					

duration impacts the improvement of TSE significantly (Liu & Liao, 2019; Yang, 2020). Furthermore, the literature supports that PD duration is an important factor in strengthening certain knowledge, as a longer PD duration provides more active learning opportunities and content-focused instructional activities than a shorter duration (Desimone, 2009; Postholm, 2012). STEM-focused pedagogy is the foundation of instructional content in the STEM field. Therefore, in this study, we ran a second model to predict the interaction effect of STEM-focused pedagogy learning and PD duration on TSE (see Model 2 in Table 4). Results indicated the effect of PD duration on increasing TSE for teachers who received STEM-focused pedagogy learning is statistically significantly higher than for teachers who did not receive STEM pedagogy learning in PD training ( $\beta$ =0.24, k=27, p<0.0001). Figure 4 plots the results that teachers who did not receive STEM-focused pedagogy training in PD demonstrated a decreasing trend in the effect of predicting TSE; whereas teachers who received STEM pedagogy learning showed an increasing trend in predicting the effect on TSE.

## Research question 3: moderation effect of PD substantive features

Meta-regression models were performed to examine differences in the average effect of the moderator variables. The summary estimates for each moderator variable are in Table 5.

In terms of the active learning by PD instructional design, the cognitive apprenticeship-based group that uses demonstration-based learning with reflection and feedback showed lower moderation effect than other approaches, though it was not significant ( $\beta = -0.94$ , p > 0.05). Furthermore, when the moderation effect of PD providing timing was examined, the results revealed that the based group, PD provided as job-embedded, was more effective than other types of timing. All other timings showed statistically significant less effectiveness in improving teachers' self-efficacy [i.e., summer institutes ( $\beta = -2.36$ , p < 0.0001), separate training during the semester/after school ( $\beta = -2.17$ , p < 0.0001), and mixed timing ( $\beta = -2.22$ , p < 0.01)].

Additionally, collaborative participation and PD contribution to sources of efficacy did not show statistically significant results. The findings indicate that collaborative participation in PD activities did not significantly improve teachers' efficacy in teaching compared to those who did not participate, and each source of efficacy contributed by PD did not have statistically significant differences in predicting teachers' self-efficacy.

We are, however, cautious in interpreting the results due to the relatively small sample size. Even though running the meta-regression model that includes all moderators could prevent an increase in Type I error, all moderators were added to the model separately out of concern for multicollinearity issues (see Appendix C). Appendix C shows similar findings that back up the results in Table 5.

### Discussion

This meta-analysis supports the significant contribution of PD as the source of teachers' self-efficacy. It provides important information regarding the components of effective PD in improving K-12 grade level STEM. To the



Fig. 3 Funnel plot

 Table 4
 Effect of STEM pedagogy learning and its interaction with PD duration

Models	Regression component	β	SE	95% CI
1	Constant	1.87***	0.50	[0.90, 2.84]
	Received STEM pedagogy learning	- 1.46**	0.56	[- 2.55, - 0.37]
2	Constant	5.48***	0.63	[4.24, 6.71]
	PD duration	- 0.23***	0.04	[- 0.30, - 0.16]
	STEM-focused pedagogy learning	- 5.41***	0.64	[- 6.67, - 4.16]
	PD duration* STEM-focused pedagogy learning	0.24***	0.04	[0.17, 0.31]

\*p < 0.05, \*\*p < 0.01, and \*\*\*p < 0.001. The reference group for Model 1: teachers who did not receive STEM-focused pedagogy learning in PD. Coding for STEM-focused pedagogy learning in Model 2: 0 = no STEM pedagogy training received; 1 = STEM pedagogy training received



Fig. 4 Interaction effect of STEM pedagogy learning and PD duration

best of our knowledge, this is the first study that examines the effect of experimentally designed PD on STEM teacher self-efficacy. The results from the present study demonstrated that teacher PD had a medium effect size (g=0.64) on improving STEM teachers' self-efficacy in classroom teaching (Cohen, 1977). This meta-analysis confirms findings from previous empirical studies (e.g., Aaron Price & Chiu, 2018; Mintzes et al., 2013).

### Importance of STEM-focused pedagogy training

The current meta-regression analysis quantitatively revealed the positive contribution of learning STEM pedagogy on STEM teachers' overall self-efficacy enhancement, even though a previous review study (Seneviratne et al., 2019) qualitatively synthesized mixed findings of the PD effect on science teachers' efficacy specifically in inquiry-based pedagogy across the literature. The results of the current meta-analysis demonstrated that teachers who received STEM pedagogy training showed an increasing trend in predicting the effect on TSE, and vice versa. It also indicated that if the PD does not contain a pedagogy component, the longer the training, the worse effect it will have. This implies that when providing STEM teachers with adequate contact time for learning

Table 5	Meta-regressior	n model for	<sup>-</sup> substantive PD	features
---------	-----------------	-------------	-----------------------------	----------

Moderator	Regression component	β	SE	95% CI
Active learning	Cognitive apprenticeship-based	- 0.94	0.79	[- 2.39, 0.71]
	Mixed	- 0.44	0.67	[- 1.76, 0.87]
Collaborative participation	Participated	0.71	0.83	[- 0.92, 2.33]
Coherence	Summer institutes	- 2.36**	0.85	[- 4.02, - 0.69]
	Separate training in semester	- 2.17*	0.93	[- 4.00, - 0.35]
	Mixed	- 2.22**	0.78	[- 3.78, - 0.66]
PD contribution to sources of efficacy	Vicarious experiences	0.52	0.76	[- 1.35, 3.82]
	Social persuasion	- 0.38	0.21	[- 3.04, 1.86]
	Emotional arousal	0.11	0.56	[- 1.82, 2.47]
	Mixed	0.21	0.58	[- 0.92, 1.35]
Intercept		2.41*	1.02	[.42, 4.41]

\**p* < 0.05, \*\**p* < 0.01, and \*\*\**p* < 0.001. Reference group of coherence: job-embedded PD providing timing; reference group of active learning in PD: Constructivismoriented instructional design; reference group of Collaborative participation: no collaborative participation; and the reference group of PD contribution to sources of efficacy: mastery experience

STEM content and pedagogical knowledge, their self-efficacy is more likely to reach the optimal level than that of those who do not have training in STEM pedagogy in PD.

