

REVIEW

Open Access



# Enhancing mathematical problem posing competence: a meta-analysis of intervention studies

Ling Zhang<sup>1\*</sup> , Gabriel J. Stylianides<sup>2</sup> and Andreas J. Stylianides<sup>3</sup>

## Abstract

Mathematical problem posing, generally defined as the process of interpreting given situations and formulating meaningful mathematical problems, is academically important, and thus several interventions have been used to enhance this competence among students and teachers. Yet little is known about the interventions' various components and their relative or combined effectiveness. In this meta-analysis of 26 intervention studies in mathematics, we identified nine intervention components and found that the interventions had a medium, positive, and significant mean weighted effect size. A stepwise meta-regression analysis revealed that intervention efficacy varied by moderators relevant to the research design, sample characteristics, and intervention characteristics. The findings obtained from this meta-analysis are expected to serve as a foundation for future efforts to design and implement (more) effective interventions to improve mathematical problem posing competence.

**Keywords** Mathematical problem posing, Interventions, Meta-analysis

## Introduction

Problem posing has long been thought of as a vital intellectual activity in scientific investigation. As Einstein and Infeld (1938) pointed out, the formulation of an interesting problem is often more important than its solution. Research and practice on problem posing are relatively new compared to those on problem solving (Brown & Walter, 1993; Cai et al., 2015), but problem posing is attracting increased attention by both researchers and practitioners. The importance of problem posing in school mathematics is underpinned, for example, by a growing body of empirical evidence showing that problem posing has the potential to support students'

mathematical understanding, problem solving ability, and creativity (Bonotto & Santo, 2015; Cai & Hwang, 2002).

Given the important role of problem posing in the teaching and learning of mathematics, research and practice have aimed to develop ways to enhance students' problem posing competence. Several studies have shown that, with appropriate instructional support, students and teachers are capable of posing interesting and important mathematical problems (e.g., Cai et al., 2015; Silver & Cai, 2005). While there is an increase in the number of interventions aimed at improving participants' mathematical problem posing competence (Cai et al., 2020), these interventions have varied widely in their design and implementation, yielding mixed results. The variation suggests a lack of consensus on what constitutes an effective intervention to promote this important competence. Particularly, questions like the following remain unanswered: What are the key components of effective interventions in enhancing mathematical problem posing competence? Are certain intervention components more important than others, and if so, for which participant

\*Correspondence:

Ling Zhang  
zlingzw@gmail.com

<sup>1</sup> School of Mathematics and Statistics, Southwest University, Chongqing, China

<sup>2</sup> Department of Education, University of Oxford, Oxford, UK

<sup>3</sup> Faculty of Education, University of Cambridge, Cambridge, UK



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

groups and under what conditions? Indeed, there is a lack of a meta-analysis of past interventions in the area of mathematical problem posing to answer questions such as these, thus making it difficult for researchers and practitioners to understand, adopt, or strive toward best practice.

In this paper, we use the term *intervention* broadly to refer to a purposeful set of actions taken to improve a situation (Stylianides & Stylianides, 2013), in this case the mathematical problem posing competence of individuals at any level, from kindergarten to secondary school students as well as prospective or in-service teachers; these actions could be delivered in various settings (e.g., classrooms or laboratories) and systematic evidence would have been collected to explore their effectiveness. Specifically, in this paper we take a step toward addressing the aforementioned research gap by reporting a synthesis of different components that were incorporated in interventions aimed at enhancing participants' mathematical problem posing competence, the findings of a meta-analysis of the treatment efficacy of these interventions, and the moderators' impact on the treatment efficacy. By doing so, we expect our findings to support researchers and practitioners in their future efforts to design and implement (more) effective interventions to improve students' or teachers' mathematical problem posing competence.

## Background

### Mathematical problem posing

There is no agreement on how *mathematical problem posing* is defined, though it is generally used to refer to “the process by which, on the basis of mathematical experience, students construct personal interpretations of concrete situations and formulate them as meaningful mathematical problems” (Stoyanova & Ellerton, 1996, p. 519). As a complex notion, mathematical problem posing has been described in different ways: as a logical process (Cai & Hwang, 2020; Cai & Rott, 2024; Stoyanova & Ellerton, 1996); as a product-oriented phenomenon (Silver, 1994); as a role-centered accomplishment shaped by the norms of particular communities (Klinshtern et al., 2015; Kontorovich, 2020); and a cognitive activity, a research or instructional tool, or a learning goal (Cai & Leikin, 2020; Liljedahl & Cai, 2021).

Despite the varied manifestations embedded in definitions, there is significant overlap among them. They all view the process of problem posing as generating or revealing something new from a set of data, and it is considered to be a form of authentic mathematical inquiry (Bonotto & Santo, 2015). Problem posing is, in fact, of central importance to the discipline of mathematics and to mathematical thinking (Kilpatrick,

1987). The advancement of mathematics requires creative imagination as the result of raising new questions, creating new possibilities, and viewing old questions from new angles (Ellerton & Clarkson, 1996). Indeed, the identification and posing of good problems was recognized to be an important part of doing high-quality mathematics decades ago (Hadamard, 1945).

If a goal of education is to prepare students for the kinds of thinking they will need in the future, it seems reasonable that “the experience of discovering and creating one's own mathematics problems ought to be part of every student's education” (Kilpatrick, 1987, p. 123) rather than reserved for candidates for advanced degrees in mathematics. Partly based on realizations such as this one, in recent years, several curriculum frameworks around the world have supported the central role of problem posing in students' mathematical education as a way of helping students learn how to think creatively and engage in mathematical inquiry (e.g., Chinese Ministry of Education, 2022; Ministry of Education of Italy, 2007; National Council of Teachers of Mathematics (NCTM), 2000; Toh et al., 2023).

In line with the growing recognition of the significance of problem posing in school mathematics, there has been a surge in research studies focused on exploring various aspects of this notion. This literature can be categorized in the following three strands (Cai & Leikin, 2020; Liljedahl & Cai, 2021): research on *problem posing as a cognitive activity*, which focuses on understanding the nature of problem posing itself and its relationship with other constructs; research on *problem posing as a tool*, which investigates how problem posing can serve to improve students' or other participants' learning of mathematics more generally; and research on *problem posing as a goal*, which focuses on how one develops the capacity for posing good problems. Theoretical arguments and empirical evidence supporting the importance of problem posing competence in school mathematics describe mathematical problem posing both as a valuable *goal* in itself and as a *tool* to accomplish broader mathematical goals through engaging in problem posing activities. For instance, mathematical problem posing can deepen mathematical understanding, advance mathematical problem solving skills, promote mathematical creativity, and foster positive attitudes toward mathematics (Cai et al., 2015; Rosli et al., 2014). As elucidated by Cai (2022), this approach of viewing problem posing as a tool emphasizes engaging participants in problem posing tasks and activities to help them achieve both wider-range cognitive and noncognitive learning goals, while, at the same time, developing their problem posing competence as they engage in these tasks.