This result highlights the importance of sustained content-focused STEM teaching pedagogy in STEM teacher PD programs. The current STEM teaching workforce remains uneven in preparation and qualifications across countries (Athanasia & Cota, 2022; National Science Board, 2022). Research pointed out that a considerable number of STEM teachers went through alternative qualification programs rather than formal teacher education programs (Rotermund & Burke, 2021). This will result in a lack of student teaching experience, formal training on teaching pedagogy, and placing in schools with high rates of minority enrollment in the STEM teaching workforce (NSB, 2016; Rotermund & Burke, 2021). Thus, a considerable number of STEM teachers may be more aware of the challenges after reporting to the teaching position, which raises the need for STEM pedagogy.

Moreover, this study shows that rather than looking at the impact of pedagogy learning alone, taking the sustained PD duration into account might be more appropriate in predicting the effects of PD training. STEM teacher PD programs should consider that the effect of learning pedagogy on improving TSE may vary depending on the contact time of training. According to Fig. 4, though the interaction effect between PD duration and STEM-focused pedagogy learning demonstrated an increasing trend in enhancing TSE, the growth curve is slow. The relatively slow growth in TSE implies that STEM pedagogy learning impacts the improvement of TSE over an extended period of time. That is, STEM-TSE cannot be expected to improve tremendously after one shot or a short period of STEM pedagogy learning. On the contrary, it takes time, effort, and patience. Policymakers and teacher PD program developers need to notice the long-term effect of learning STEM pedagogy. Therefore, providing ongoing and consistent PD training to develop teachers' STEM pedagogy calls for more effort. Future research may need to investigate the optimal duration and intensity of STEM instructional strategy training.

### Importance of multi-facet PD design

Previous studies indicated that, among the four sources of efficacy, the mastery experience supported by teacher PD has the strongest impact on TSE (Tschannen-Moran & McMaster, 2009). However, our results demonstrated no significant difference in the contribution of PD to each aspect of the source of efficacy when predicting STEM-TSE. This finding shed light on the implications for future STEM PD programs: rather than focusing on developing teachers' mastery experience in teaching STEM subjects, the effective PD design should balance the proportion of supporting mastery experience as well as other sources of efficacy so that to reach the proximal effect on the improvement of TSE.

More specifically, the results of moderation effects by collaborative participation and PD instruction design further support the importance of contributing to sources of TSE through multi-facet PD designs. In the present meta-analysis, the effects on TSE did not differ by whether participants participated in collaborative activities in PD. Among the included studies, teacher collaborative participation was mostly designed to support verbal persuasion such as small group discussions, and peer sharing and feedback (i.e., Aaron Price & Chiu, 2018; Kaschalk-Woods et al., 2021). While embedding collaborative learning opportunities in STEM PD training, rather than only supporting verbal persuasion experience by receiving feedback from peers, implementing more higher-order thinking activities such as argumentbased inquiry and more hands-on collaborative activities (i.e., STEM experiments performed by group) to support affective states and mastery experience may boost the effects on training outcomes. Additionally, receiving feedback from experts on their practice (i.e., Nadelson et al., 2013) and collaborating with the higher education research community in STEM project teaching and learning (i.e., McCollough et al., 2016) might be the support for a multi-facet PD design with by contributing to multiple sources of efficacy.

Further, regarding the active learning reflected by PD instructional design, this study found that there were no statistically significant differences among the PD instructional approaches, such as constructivist-oriented PD instructional design, cognitive apprenticeship-based PD, and a mixed approach of both designs. Therefore, to maximize the PD effects, when designing future PD training programs, policymakers and program developers should consider the integration of mastery experience along with verbal persuasion, vicarious experience, and effective states. In addition to affording vicarious experience by encouraging STEM teachers to observe exemplified teachers' modeling and share thoughts, developing teachers' cognitive and metacognitive development by scaffolding the authentic experience with the practice of STEM pedagogy and following the STEM subject characteristics to explore the tasks and lesson designs for their real class teaching might be more effective.

Therefore, providing multi-facet PD designs that include apprentice experience to support vicarious experience, constructivism-oriented instruction supporting mastery experience with learning by doing, and collaborative participation in enhancing verbal persuasion, as well as creating a comfortable training environment (e.g., expert on-site support in classroom teaching, training format option, or providing encouragement and assistance throughout the PD) for supporting emotional arousal might be more effective in increasing STEM teachers' overall level of self-efficacy in classroom teaching.

### PD providing timing matters

This study found that job-embedded PD training tailored for their classroom instruction leads to the most effective impact among all types of PD providing timing in terms of the representation of the coherence PD feature. In line with previous empirical research, ongoing job-embedded PD training could maximize the effect on TSE (Althauser, 2015; Croft et al., 2010). Our findings support that job-embedded TPD connects current PD content with STEM teachers' classroom teaching while also providing agile and adapted support to teachers' needs.