Despite widespread recognition of problem posing as an important intellectual competence in school mathematics and research that has shown that students and teachers are capable of posing worthwhile mathematical problems, participants often pose problems that are nonmathematical, irrelevant, unsolvable, unclear, or have errors (Cai & Hwang, 2002; Joaquin, 2023; Silver, 1994; Silver & Cai, 1996; Zhang et al., 2022a). Several hypotheses have been offered for these difficulties. For example, Crespo and Sinclair (2008) hypothesized that the difficulties might relate to a lack of opportunity for participants to explore the problem situation adequately during the problem posing process. English (1997) proposed that participants might lack foundations in problem posing. Ellerton (2013) argued that the difficulties might arise from little or no opportunity for participants to be involved in problem posing. Indeed, most of the mathematical problems a learner encounters during their education have been posed and formulated by others – the teacher or the textbook author (Kilpatrick, 1987).

Several efforts to address these difficulties have been made. For example, some researchers attempted to provide participants with more opportunities for exploration of mathematical situations (Crespo, 2003; English, 1998), while others explored the characteristics of disciplinary practice in order to identify strategies to facilitate high-quality problem posing (Brown & Walter, 1993; Milinković, 2015). While the results of most such efforts generally suggest that it is feasible to improve participants' mathematical problem posing competence, there is a wide variation in the design of the interventions, their research participants, and the instruments they used to measure the variables of interest including the outcomes. Thus, it is not clear what intervention designs are effective, with respect to what outcome measures, and for whom (Cai et al., 2015). A meta-analysis of these interventions and their effect on mathematical problem posing competence is sorely needed.

### **Interventions to enhance mathematical problem posing competence**

*Mathematical problem posing competence* refers to the criterion behavior as well as the knowledge, cognitive skills, and affective-motivational dispositions that underlie that behavior during engaging in the activity of mathematical problem posing (Zhang et al., 2023). Conceptually, it is assumed to involve a multitude of cognitive and affective states that are changing throughout the duration of the problem posing activity and cannot all be directly observed but rather must be inferred from observed behavior (Blömeke et al., 2015). However, the development of participants' mathematical problem posing competence has been documented to result in

particular positive observed cognitive outcomes, such as higher quality and quantity of posed problems, and more positive affective-motivational dispositions that underlie the cognitive outcomes (Bicer et al., 2020; Cai & Leikin, 2020; Zhang et al., 2023).

As discussed earlier, although there are several interventions that aimed to enhance students' and teachers' mathematical problem posing competence, these have not been systematically reviewed. There are a few reviews related to mathematical problem posing that align predominantly with the "problem posing as a tool" perspective as compared to the other perspectives we discussed earlier. For example, Rosli et al. (2014), Kul and Çelik (2020), and Wang et al. (2022) conducted a meta-analysis on the effects of engaging in problem posing activities on students' learning of mathematics. Other reviews examined the effects of engaging in problem posing activities on particular learning goals such as problem solving (Kopparla et al., 2019; Priest, 2009), and mathematical attitudes and achievement (Bevan et al., 2019). While these reviews occasionally ventured into a few studies examining the effect of engaging in problem posing activities on the development of problem posing, their primary focus remained on demonstrating the merit of *problem posing as a tool* on a broad-based impact for learning mathematics as opposed to examining *problem posing as a goal*. Accordingly, the aforementioned reviews are informative but insufficient to reveal what might constitute effective interventions, including but not limited to methods of engaging participants in problem posing activities, to enhance mathematical problem posing competence. To the best of our knowledge no attempt has been made to synthesize interventions that aimed to impact positively on participants' mathematical problem posing competence, that is, intervention studies treating *problem posing as a goal*. Hence, a systematic approach for reviewing the body of empirical research is needed to understand what intervention components might be important to include and potential moderators of treatment efficacy.

### **Intervention components**

According to a constructivism-oriented viewpoint, special attention in analyzing the core components of interventions should be paid in order to stimulate meaningful reflection (Danusso et al., 2010) including questions like, "what works?" To address this issue, Harden and Thomas (2005) described the intervention development as "ideas" for actions to affect outcome 'X,' and they suggested we think about questions like, "how do people experience 'X'?" or "what factors make it more/less likely that 'X' occurs?" Interventions (or "ideas" for actions) designed by educational researchers are invariably inspired by

theories of learning, cognition, motivation, or development (Pressley et al., 2006). As far as intervention design for enhancing participants' mathematical problem posing competence is concerned, the "ideas" for actions could emerge in the analysis of the reasons why participants have difficulties posing mathematical problems. For example, we discussed such reasons earlier including participants lacking the foundation of problem posing or opportunities to engage in problem posing or explore problem posing situations (Ellerton, 2013; English, 1997). Therefore, the interventions may provide instructional practice or offer relevant resources in response to what participants are lacking. Also, the "ideas" for actions can be inspired by evidence accumulation of strategies used for generating mathematical problems such as the "what-if-not" strategy (Brown & Walter, 1993), which involves participants listing the elements of the problem and then generating the new problem by asking "what if not the element  $k$ ."

In conclusion, prior research provided necessary theoretical foundation for the development of interventions including possible instructional practices, resources, or strategies enabling informed decisions about how to shape and organize the particular aspects of treatments (Pressley et al., 2006). In the present review, we adopted Bicer's (2021) definition of "instructional practices in mathematics education" and Boller et al.'s (2014) typology of "educational quality improvement interventions" to identify the intervention components that were incorporated in interventions that aimed to improve participants' mathematical problem posing competence. These components included *activity-based practice* that participants were required to experience (e.g., problem posing activity), *method-based assistance* that helped participants to pose problems (e.g., problem posing strategies, technology), and *environment-based support* that guided interaction (e.g., peer discussion).

#### **Potential moderators of treatment efficacy**

Empirical studies have been conducted to examine the effects of interventions on participants' mathematical problem posing competence. Given the wide range of intervention designs and implementations, it is not surprising that there is heterogeneity in effect sizes between studies. Knowledge about study features (i.e., moderators) that can explain the heterogeneity in effect sizes can be useful for researchers to evaluate the effectiveness of existing interventions and design new potentially effective interventions (Li et al., 2020a). Even though no meta-analysis has been conducted to examine the moderating effect of intervention designs on improving participants' mathematical problem posing competence, in this review we followed previous meta-analyses with respect to

mathematical learning (e.g., Myers et al., 2022; Niu et al., 2013) and grouped these moderators based on research design, sample characteristics, and intervention characteristics. In what follows, we discuss separately each group of moderators.