### **Conclusion and future directions**

This meta-analysis backs up the considerable role that PD plays in TSE and sheds light on key PD elements that can help boost STEM-TSE. As far as we are aware, this is the first study that particularly looks at the overall impact of experimentally designed PD on STEM-TSE. The findings are valuable for school personnel and policymakers in providing insights for future STEM teacher PD programs. The significant positive interaction effects of PD duration and STEM pedagogy learning on TSE confirm the importance of ongoing STEM pedagogy learning to boost TSE at a proximal level. Regarding other PD features, the findings suggest that, rather than focusing on developing teachers' mastery experience, future PD training programs might be more effective in raising TSE, if they integrate all sources of efficacy by enabling teachers to plan lessons by adhering to the STEM subject characteristics, watching exemplary teachers teach, closely linking PD content to teachers' everyday instruction, encouraging teachers' reflection and feedback, and creating a comfortable training environment. Additionally, ongoing job-embedded PD timing matters in the enhancement of STEM-TSE.

Although the scope and rigor of this meta-analysis permit great confidence in the results, there are several limits to the generalizability of the findings that should be noted. First, the sample size is relatively small. Given the small sample size of this study, research results should be interpreted with caution. Second, despite including research from seven different countries, our analyses were limited to studies published in English, which might have influenced our effects. However, excluding non-English articles is unlikely to substantially alter the overall meta-analytic conclusions (Morrison et al., 2012; Nussbaumer-Streit et al., 2020). Moreover, meta-analyses are not exempt from coder/rater bias, which threatens internal validity; in the same manner, it affects individual studies. To minimize this threat, we employed Cohen's Kappa, shown to generate valid scores.

Concerning recommendations for researchers, conducting studies to investigate the effective duration or dosage of STEM instructional strategy training on improving TSE would help understand the best practices for introducing STEM pedagogy to teachers. Further, the investigation of the impact of in-service PD training on the self-efficacy of STEM teachers who were certified through alternative programs is still rare in the literature. Future research may seek best practices to develop the knowledge and skills of this group of teachers. Meanwhile, this study used the most rigorous approach to investigate the effect of PD on STEM-TSE by only including pre/post and control group design studies. The final 21 eligible studies in the current meta-analysis were screened from the initial sample of 6,365 indicating that a great number of experimental research in the STEM teacher PD field lacks a control group in the literature. A control group ensures the internal validity of studies and creates a benchmark for comparing the experimental results with experimental groups (Campbell & Stanley, 1963; Guetterman et al., 2018; Shadish et al., 2002). Therefore, future research should include comparison groups when studying teacher PD in STEM education.

Appendix A Previous reviews

Author	Thibaut et al. (2018)		
Торіс	Instructional practices for secondary education of integrated STEM		
Research design	Systematic review (included 23 studies, years ranging from 2000 to 2017)		
Grade level	Secondary		
Sample	In-service teachers		
Findings	This study provides a well-defined framework for instructional practices in integrated STEM in secondary educa- tion. The framework contains five key principles: integration of STEM content, problem-centered learning, inquiry- based learning, design-based learning, and cooperative learning		
Author	Chai (2019)		
Торіс	Perceptions of STEM TPD of TPACK		
Research design	Content analysis (included 20 studies, years ranging from 2011 to 2017)		
Grade level	Not specified		
Sample	In-service teachers		
Findings	A descriptive theory-in-practice based on the reported studies was constructed. Current STEM TPD research is featured by content, pedagogy, and technology and is treated as properties of the core phenomenon of teacher PD for STEM		
Author	Margot and Kettler (2019)		
Торіс	Teachers' perception of STEM integration and education		
Research design	Systematic review (included 25 studies, years ranging from 2000 to 2016)		
Grade level	PreK-12		
Sample	In-service teachers		
Findings	Teachers reported barriers such as pedagogical challenges, curriculum challenges, structural challenges, concerns about students, concerns about assessments, and lack of teacher support. Teachers felt supports that would improve their effort to implement STEM education included collaboration with peers, quality curriculum, district support, prior experiences, and effective PD		
Author	Seneviratne et al. (2019)		
Торіс	Interconnection among PD programs, teacher's sense of efficacy and enactment of inquiry-based instruction		
Research design	Literature review (included 46 studies, years ranging from 2004 to 2018)		
Grade level	K-12		
Sample	Pre-service and in-service teachers		
Findings	This study concludes that there are mixed findings in the literature regarding the effect of PD on science TSE in inquiry-based teaching. There might be the mediator or moderator effect of self-efficacy in the relationship between PD and the enactment of inquiry-based science teaching, particularly, for pre-service teachers		
Author	Lo (2021)		
Торіс	Design principles for effective TPD related to STEM integration		
Research design	Systematic review (included 48 studies, years ranging from 2015 to 2020)		
Grade level	K-12		
Sample	In-service teachers		
Findings	CK, PCK, and sample STEM instructional materials are the three most frequently reported elements of effective TPD programs. However, teachers encounter various obstacles to implementing integrated STEM, including pedagogi- cal challenges and structural challenges		

Author	Huang et al. (2022)	
Торіс	Trend of STEM TPD	
Research design	Systematic review (included 76 studies, years ranging from 2006 to 2020)	
Grade level	K-12	
Sample	Pre-service and in-service teachers	
Findings	Knowledge delivered in the STEM TPD were mostly PK and PCK, and a certain number of PD focused on TCK and TPACK. Most of the studies adopted mixed PD approaches and featured activities like learning by design/ doing, reflective learning, and group work	
Author	Mohamad Hasim et al. (2022)	
Торіс	Impact of STEM PD on teacher outcomes	
Research design	Systematic review (included 21 studies, years ranging from 2017 to 2021)	
Grade level	Not specified	
Sample	In-service teachers	
Findings	STEM PD activities are engineering-based, inquiry-based, problem-based, project-based, or integrated STEM affect teachers' knowledge and practice; teachers can enhance their understanding of integrated STEM and improve their understanding of each STEM discipline through these activities	
Author	Gonzalez et al. (2022)	
Торіс	Impact of STEM TPD on teacher knowledge, instruction, and student achievement	
Research design	Meta-analysis (included 37 studies, years ranging from 1989 to 2018	
Grade level	PreK-12	
Sample	In-service teachers	
Findings	Teachers who participated in STEM classroom interventions experienced improvements in CK, PCK and classroom instruction. Programs with larger impacts on teacher practice yielded larger effects on student achievement	

## Appendix **B**

Formulas of calculating Hedge's g (Morris, 2008).

g=J\*d

Where

$$d = [(M_{\text{post.m}} - M_{\text{pre.m}}) - (M_{\text{post.c}} - M_{\text{pre.c}})]/SD_{\text{pooled.pre}}$$

$$J = 1 - (3/(4 * (Nm + Nc) - 9))$$

The variance of Hedge's g was calculated by formula:

$$V_{g\Delta} = \left(\frac{N_m + N_c}{N_m * N_c} + \frac{d^2}{2*(N_m + N_c)}\right) * J^2(4)$$

### **Appendix C**

## Back-up meta-regression model for substantive PD features

Moderator	Regression component	β	SE	95% CI
Active learning	Constant	1.12**	0.39	[0.35, 1.89]
	Cognitive apprenticeship- based	- 1.04	0.53	[- 2.08, 0.00]
	Mixed	- 0.29	0.57	[- 1.39, 0.82]
Collaborative	Constant	0.31	0.17	[-0.11, 0.72]
participation	Participated	0.50	0.37	[- 0.31, 1.32]
Coherence	Constant	2.66***	0.56	[1.56, 3.76]
	Summer insti- tutes	- 2.45***	0.66	[- 3.75, - 1.15]
	Separate training in semester	- 2.40***	0.65	[- 3.67, - 1.13]
	Mixed	- 2.06**	0.70	[- 3.43, - 0.701]
PD contribution	Constant	0.59	0.16	[0.16, 1.01]
to sources of efficacy	Vicarious experi- ences	0.11	0.16	[- 0.32, 0.54]
	Social persuasion	- 0.03	0.17	[- 0.45, 0.40]
	Emotional arousal	- 0.60	0.24	[- 1.23, 0.03]
	Mixed	0.33	0.51	[- 0.80, 1.47]

\* $\rho$  < 0.05, \*\* $\rho$  < 0.01, and \*\*\* $\rho$  < 0.001; this table presents the results of metaregression with each moderator ran separately.

#### Abbreviations

PD	Professional development
STEM	Science, Technology, Engineering, and Mathematics
TSE	Teacher self-efficacy
STEM-TSE	STEM teachers self-efficacy

### **Supplementary Information**

The online version contains supplementary material available at https://doi. org/10.1186/s40594-023-00422-x.

Additional file 1: Search queries used for each database.

Additional file 2: Eligible studies included as sample of the current meta-analysis.

Additional file 3: Description of the included studies.

#### Acknowledgements

Not applicable.

### Author contributions

All authors contributed to the study's conception and design. XZ wrote the first draft of the manuscript. LS helped with the literature search and data coding. ZX and YP helped with data analysis and results reporting. All authors provided ethics and commented on previous versions of the manuscript. All authors read and approved the final manuscript.

#### Funding

Not applicable.

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

### Declarations

#### **Competing interests**

The authors declare that they have no competing interests.

Received: 10 October 2022 Accepted: 27 April 2023 Published online: 22 May 2023

### References

- Aaron Price, C., & Chiu, A. (2018). An experimental study of a museum-based, science PD programme's impact on teachers and their students. *International Journal of Science Education*, 40(9), 941–960. https://doi.org/10. 1080/09500693.2018.1457816
- Althauser, K. (2015). Job-embedded professional development: Its impact on teacher self-efficacy and student performance. *Teacher Development*, *19*(2), 210–225. https://doi.org/10.1080/13664530.2015.1011346
- Athanasia, G. & Cota, J. (2022). *The U.S. should strengthen STEM education to remain globally competitive*. Center For Strategic and International Studies. https://www.csis.org/blogs/perspectives-innovation/us-should-strengthen-stem-education-remain-globally-competitive.
- Atiles, J. T., Jones, J. L., & Anderson, J. A. (2013). More than a read-aloud: Preparing and inspiring early childhood teachers to develop our future scientists. *Teacher Education and Practice*, 26(2), 285–300.
- Badampudi, D., Wohlin, C., & Petersen, K. (2015). Experiences from using snowballing and database searches in systematic literature studies. In Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering (pp. 1–10).
- Balta, N., Arslan, M., & Duru, H. (2015). The effect of in-service training courses on teacher achievement: A meta-analysis study. *Journal of Education* and Training Studies, 3(5), 254–263. https://doi.org/10.11114/jets.v3i5. 1037
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191. https://doi.org/10.1037/0033-295X.84.2.191
- Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology*, 4(3), 359. https://doi.org/10.1521/jscp.1986.4.3.359