Regarding research design, we used Garzón et al.'s (2020) typology that grouped studies as *between-participants design* studies involving experimental and control treatments to measure the raw difference between treatments (pretest–posttest–control, posttest only with control) and *within-participants design* studies employing a single-group pretest and posttest design (single-group pretest–posttest). Within-participants designs, as highlighted by Cohen (1988) and Maxwell and Delaney (2004), benefit from increased statistical power due to minimized variability and reduced error variance, potentially leading to larger observed effect sizes. Furthermore, Niu et al. (2013) argued that within-participants designs are vulnerable to most threats to internal validity such as maturation, history, and testing, since they lack a control group, which might also contribute to larger observed treatment effect sizes compared to between-participants designs. In their review, Niu et al. (2013) empirically indicated that within-participants design studies (mean  $ES=0.312$ ,  $SE=0.087$ ) had significantly larger effect sizes than between-participants design studies (mean  $ES=0.120$ ,  $SE=0.074$ ), with a significance level of 0.10. Although this level of significance is less stringent than 0.05, it still suggests that within-participants designs tend to yield larger observed effects due to their inherent methodological characteristics. Hence, we hypothesize that studies adopting a within-participants design will yield a higher mean treatment effect size than studies using another design.

Regarding sample characteristics that could help determine for which group of participants interventions may be most useful, we considered *sample level* (K–12 students vs. prospective teachers vs. in-service teachers) and *sample size* (small group vs. medium group vs. large group). Although prior meta-analyses in mathematics have attempted to examine the moderating effect of sample characteristics, their findings have been largely inconsistent. For example, for *sample level*, Rosli et al. (2014) found that prospective teachers were strongly influenced by engaging in problem posing activities across all mathematical learning outcomes compared to grade 4–12 students. However, Wang et al. (2022) concluded that there was not enough evidence that sample level was a moderator for the effect of problem posing strategies on mathematical learning achievements. Silver (1994) suggested that students who have been exposed to traditional forms of mathematics teaching for a long time (e.g., students in higher grade levels) and were relatively successful in

learning mathematics in this style of teaching were more likely to have a lower motivation level in posing problems compared to younger students. In addition, Voica and Pelczer (2009) found that in-service teachers' pedagogical knowledge and classroom experience constrained their views of the problems they could pose. While we recognize that problem posing may hold varying significance for teachers (in-service or prospective) and students from a pedagogical standpoint, our understanding of the differences in treatment efficacy across sample level, such as students versus prospective or in-service teachers, remain limited. Thus, we expect to see greater improvement among younger participants and seek to investigate if the learners' level notably affects the efficacy of the treatment. Regarding *sample size*, some intervention studies in learning achievement showed that effect sizes of different sample sizes differed significantly (Zheng et al., 2020) while others reported no significant moderation effect (Borde et al., 2017). Considering that mathematical problem posing was a relatively new activity compared to problem solving for many participants (Cai et al., 2015; Zhang et al., 2022a, 2022b), we hypothesize that assistance through smaller group may contribute to more significant gains.

Regarding intervention characteristics that could help determine the conditions under which interventions are most effective (Myers et al., 2022), we considered *intervention duration* (short duration vs. medium duration vs. long duration), the *number of intervention components* (single component vs. multiple components), the *mode of intervention components* (activity-based practice vs. method-base assistance vs. environment-based support vs. mixed). Regarding *intervention duration*, the results from prior research have been inconsistent. Wang et al. (2022) found that longer-duration interventions were associated with larger improvement in students' mathematical dispositions compared to shorter-duration interventions, but several other studies found that delivering medium-duration interventions was the primary source of heterogeneity and influenced the most the effect size of learning achievement (Liu & Pásztor, 2022; Zheng et al., 2020). Therefore, intervention duration may have a significant impact on the improvement of participants' mathematical problem posing competence. However, the direction and strength of this impact may vary depending on other moderators, such as the target sample characteristics and the intervention delivery method. Regarding the *number and mode of intervention components*, intervention studies often have several components implemented across one or more settings by different intervention agents, with a general lack of consensus on what causes or contributes to specific outcomes associated with a particular intervention design (Sheridan

et al., 2019). The number and the mode of core intervention components that contribute to positive outcomes in mathematical problem posing competence have not been empirically determined. Such information is necessary to direct the design and implementation of effective interventions (Damschroder & Hagedorn, 2011; Sheridan et al., 2019). Hence, we hypothesize that different intervention components are associated with different levels of effectiveness at improving participants' mathematical problem posing competence.

### The focus of this meta-analysis

To take a step toward understanding the impact of existing published interventions on participants' mathematical problem posing competence, we conducted a meta-analysis of this body of research to address the following three research questions.

RQ1: What components were incorporated in published interventions for enhancing participants' mathematical problem posing competence?

RQ2: What are the overall treatment effects of the published interventions on participants' mathematical problem posing competence?

RQ3: What moderators (e.g., research design, sample characteristics, and intervention characteristics) influenced the effectiveness of published interventions on participants' mathematical problem posing competence?

At the instructional design level, we aim to cast light on the components that researchers incorporated in interventions for improving mathematical problem posing competence so as to deepen understanding of the mechanisms by which this competence can be enhanced (RQ1). In addition, we are interested in the overall treatment efficacy of published interventions on participants' mathematical problem posing competence (RQ2) and in the moderators' effect on treatment efficacy (RQ3) so as to cast light on what works best and inform the future design of (more) effective interventions.

## Methods

### Literature search

We followed standardized guidelines for systematic reviews by Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) (Page et al., 2021). In July 2021 we searched electronically the following databases, which we identified based on Depaepe et al.'s (2013) review of commonly used databases in mathematics education research: Web of Science, Educational Resources Information Center (ERIC), PsycINFO, and Springer. In order to make the searching as

comprehensive as possible, the query string with Boolean operators was set as follows, partially adapted from the searching word used in Lo et al.'s (2017) and Wang et al.'s (2022) reviews: (math\* OR algebra OR trigonometry OR geometry OR calculus OR statistics) AND (“problem posing” OR problem-posing).<sup>1</sup> In an effort to ensure all eligible publications were identified, the first 15 pages of Google Scholar search results (10 publications per page, ordered by relevance) were cross-referenced with the compiled inclusion bibliography, using the same search terms.

### **Inclusion and exclusion criteria**

The typology of research literature on mathematical problem posing (Cai & Leikin, 2020; Liljedahl & Cai, 2021) that we discussed earlier included the following three strands: research on *problem posing as a cognitive activity*, research on *problem posing as a tool*, and research on *problem posing as a goal*. We considered the last research strand, namely, problem posing as a goal regarding how one develops the capacity of posing good problems, as the only one relevant to our review given our particular focus, and we formulated the exclusion criteria so as to filter out intervention studies belonging to the other two strands. To incorporate as many pertinent publications as possible and differentiate the publications between *problem posing as a goal*, *problem posing as a cognitive activity*, and *problem posing as a tool*, we followed Stylianides et al. (2024) and examined the (main) *focus* of each publication. If the publication mentioned that enhancing participants' mathematical problem posing competence was one of the (or the only) primary research aims in its title or abstract, we classified it in this review as research on problem posing as a goal. In particular, the inclusion and exclusion criteria were considered by stages as follows.

In the stage of title and abstract screening (stage 1), the literature had to (a) be peer-reviewed and published in journal articles or book chapters; (b) be printed in English; (c) be published between January 1990 and June 2021; and (d) include the term “mathematical problem posing” or “problem posing” (in the subject of mathematics) in the title and/or abstract/keywords. Publications

that met the inclusion criteria underwent full review (stage 2), while those that did not were excluded. Additionally, publications for which a copy could not be obtained were also excluded.