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

- Bandura, A., & Walters, R. H. (1963). *Social learning and personality development*. Holt Rinehart and Winston.
- Blank, R. K., & De Las Alas, N. (2009). The effects of teacher professional development on gains in student achievement: How meta analysis provides scientific evidence useful to education leaders. Council of Chief State School Officers. One Massachusetts Avenue NW Suite 700, Washington, DC 20001.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011). Introduction to meta-analysis. Wiley.
- Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 1–76). Rand McNally.
- Carnevale, A. P., Smith, N., & Melton, M. (2011). STEM: Science Technology Engineering Mathematics. *Georgetown University Center on Education and the Workforce.*
- Chai, C. S. (2019). Teacher professional development for science, technology, engineering, and mathematics (STEM) education: A review from the perspectives of technological pedagogical content (TPACK). *The Asia-Pacific Education Researcher, 28*(1), 5–13. https://doi.org/10.1007/ s40299-018-0400-7
- Chen, G. (2020). A visual learning analytics (VLA) approach to video-based teacher professional development: Impact on teachers' beliefs,

self-efficacy, and classroom talk practice. *Computers & Education, 144*, 103670. https://doi.org/10.1016/j.compedu.2019.103670

Coe, R., Aloisi, C., Higgins, S., & Major, L. E. (2014). What makes great teaching? review of the underpinning research. Project Report. Sutton Trust, London. https://dro.dur.ac.uk/13747/.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences. Routledge.

- Collins, A., Brown, J. S., & Newman, S. E. (1987). Cognitive apprenticeship: Teaching the craft of reading, writing and mathematics (Technical Report No. 403). BBN Laboratories, Cambridge, MA. Centre for the Study of Reading, University of Illinois.
- Collins, A., Brown, J. S., & Newman, S. E. (2018). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In *Knowing, Learning, and Instruction* (pp. 453–494). Routledge, USA.
- Croft, A., Coggshall, J. G., Dolan, M., & Powers, E. (2010). Job-Embedded professional development: What it is, who is responsible, and how to get it done well. Issue brief. National Comprehensive Center for Teacher Quality.
- Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education Policy Analysis Archives*, 8, 1–1. https://doi.org/10.14507/epaa.v8n1. 2000
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). Effective Teacher Professional Development. Research Brief. *Learning Policy Institute*. https:// learningpolicyinstitute.org/product/teacher-prof-dev.
- Darling-Hammond, L., & McLaughlin, M. W. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan*, 76(8), 597–604. https://doi.org/10.1177/003172171109200622
- Deniz, H., & Akerson, V. (2013). Examining the impact of a professional development program on elementary teachers' views of nature of science and nature of scientific inquiry, and science teaching efficacy beliefs. *The Electronic Journal for Research in Science & Mathematics Education*, 17(3). https://hdl.handle.net/2022/22230
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educa-tional Researcher, 38*(3), 181–199. https://doi.org/10.3102/0013189X08 331140
- Didion, L., Toste, J. R., & Filderman, M. J. (2020). Teacher professional development and student reading achievement: A meta-analytic review of the effects. *Journal of Research on Educational Effectiveness*, *13*(1), 29–66. https://doi.org/10.1080/19345747.2019.1670884
- Duran, E., Ballone-Duran, L., Haney, J., & Beltyukova, S. (2009). The impact of a professional development program integrating informal science education on early childhood teachers' self-efficacy and beliefs about inquiry-based science teaching. *Journal of Elementary Science Education*, 21(4), 53–70. https://doi.org/10.1007/BF03182357
- Duval, S., & Tweedie, R. (2000). Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000. 00455.x.
- Egert, F., Fukkink, R. G., & Eckhardt, A. G. (2018). Impact of in-service professional development programs for early childhood teachers on quality ratings and child outcomes: A meta-analysis. *Review of Educational Research*, 88(3), 401–433. https://doi.org/10.3102/0034654317751918.
- Egger, M., Smith, G. D., 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
- 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
- Enochs, L. G., Smith, P. L., & Huinker, D. (2010). Establishing factorial validity of the mathematics teaching efficacy beliefs instrument. *School Science* and Mathematics, 100(4), 194–202. https://doi.org/10.1111/j.1949-8594. 2000.tb17256.x
- Ertmer, P. A., Schlosser, S., Clase, K., & Adedokun, O. (2014). The grand challenge: Helping teachers learn/teach cutting-edge science via a PBL approach. *Interdisciplinary Journal of Problem-Based Learning*, 8(1), 1–19. https://doi. org/10.7771/1541-5015.1407
- Evans, B. R. (2011). Content knowledge, attitudes, and self-efficacy in the mathematics New York City Teaching Fellows (NYCTF) program. *School Science and Mathematics*, *111*, 225–235. https://doi.org/10.1111/j.1949-8594.2011.00081.x