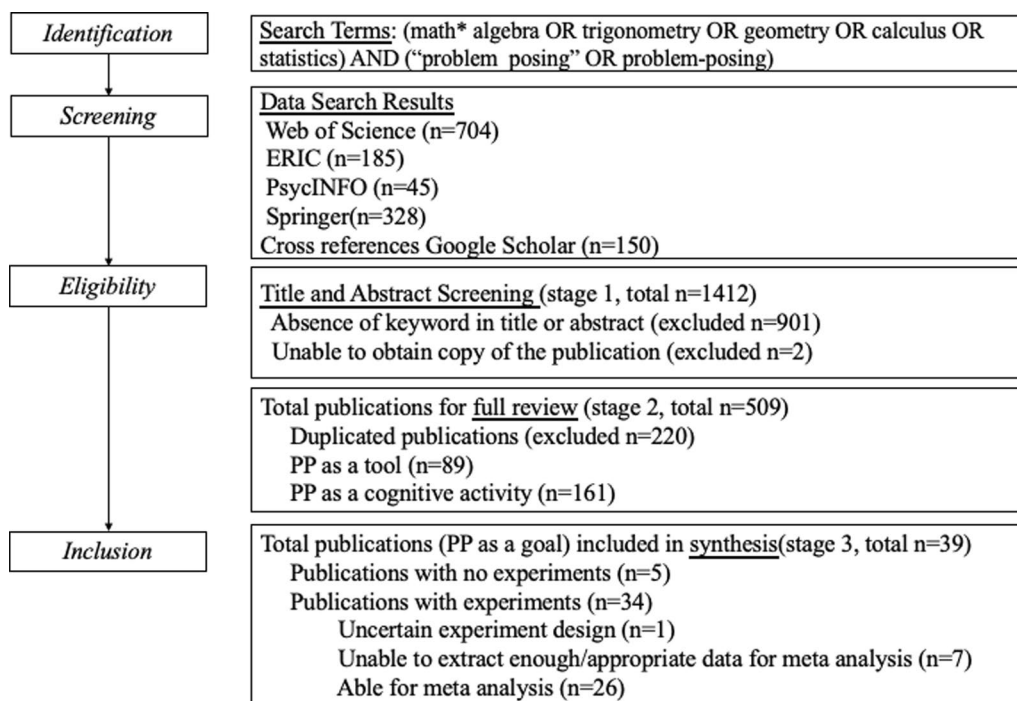
In the stage of full review (stage 2), the publications that reported research on *problem posing as a goal*, as we explained previously, were included. The number of these publications was reduced using the following exclusion criteria: (a) duplicate publications from different databases, or book chapters if there were journal articles that reported the same data/analysis or findings (journal articles were typically more elaborate); and (b) publications whose full review revealed that *problem posing as a goal* was in fact not one of the main aims.

After obtaining the set of publications based on the previous criteria for more detailed review, more restrictive criteria were applied to further select eligible studies for the meta-analysis (stage 3): (a) publications that reported on at least one type of intervention component related to mathematical problem posing along with statistical evidence; (b) publications that measured participants' mathematical problem posing and reported enough quantitative data such that the effect size could be calculated; (c) publications that considered the intervention implementation effects including the comparison of treatment and control groups or the comparison of a single pre- and post-group.

### **Publications identified and selected**

In the stage of title and abstract screening (stage 1), two research assistants (masters students majoring in mathematics education) classified 30 publications randomly selected from a total of 1412 publications. The inter-rater agreement for judging whether these 30 articles should be included for full review was 100%. They then independently reviewed the remaining 1382 publications. The sample was reduced to 509 publications, following title and abstract screening and removal of 2 publications for which a copy could not be obtained, for full review (stage 2). This second stage resulted in the exclusion of 220 duplicate publications, 89 publications that viewed problem posing as a tool (e.g., Bicer et al., 2020; Darhim et al., 2021), and 161 publications that viewed problem posing as a cognitive activity or as an independent variable (e.g., Van Harpen & Presmeg, 2013). Finally, a total of 39 publications met the inclusion criteria and were retained for systematic review (stage 3). Among them, five publications presented only the suggested intervention without an experiment (Abramovich & Cho, 2015; Aydin & Monaghan, 2018; Contreras, 2007; Lavy & Bershadsky, 2003; Milinković, 2015), and nine others reported uncertain experimental design information like missing experiment details and intervention duration

<sup>1</sup> Given that each database relies on slightly different term entry formatting, we provide the precise search terms for each database as follows. Web of Science: TS=(math\* OR algebra OR trigonometry OR geometry OR calculus OR statistics) AND TS=(“problem posing” OR problem-posing); ERIC: AB=(math\* OR algebra OR trigonometry OR geometry OR calculus OR statistics) AND AB=(“problem posing” OR problem-posing); PsycINFO: AB=(math\* OR algebra OR trigonometry OR geometry OR calculus OR statistics) AND AB=(“problem posing” OR problem-posing); and Springer (google scholar): the exact phrase=mathematical problem posing, the exact phrase=mathematics problem posing, the exact phrase=math problem posing.



**Fig. 1** Search flow summary

(Xia et al., 2007) or insufficient statistical data (Abu-Elwan, 2007; Bonotto, 2010; Courtney et al., 2014; Crespo, 2003; Kwon & Capraro, 2018; Öçal et al., 2020; Otun & Njoku, 2020); the remaining 26 publications employed at least one intervention component and reported statistical data for calculation of effect sizes and thus were used for the meta-analysis. Figure 1 presents the search flow summary.

**Data extraction and coding**

Two trained research assistants extracted the following data (where available) from the 26 included publications in the meta-analysis: (a) publication information including the study details (i.e., DOI, author names, publication year, country of origin) and type of publication (journal article, book chapter); (b) intervention information including the research design based on Garzón et al.’s (2020) typology (i.e., pretest–posttest-control that evaluates participants before and after the treatment, posttest only with control that evaluates participants only after the treatment, and pretest–posttest that evaluates a single group of participants before and after the treatment), sample characteristics (i.e., age and grade relevant to sample level, sample size<sup>2</sup>), and intervention characteristics

(i.e., intervention duration,<sup>3</sup> intervention components, the number of intervention components, and the mode of intervention components); and (c) measured outcome resulting from the intervention including the type of outcome (i.e., the quantity of posed problems, the quality of posed problems and noncognitive aspects about problem posing) and statistical outcome (i.e., the effect size or some other relevant data reflecting the level of participants’ mathematical problem posing competence).

Particularly, in terms of the data coding of the “intervention components” for RQ1, we considered the question “how participants experienced problem posing?” according to the guidance for systematic reviews proposed by Harden and Thomas (2005). The information from the procedure/design of the methodology part in each study that conducted an experiment was parsed into discrete categories of intervention components. If the study suggested intervention components without reporting on an experiment, we extracted the data from the description of the suggested intervention components and any evidence or examples that were provided

<sup>2</sup> We considered three categories of sample size as recommended by Stevens et al. (2018): less than 25; 25 to 50; and more than 50. We refer to these categories as “small group,” “medium group,” and “large group,” respectively.

<sup>3</sup> We considered four categories of intervention duration as recommended by Chauhan (2017): less than 1 day; 1 day to 1 week; 1 week to 1 month; and more than 1 month. We refer to the first category as “short duration,” to the second and third categories combined as “medium duration,” and to the last category as “long duration.” In this review, the intervention duration refers to the length of time over which an intervention was implemented or spread across. We reported the duration of specific sessions of the reviewed studies (where possible) in Table 2 in appendix.

as rationale for the suggestions. The constant comparative method (Strauss & Corbin, 2008) was adopted and used for identifying specific components that belonged to particular categories of intervention components that we discussed earlier: *activity-based practice*, *method-based assistance*, and *environment-based support*. Based on the categories of intervention components, we identified *the mode of intervention components* as single-based support (e.g., activity-based practice only), two-based support (e.g., activity-based practice combined with method-based assistance), or mixed-based support (three categories of components combined).

According to the criteria of data extraction, all research members initially examined a random selection of two articles from the 26 included studies in the meta-analysis. Results were discussed to ensure agreement and consistency in data extraction across research assistants. Two research assistants conducted the audit of the extracted data. The inter-rater reliability was 0.9 calculated by Cohen's kappa statistic among each coding point from 26 reviewed studies (Cohen, 1992). All disagreements were discussed until consensus was reached.

#### Effect sizes (ESs)

A meta-analysis integrating studies with different research designs allows us to accumulate a larger sample and provide a more complete overview of interventions on a particular topic, which avoids sample noise that can lead to an incorrect or inconclusive interpretation of the results (Lipsey & Wilson, 2001). In this review, where possible and according to Garzón et al.'s (2020) typology, we included between-participants design (i.e., pretest–posttest–control and posttest only with control) to measure the raw difference between treatments (raw-score metric) and within-participants design (single-group pretest–posttest) to evaluate the change difference after treatment (change-score metric). In order to balance the synthesis of the best quality evidence as well as describe the extent of change attributable to the mathematical problem posing competence through interventions, we opted for the raw-score metric as the common metric and we transformed the change-score metric effect size (ES) into a raw-score metric ES, as recommended by Morris and Deshon (2002). The equation  $d_{BP} = d_{WP} \sqrt{2(1 - \rho)}$  was used, where  $d_{BP}$  is the transformed ES for the raw-score metric and  $d_{WP}$  is the ES for the change-score metric.<sup>4</sup>

Specifically, the ESs from each comparison were typically calculated with Hedge's  $g$  (Lipsey & Wilson, 2001). The calculation of the ES was performed on an individual basis because some studies employed more than one outcome measure of problem posing competence. We calculated the Glass's  $\Delta$  effect size and standard error of each learning outcome in a study by the extracted data (i.e., the mean score of the experimental group on pretest and posttest,  $t$ -statistic, chi-square, sample size, etc.) and corresponding statistical formula (Lipsey & Wilson, 2001). Then, we applied the Hedge's  $g$  adjusted estimate to each ES index for correcting the sampling bias (Lipsey & Wilson, 2001). Where a study provided several ESs with respect to one particular aspect of outcome, such as the quantity of posed problems, the quality of posed problems, or a noncognitive outcome, we averaged the ESs and standard deviations to calculate the overall ES (Bernard et al., 2004). We finally grouped similar research outcomes for testing the homogeneity analysis of the effect size distribution by the Hedge's  $g$  indices.

#### Risk of publication bias

Published research only comprises a proportion of all the research conducted. However, unpublished research may differ significantly from published research due to selectivity of what gets published (Song et al., 2013). Also, studies with significant outcomes tend to get published more than those with nonsignificant outcomes (Stern & Simes, 1997). If only published papers are used for a meta-analysis, the results may be biased (Sutton, 2009), which is a major threat to meta-analytic validity. Therefore, to assess the publication bias we used a symmetric distribution of effect size as indicated by a statistically significant Egger's test (Balduzzi et al., 2019), trim and fill analysis and mixed-effects meta-regression test for funnel plot asymmetry.

#### Moderator analysis

Since there are multiple moderators, they may amplify or attenuate each other's effect on treatment effectiveness. A stepwise meta-regression analysis was used to explain the sources of differences (i.e., heterogeneity) between studies and explore moderators that impact on the treatment efficacy (Hedges, 1982). We included all the potential moderators into a stepwise meta-regression model to investigate whether particular moderators explained any of the heterogeneity of treatment effects between studies. The weighted least squares approach to estimate regression coefficients was used and the weights were based upon the random effects model to approximate inverse variance. We used small-sample adjusted  $t$ -test to determine if there was a relationship between moderators and effect sizes in the population as well as adjusted

<sup>4</sup> This and all other analyses conducted for the purposes of this paper were performed using R (version 4.1.2).



*F*-test to assess model fit (Tipton & Pustejovsky, 2015). In addition, outliers might cause the rising residual heterogeneity and the increasing mean estimated effect size (Viechtbauer & Cheung, 2010). To identify the potential outliers, we followed Myers' (2022) method and considered a value an outlier if it exceeded the 75th percentile by a factor of 1.5 times the interquartile range. The results of this calculation ranged from  $-1.033$  to  $2.588$ , so we removed one effect size above this threshold. The study of Kalmpourtzis (2019) had effect size ( $g=3.32$ ) and was considered to be an outlier. The sensitivity analysis of the models with and without outliers was performed to examine the robustness of our results to the outlier (Harwell & Maeda, 2008).

## Results

### Description of selected studies

In Table 2 in Appendix, we summarized all 26 studies that were included in the meta-analysis. Figure 2 presents the intervention duration and participant age of those 26 studies that reported experimental findings. These studies were published from 1996 (Silver et al., 1996) to 2021 (Ayvaz & Durmus, 2021; Cai & Hwang, 2021; Leavy & Hourigan, 2022)<sup>5</sup>, but most of them (19/26) were published after 2010. The studies were conducted in ten different countries/districts (Australia, China, Cyprus, Greece, Indonesia, Ireland, Israel, Japan, Turkey, USA), and included a range of participants: kindergarten students, elementary school students, secondary school students (mostly ages 5 to 18, 13 studies), university students preparing to become elementary or secondary school teachers (mostly ages 18–22, 10 studies), and in-service teachers (mostly ages over 23, 3 studies). The duration of the interventions in these studies ranged from less than 1 day (9 studies), between 1 day and less than 1 week (2 studies), more than 1 week but less than a month (5 studies), and more than 1 month (10 studies).

### Intervention components incorporated in studies

To address RQ1, we used evidence from the 26 studies included in the meta-analysis and identified the following intervention components that the studies used for enhancing participants' mathematical problem posing competence (see last column of Table 2 in Appendix). The intervention components were categorized as *activity-based practice*, *method-based assistance*, and *environment-based support*, and are summarized in Fig. 3. Next, we elaborate on each category of intervention component separately.