- Ferand, N. K., DiBenedetto, C. A., Thoron, A. C., & Myers, B. E. (2020). Agriscience teacher professional development focused on teaching STEM principles in the Floriculture curriculum. *Journal of Agricultural Education*, *61*(4), 189–202. https://doi.org/10.5032/jae.2020.04189
- Fischer, C., Fishman, B., Dede, C., Eisenkraft, A., Frumin, K., Foster, B., & McCoy, A. (2018). Investigating relationships between school context, teacher professional development, teaching practices, and student achievement in response to a nationwide science reform. *Teaching and Teacher Education*, 72, 107–121. https://doi.org/10.1016/j.tate.2018.02.011
- Friday Institute for Educational Innovation. (2012). *Teacher efficacy and beliefs* toward STEM survey. Raleigh, NC: Author.
- Fuller, E. J., & Pendola, A. (2019). Teacher preparation and teacher retention: Examining the relationship for beginning STEM teachers. American Association for the Advancement of Science ARISE commissioned paper series.
- Fulton, K., & Britton, T. (2011). STEM teachers in professional learning communities: From good teachers to great teaching. National Commission on Teaching and America's Future.
- Gardner, K., Glassmeyer, D., & Worthy, R. (2019). Impacts of STEM professional development on teachers' knowledge, self-efficacy, and practice. *Frontiers in Education*, 4(26), 1–10. https://doi.org/10.3389/feduc.2019.00026
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915–945. https://doi.org/10.3102/00028312038004915
- Gonzalez, K., Lynch, K., & Hill, H. C. (2022). A meta-analysis of the experimental evidence linking STEM classroom interventions to teacher knowledge, classroom instruction, and student achievement. *EdWorkingPaper*, 22–515. https://doi.org/10.26300/d9kc-4264
- Gosselin, D. C., Thomas, J., Redmond, A., Larson-Miller, C., Yendra, S., Bonnstetter, R. J., & Slater, T. F. (2010). Laboratory earth: A model of online K-12 teacher coursework. *Journal of Geoscience Education, 58*(4), 203–213. https://doi.org/10.5408/1.3534859
- Guetterman, T. C., Creswell, J. W., Deutsch, C., & Gallo, J. J. (2018). Skills development and academic productivity of scholars in the NIH mixed methods research training program for the health sciences. *International Journal* of *Multiple Research Approaches*, 10(1), 373. https://doi.org/10.29034/ ijmra.v10n1a25
- Guzey, S. S., Caskurlu, S., & Kozan, K. (2020). Integrated STEM pedagogies and student learning. *Handbook of research on STEM education* (pp. 65–75). Routledge.
- Hammack, R., Gannon, P., Foreman, C., & Meyer, E. (2020). Impacts of professional development focused on teaching engineering applications of mathematics and science. *School Science and Mathematics*, 120(7), 413–424. https://doi.org/10.1111/ssm.12430
- Haney, J. J., Wang, J., Keil, C., & Zoffel, J. (2007). Enhancing teachers' beliefs and practices through problem-based learning focused on pertinent issues of environmental health science. *The Journal of Environmental Education*, 38(4), 25–33. https://doi.org/10.3200/JOEE.38.4.25-33
- Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing meta-analysis with R: A hands-on guide*. Chapman and Hall/CRC.
- Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, 1(1), 39–65. https://doi.org/10.1002/jrsm.5
- Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557.
- Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal*, 105(1), 11–30. https://doi.org/10.1086/428763
- Hirsh, S. (2009). A new definition. Journal of Staff Development, 30(4), 10–16.
- Hoi, C. K. W., Zhou, M., Teo, T., & Nie, Y. (2017). Measuring efficacy sources: Development and validation of the Sources of Teacher Efficacy Questionnaire (STEQ) for Chinese teachers. *Psychology in the Schools, 54*(7), 756–769. https://doi.org/10.1002/pits.22025
- Huang, B., Jong, M. S. Y., Tu, Y. F., Hwang, G. J., Chai, C. S., & Jiang, M. Y. C. (2022). Trends and exemplary practices of STEM teacher professional development programs in K-12 contexts: A systematic review of empirical studies. *Computers & Education, 189*, 104577. https://doi.org/10.1016/j. compedu.2022.104577

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

Kasalak, G., & Dagyar, M. (2020). The relationship between teacher Self-Efficacy and teacher job satisfaction: A meta-analysis of the Teaching and Learning International Survey (TALIS). *Educational Sciences: Theory and Practice*, 20(3), 16–33. https://doi.org/10.12738/jestp.2020.3.002.

Kaschalk-Woods, E., Fly, A. D., Foland, E. B., Dickinson, S. L., & Chen, X. (2021). Nutrition curriculum training and implementation improves teachers' self-Efficacy, knowledge, and outcome expectations. *Journal of Nutrition Education and Behavior, 53*(2), 142–150. https://doi.org/10.1016/j.jneb. 2020.11.011

Kelley, T. R., Knowles, J. G., Holland, J. D., & Han, J. (2020). Increasing high school teachers self-efficacy for integrated STEM instruction through a collaborative community of practice. *International Journal of STEM Education*, 7(1), 1–13. https://doi.org/10.1186/s40594-020-00211-w

Kelly, G. J., Cunningham, C. M., & Ricketts, A. (2017). Engaging in identity work through engineering practices in elementary classrooms. *Linguistics* and Education, 39, 48–59. https://doi.org/10.1016/j.linged.2017.05.003

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 945–980. https://doi.org/10. 3102/0034654315626800

Knowles, J. G. (2017). Impacts of professional development in integrated STEM education on teacher self-efficacy, outcome expectancy, and STEM career awareness [Doctoral dissertation, Purdue University]. ProQuest Dissertations and Theses.

Kolb, D. (1984). Experiential learning. Prentice-Hall.

Kowalski, S. M., Taylor, J. A., Askinas, K. M., Wang, Q., Zhang, Q., Maddix, W. P., & Tipton, E. (2020). Examining factors contributing to variation in effect size estimates of teacher outcomes from studies of science teacher professional development. *Journal of Research on Educational Effectiveness*, 13(3), 430–458. https://doi.org/10.1080/19345747.2020.1726538

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

Liu, Y., & Liao, W. (2019). Professional development and teacher efficacy: Evidence from the 2013 TALIS. School Effectiveness and School Improvement, 30(4), 487–509. https://doi.org/10.1080/09243453.2019.1612454

Lo, C. K. (2021). Design principles for effective teacher professional development in integrated STEM education. *Educational Technology and Society*, 24(4), 136–152.