Regarding *activity-based practice*, the intervention components included “overview of what problem posing is” in 26.9% of the reviewed studies (WPP,  $n=7$ ), “discussion of what ‘good’ problems are” in 15.4% of the reviewed studies (WGP,  $n=4$ ), “engagement with problem posing activities” in 65.4% of the reviewed studies (PPA,  $n=17$ ), and “evaluation of posed problems” in 19.2% of the reviewed studies (EPP,  $n=5$ ). Establishing knowledge of what problem posing is (WPP) and value judgements of the products of the problem posing activity (WGP) are important and pervasive for participants' subsequent engagement in problem posing (Cai et al., 2020). The most common component (PPA), included in almost half of the reviewed studies, was to set branches of problem posing activities for participants to engage in. It provided directly experience and practices for participants to generate problems (Cai & Hwang, 2021; English, 1997). Evaluation of the problems posed by posers or presented by researchers (EPP) is an approach that enables researchers to gather evidence of how participants make judgements about their problem posing performance and the rationale behind problem selection and modification.

Regarding *method-based assistance*, reviewed studies which fell in this category offered such assistance in the form of the following intervention components: “comprehension of the problem posing situation” in 15.4% of the reviewed studies (CPPS,  $n=4$ ), “use of strategies involved in problem posing” in 19.2% of the reviewed studies (SPP,  $n=5$ ), “use of problem posing examples” in 15.4% of the reviewed studies (PPE,  $n=4$ ), and “use of technology in problem posing” in 19.2% of the reviewed studies (TPP,  $n=5$ ). CPPS helps participants gain familiarity with the situation of problem posing tasks to push and pull at the constraints of it, to become aware of its various characteristics, possible tensions, etc. (Hawkins, 2000). Studies in this category equipped participants with scaffolding involved in problem posing, such as with strategies (e.g., what-if-not strategies, SPP) and examples (PPE) to reduce their entry barrier. In addition, the use of technology (TPP) in an intervention was often intended to help participants better engage in particular intervention components (e.g., SPP or PPE). Technology also contributes to better applying realistic or game scenarios to problem posing (Aydin & Monaghan, 2018).

Regarding *environment-based support*, 46.2% of the reviewed studies attempted to incorporate such support in the form of “creation of an interactive learning environment” (ILE,  $n=12$ ) in the interventions. Interactive support leads to participants' feeling of safety and appreciation, together with an increased interest in within-resolution problem posing and openness for trying new things (Schindler & Bakker, 2020). It could be embedded

<sup>5</sup> Leavy & Hourigan (2022) was available online in 2021 (and so it was included in our review) even though it was officially published in 2022.

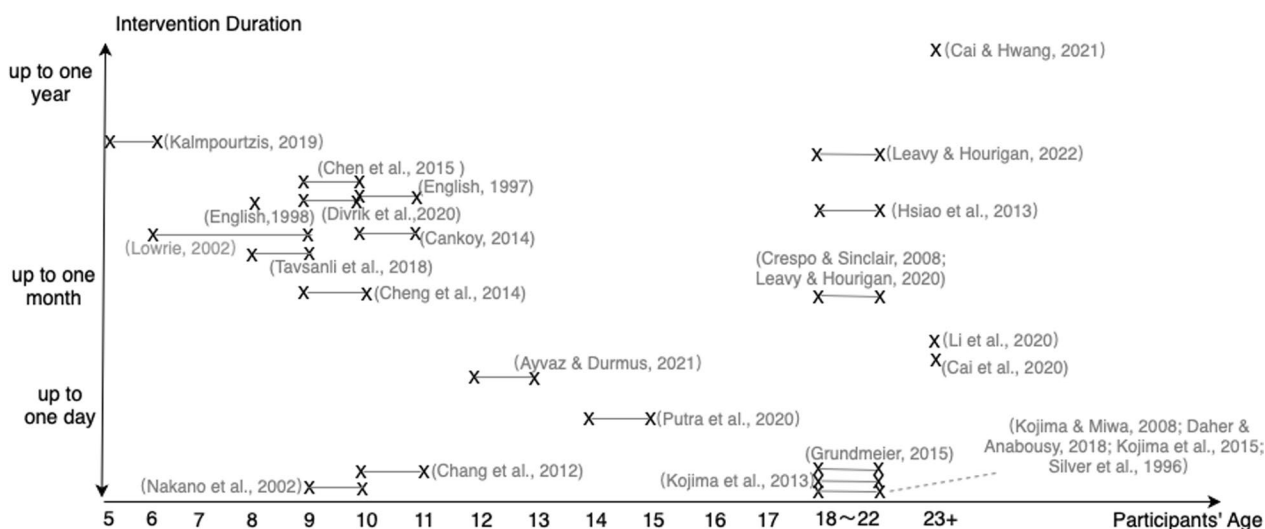


Fig. 2 Outline of the studies included in the meta-analysis (n = 26). Note All reviewed publications in this figure are included in the reference list

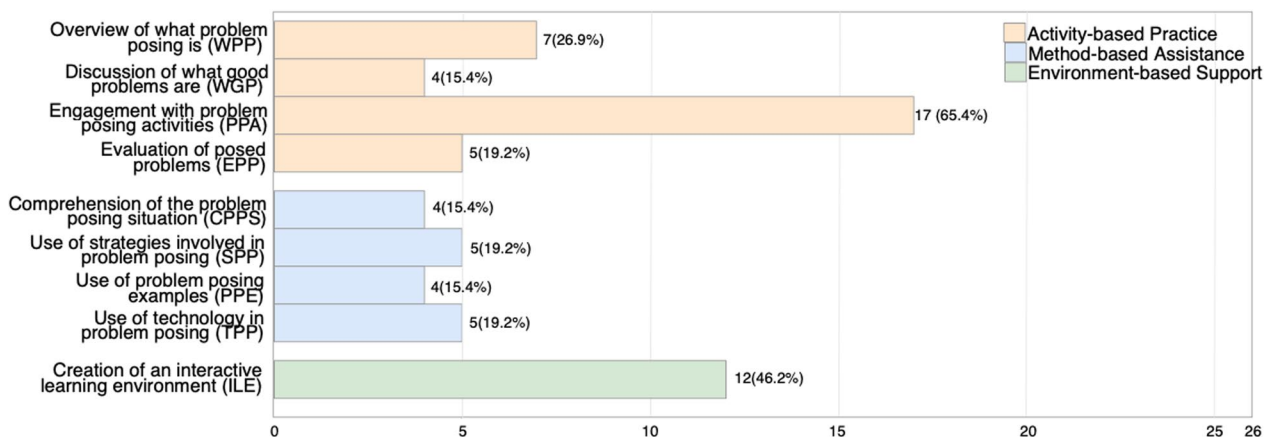


Fig. 3 Frequency of studies from those included in the meta-analysis (n = 26) that incorporated a particular category of intervention component