Long, C. L. S. (2015). The impact of asynchronous online course design for professional development on science-teacher self-efficacy [Doctoral dissertation, Oklahoma State University]. ProQuest Dissertations and Theses.

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2009). Designing professional development for teachers of science and mathematics. Corwin Press.

Maass, K., Sorge, S., Romero-Ariza, M., Hesse, A., & Straser, O. (2022). Promoting active citizenship in mathematics and science teaching. *International Journal of Science and Mathematics Education, 20*(4), 727–746. https:// doi.org/10.1007/s10763-021-10182-1

Margot, K. C., & Kettler, T. (2019). Teachers' perception of STEM integration and education: A systematic literature review. *International Journal of STEM Education*, 6(1), 1–16. https://doi.org/10.1186/s40594-018-0151-2

Mathur, M. B., & VanderWeele, T. J. (2020). Sensitivity analysis for publication bias in meta-analyses. *Journal of the Royal Statistical Society. Series C, Applied Statistics*, 69(5), 1091. https://doi.org/10.1111/rssc.12440

McCollough, C., Jeffery, T., Moore, K., & Champion, J. (2016). Improving middle grades STEM teacher content knowledge and pedagogical practices through a school-university partnership. *School-University Partnerships*, 9(2), 50–59.

Mintzes, J. J., Marcum, B., Messerschmidt-Yates, C., & Mark, A. (2013). Enhancing self-efficacy in elementary science teaching with professional learning communities. *Journal of Science Teacher Education*, 24(7), 1201–1218. https://doi.org/10.1007/s10972-012-9320-1

Mohamad Hasim, S., Rosli, R., Halim, L., Capraro, M. M., & Capraro, R. M. (2022). STEM professional development activities and their impact on teacher knowledge and instructional practices. *Mathematics*, *10*(7), 1109. https://doi.org/10.3390/math10071109 Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, *11*(2), 364–386. https://doi.org/10.1177/1094428106291059

Morrison, A., Polisena, J., Husereau, D., Moulton, K., Clark, M., Fiander, M., & Rabb, D. (2012). The effect of English-language restriction on systematic review-based meta-analyses: A systematic review of empirical studies. *International Journal of Technology Assessment in Health Care, 28*(2), 138–144. https://doi.org/10.1017/S0266462312000086

Nadelson, L. S., Callahan, J., Pyke, P., Hay, A., Dance, M., & Pfiester, J. (2013). Teacher STEM perception and preparation: Inquiry-based STEM professional development for elementary teachers. *The Journal of Educational Research*, *106*(2), 157–168. https://doi.org/10.1080/00220671.2012. 667014

National Science Board. (2016). *Science and Engineering Indicators 2016*. Arlington, VA: National Science Foundation, NSB-2016-1.

National Science Board. (2022). Science and Engineering Indicators 2022: The State of U.S. Science and Engineering. National Science Foundation, NSB-2022-1. Alexandria, VA. Available at https://ncses.nsf.gov/pubs/ nsb20221.

Nguyen, T. D., & Redding, C. (2018). Changes in the demographics, qualifications, and turnover of American STEM teachers, 1988–2012. *AERA Open*, 4(3), 1–13. https://doi.org/10.1177/2332858418802790

Nugent, G., Kunz, G., Houston, J., Wu, C., Patwardhan, I., Lee, S., & Luo, L. (2018). The effectiveness of a summer institute and remotely delivered science instructional coaching in middle and high school. *Journal of Science Teacher Education*, 29(8), 760–784. https://doi.org/10.1080/1046560X. 2018.1514193

Nussbaumer-Streit, B., Klerings, I., Dobrescu, A. I., Persad, E., Stevens, A., Garritty, C., Kamel, C., Affengruber, L., King, V. J., & Gartlehner, G. (2020). Excluding non-English publications from evidence-syntheses did not change conclusions: A meta-epidemiological study. *Journal of Clinical Epidemiology*, 118, 42–54. https://doi.org/10.1016/j.jclinepi.2019.10.011

Oliveira, S., Roberto, M. S., Veiga-Simão, A. M., & Marques-Pinto, A. (2021). A meta-analysis of the impact of social and emotional learning interventions on teachers' burnout symptoms. *Educational Psychology Review*, 33(4), 1779–1808. https://doi.org/10.1007/s10648-021-09612-x

Organization for Economic Co-operation and Development (OECD). (2009). Education at a Glance 2009: OECD Indicator. Retrieved from https:// www.oecd.org/education/skills-beyond-school/43636332.pdf

Park, D. Y., Park, M. H., & Bates, A. B. (2018). Exploring young children's understanding about the concept of volume through engineering design in a STEM activity: A case study. *International Journal of Science and Mathematics Education*, 16, 275–294. https://doi.org/10.1007/ s10763-016-9776-0.

Parker, M., Ficklin, K., & Mishra, M. (2020). Teacher self-efficacy in a rural K-5 Setting: Quantitative research on the influence of engineering professional development. *Contemporary Issues in Technology and Teacher Education*, 20(4), 704–729.