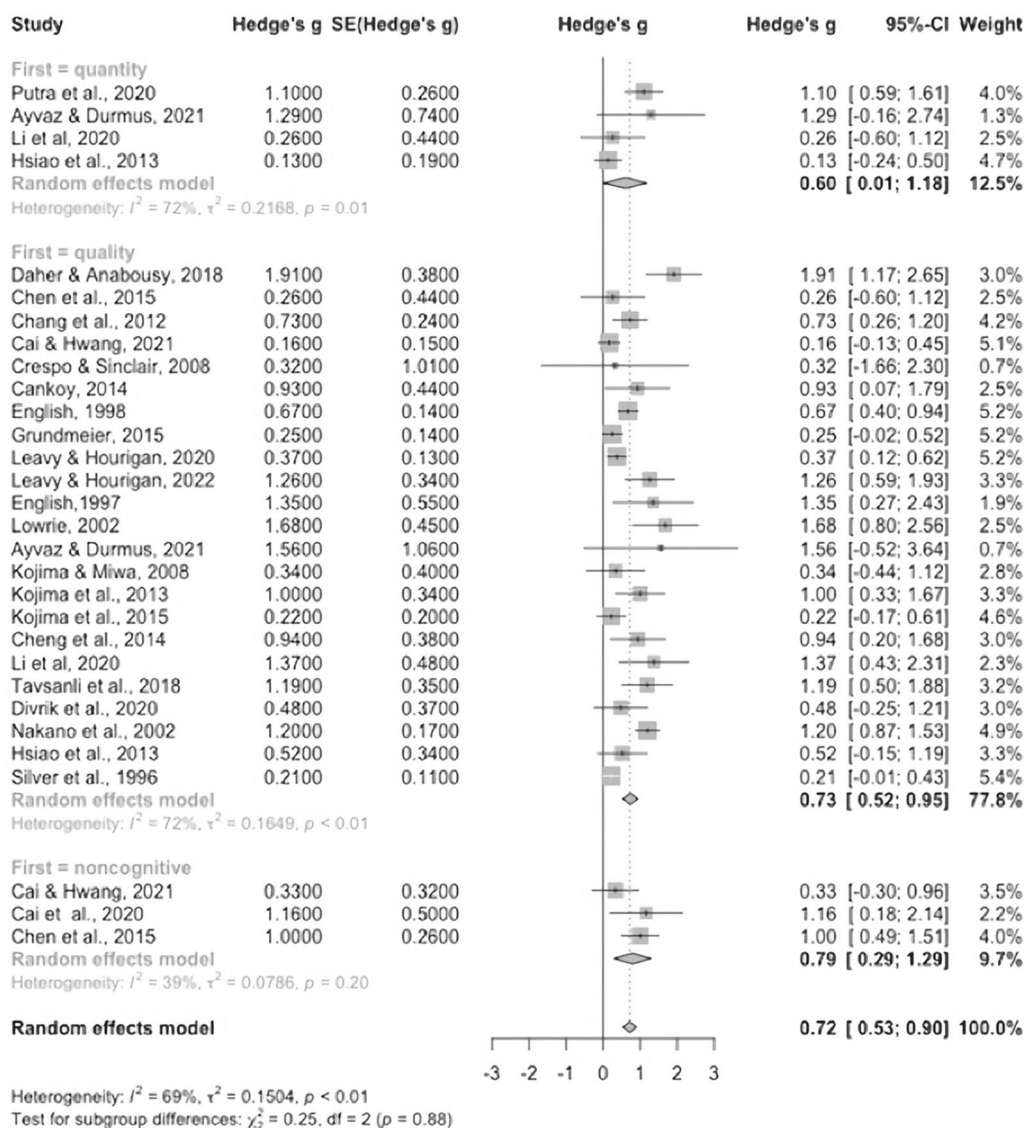
in any type of intervention components (e.g., Cai & Hwang, 2021; English, 1997).

**The treatment effect on mathematical problem posing competence**

To address RQ2, the forest plot in Fig. 4 presents the overall treatment effect on the clusters of measured outcomes, including the quantity and quality of the posed problems, and noncognitive outcomes. For the random effect model, the mean effect across 30 effect sizes from the 25 studies that included no outliers was medium, positive, and significant ( $g=0.72$ , 95% CI=[0.53, 0.90],  $p<0.001$ ) according to Hedge’s (1982) general benchmarks. Regarding each cluster of measured outcomes, the treatment effect on the quality of posed problems was larger than that on the quantity of posed problems,

and smaller than that on the noncognitive outcomes. The results showed that the effect sizes on these three clusters were 0.73, 0.60, and 0.79, respectively. However, we found no between-group variance across clusters of measured outcomes ( $p=0.88>0.05$ ), demonstrating that the interventions positively affected the quantity and quality of posed problems and noncognitive aspects of problem posing without any difference.

For the model without outliers that included 31 effect sizes from all 26 studies, the test of heterogeneity showed a large and significant residual heterogeneity estimate across studies ( $Q_{B(30)}=121.62$ ,  $I^2=75.3%$ ). The sensitivity analysis showed that removing the outlier did not substantively alter the magnitude and direction of the point estimates generated by the model that included the outlier. Although excluding the outlier reduced the



**Fig. 4** Overall treatment effect on the clusters of measured outcomes. Note We excluded the outlier (the study by Kalmipourtzis, 2019) because its unusually huge effect size of  $g = 3.32$  would likely bias the overall ES (Lipsey & Wilson, 2001)

amount of residual heterogeneity obtained using all data points by nearly 6.1%, the sensitivity analysis estimates still showed considerable heterogeneity in the effect sizes ( $Q_{B(29)} = 94.25$ ,  $I^2 = 69.2\%$ ). This showed there was significant variation among the studies even after removing influential data points.

**The moderators of treatment efficacy**

The variance inflation factor test indicated that the VIF value of several moderators was more than 10, which suggested that multicollinearity existed between moderators. To address RQ3, we used a stepwise meta-regression model as our predictive model to simultaneously

test the influence of all moderators on treatment efficacy. Using stepwise meta-regression, an initial feature selection step was performed to determine what moderators were suitable for inclusion in the final model. A criterion of  $p < 0.10$  was set for inclusion in the model (Carrara et al., 2018). The model explains a statistically significant portion of the variance ( $F(11,18) = 3.474$ ,  $p < 0.01$ ),  $R^2 = 0.68$ , and consists of seven moderators as variables: research design, sample level, sample size, number of intervention components, mode of intervention components, method-based assistance, and environment-based support.