Patrick, E., Dentzau, M., Roseler, K., & Southerland, S. A. (2014). Examining the influence of RETs on science teacher beliefs and practice. *Science Education*, 98(6), 1077–1108. https://doi.org/10.1002/sce.21127

Peters-Burton, E. E., Merz, S. A., Ramirez, E. M., & Saroughi, M. (2015). The effect of cognitive apprenticeship-based professional development on teacher self-efficacy of science teaching, motivation, knowledge calibration, and perceptions of inquiry-based teaching. *Journal of Science Teacher Education*, 26(6), 525–548. https://doi.org/10.1007/ s10972-015-9436-1

Piaget, J. (1973). To understand is to invent: The future of education. Penguin Books.

Pinner, P. C. (2012). Efficacy development in science: Investigating the effects of the Teacher-to-Teacher (T2T) professional development model in Hilo elementary schools. [Doctoral dissertation, Concordia University Chicago]. ProQuest Dissertations and Theses.

Postholm, M. B. (2012). Teachers' professional development: A theoretical review. *Educational Research*, *54*(4), 405–429. https://doi.org/10.1080/00131881.2012.734725

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

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. https://doi.org/10.1016/j.tate.2013.07.012

- Ross, J., & Bruce, C. (2007). Professional development effects on teacher efficacy: Results of randomized field trial. *The Journal of Educational Research*, *101*(1), 50–60. https://doi.org/10.3200/JOER.101.1.50-60
- Rotermund, S., & Burke, A. (2021). *Elementary and Secondary STEM Education*. Science & Engineering Indicators 2022. NSB-2021-1. National Science Foundation.
- Sancar, R., Atal, D., & Deryakulu, D. (2021). A new framework for teachers' professional development. *Teaching and Teacher Education, 101*, 103305. https://doi.org/10.1016/j.tate.2021.103305
- Seneviratne, K., Hamid, J. A., Khatibi, A., Azam, F., & Sudasinghe, S. (2019). Multifaceted professional development designs for science teachers' selfefficacy for inquiry-based teaching: A critical review. Universal Journal of Educational Research, 7(7), 1595–1611. https://doi.org/10.13189/ujer. 2019.070714
- Shabani, K., Khatib, M., & Ebadi, S. (2010). Vygotsky's zone of proximal development: Instructional implications and teachers' professional development. *English Language Teaching*, 3(4), 237–248. https://doi.org/10. 5539/elt.v3n4p237.
- Shadish, W., Cook, T., & Campbell, D. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin. ISBN: 0395615569
- Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods*, 5(1), 13–30. https://doi.org/10.1002/jrsm.1091
- Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex metaanalytic 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
- Thibaut, L., Ceuppens, S., De Loof, H., De Meester, J., Goovaerts, L., Struyf, A., & Depaepe, F. (2018). Integrated STEM education: A systematic review of instructional practices in secondary education. *European Journal of STEM Education*, 3(1), 2. https://doi.org/10.20897/ejsteme/85525
- Thurlings, M., & den Brok, P. (2017). Learning outcomes of teacher professional development activities: A meta-study. *Educational Review*, 69(5), 554–576. https://doi.org/10.1080/00131911.2017.1281226
- Tschannen-Moran, M., & Hoy, A. W. (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., & Hoy, A. W. (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., & 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
- Ufnar, J. A., & Shepherd, V. L. (2019). The Scientist in the classroom partnership program: An innovative teacher professional development model. *Professional Development in Education*, *45*(4), 642–658. https://doi.org/ 10.1080/19415257.2018.1474487
- van Aalderen-Smeets, S., & Walma van der Molen, J. (2013). Measuring primary teachers' attitudes toward teaching science: Development of the dimensions of attitude toward science (DAS) instrument. *International Journal of Science Education*, *35*(4), 577–600. https://doi.org/10.1080/ 09500693.2012.755576
- van Aalderen-Smeets, S. I., & Walma van der Molen, J. H. (2015). Improving primary teachers' attitudes toward science by attitude-focused professional development. *Journal of Research in Science Teaching*, *52*(5), 710–734. https://doi.org/10.1002/tea.21218.
- 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., & 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.
- Vygotsky, L. S. (1986). Thought and language-Revised edition. Cambridge, MA: Massachusetts Institute of Technology Press. (Originally published in Russian in 1934.)

- Vygotsky, L. S. (1987). Thinking and speech. In R. W. Rieber, & A. S. Carton (Eds.), The collected works of L. S. Vygotsky (Vol. 1), Problems of General Psychology (pp. 39–285). New York: Plenum Press. (Original Work Published 1934)
- Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.
- Wang, H. H., & Nam, Y. (2015). Exploring the impact of a STEM integration teacher professional development program on secondary science and mathematics teachers' perceptions of engineering and their attitude toward engineering integrated teaching. *Journal of the Korean Earth Science Society*, 36(5), 484–499. https://doi.org/10.5467/JKESS.2015.36.5. 484
- Wenner, G. (1995). Science knowledge and efficacy beliefs among preservice elementary teachers: A follow-up study. *Journal of Science Education* and Technology, 4(4), 307–315. https://doi.org/10.1007/BF02211262
- Xie, Y., Fang, M., & Shauman, K. (2015). STEM education. Annual Review of Sociology, 41, 331. https://doi.org/10.1146/annurev-soc-071312-145659
- Yang, H. (2020). The effects of professional development experience on teacher self-efficacy: Analysis of an international dataset using Bayesian multilevel models. *Professional Development in Education*, *4*6(5), 797–811. https://doi.org/10.1080/19415257.2019.1643393
- Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, *25*(1), 82–91. https://doi.org/10.1006/ceps. 1999.1016

### **Publisher's Note**

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

# Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- ► Rigorous peer review
- Open access: articles freely available online
- ► High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at > springeropen.com