**Table 1** Stepwise meta-regression results depending on moderators

	$\beta$	SE	$t^a$	$p$
<i>Research design (vs. pre-post-control)</i>				
Post-only-control	0.40	0.25	1.610	0.125
Pre-post	0.58	0.28	2.085 <sup>†</sup>	0.052
<i>Sample level (vs. K-12 students)</i>				
Prospective teachers	-0.89	0.21	-4.189***	<0.001
In-service teachers	-0.96	0.25	-3.788**	0.001
<i>Sample size (vs. small &lt; 25)</i>				
Medium (25-50)	-0.41	0.39	-1.038	0.313
Large (> 50)	-0.72	0.365	-1.973 <sup>†</sup>	0.064
<i>Number of components (vs. single one)</i>				
More than one	0.63	0.29	2.202*	0.041
<i>Mode of components (vs. one-based support)</i>				
Two-based support	-0.64	0.30	-2.174*	0.043
Mixed-based support	-1.86	0.57	-3.289**	0.004
<i>Intervention components (vs. nonexistent)</i>				
Method-based assistance	0.84	0.44	1.905 <sup>†</sup>	0.073
Environment-based support	0.83	0.38	2.154*	0.045

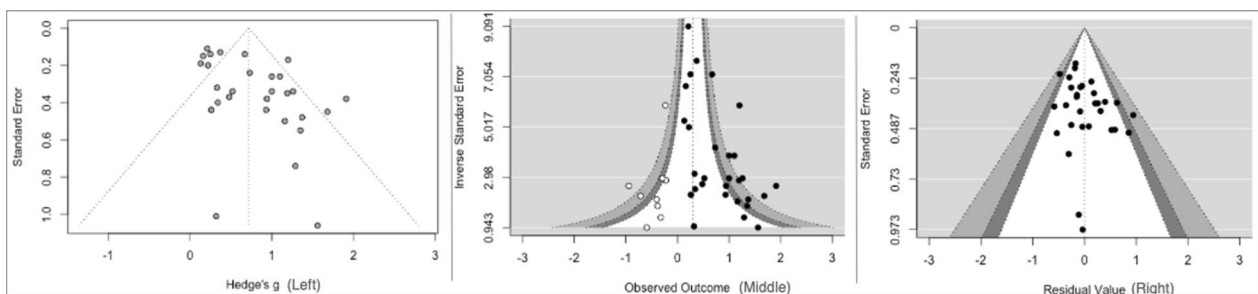
<sup>a</sup> \*\*\* $p < 0.001$ , two tailed; \*\* $p < 0.01$ , two tailed; \* $p < 0.05$ , two tailed; <sup>†</sup> $p < 0.1$ , two tailed. <sup>b</sup> We excluded the outlier (the study by Kalmpourtzis, 2019) because its unusually huge effect size of  $g = 3.32$  would likely bias the overall ES (Lipsey & Wilson, 2001).

The results of meta-regression including these seven moderators are shown in Table 1. When controlling other moderators, the following results were obtained. Regarding the research design, the overall effect size of pre-post design was on average 58% higher than that of pre-post-control design ( $t = 2.085, p < 0.1$ ). Regarding the sample level, interventions delivered to K-12 students generated an average of 89% and 96% higher effect size than those delivered to prospective teachers ( $t = -4.189, p < 0.001$ ) and in-service teachers ( $t = -3.788, p < 0.01$ ), respectively. Regarding the sample size, interventions implemented with a small group

of students (less than 25) produced an average of 72% higher effect size than those implemented with a large group ( $t = -1.973, p < 0.1$ ). In terms of the number and mode of intervention components, the results indicated that interventions employed with more than one intervention component had an average of 63% higher effect size compared to those employed with a single intervention component ( $t = 2.202, p < 0.05$ ). Similarly, the overall effect size of interventions applied in one-based support (i.e., activity-based practice, method-based assistance, or environment-based support) was on average 64% and 186% higher than that of interventions applied in two-based support ( $t = -2.174, p < 0.05$ ) and mixed-based support ( $t = -3.289, p < 0.01$ ), respectively. Regarding types of intervention components, we found that the effect sizes of interventions that incorporated method-based assistance or environment-based support were on average 84% or 83% higher than those of interventions without method-based assistance ( $t = 1.905, p < 0.1$ ) or environment-based assistance ( $t = 2.154, p < 0.05$ ), respectively.

**Publication bias**

Results of the Egger’s test showed the coefficient for the modified effect standard deviation was significant for the models ( $p < 0.05$ ), indicating that the effect size distribution was asymmetric (funnel plot was shown in Fig. 5, left). Given that any factor which is associated with both study effect and study size could confound the true association and cause as asymmetric funnel (Peters et al., 2008), we applied the trim and fill analysis of the random effects model imputed nine missing negative studies and reduced the point estimate of  $r$  to 0.5109, as shown by its confidence interval (95% CI = [0.304, 0.715]) and the heterogeneity test ( $Q_{(38)} = 150.5, p < 0.05$ ). Therefore, the true correlation between behaviors is likely to be of strong magnitude



**Fig. 5** Funnel plots illustrating the assessment of publication bias in the meta-analysis

and not enough evidence showed that plot asymmetry was caused by publication bias (Duval & Tweedie, 2000) (funnel plot is shown in Fig. 5, middle). The mixed-effects meta-regression test for funnel plot asymmetry showed that the coefficient for modified effect standard deviation was not significant for the models ( $z = 1.719$ ,  $p > 0.05$ , Fig. 5, right). Accordingly, we concluded there was not enough evidence to suggest publication bias.

## Discussion

In this review we identified nine typical intervention components—under the broad categories of *activity-based practice*, *method-based assistance*, or *environment-based support*—that had been parts of interventions published in the literature (RQ1). We also conducted a meta-analysis to examine the treatment efficacy on each cluster of measured outcomes in regard with mathematical problem posing competence. In addition, we conducted a meta-regression analysis to determine if variability in the interventions' effect sizes was associated with six kinds of moderators related to the research design, the sample, and the intervention characteristics. We next discuss our results in regard to *treatment effect* (RQ2) and *moderators' influence on the treatment efficacy* (RQ3).

### The treatment effect on mathematical problem posing competence

Regarding RQ2, our results showed that the interventions had a medium, significant, and positive impact on participants' mathematical problem posing competence (without outliers:  $g = 0.72$ ,  $p < 0.001$ ). The estimates without outliers suggest that, compared to the participants in the control groups, participants in the intervention groups demonstrated about 0.72 SD unit improvement in their mathematical problem posing scores. The mean treatment estimates without outliers indicated that approximately 76% of the students in the intervention groups scored above the mean of their peers in the control groups (Lipsey et al., 2012).

In addition, we found that the treatment effect on noncognitive outcomes of problem posing was larger than the effect on the outcome of quality of posed problems and that the latter was larger than the effect on the outcome of quantity of posed problems. However, the between-group variance was not significant.

Several research studies found that noncognitive factors have potential to improve cognitive skills (Frank, 2020; Holmlund & Silva, 2014). Thus, it would be important to further examine changes in participants' cognitive skills on mathematical problem posing as we examine changes in their respective noncognitive skills. Furthermore, while the quantity of posed problems as an outcome measure could reflect posers' fluency in problem posing, it does not mean that posing more problems represents an enhanced problem posing competence, not least because participants can generate lots of problems by changing the values of the variables in the first posed problem (Zhang et al., 2022a).

### The moderators' influence on treatment efficacy

Regarding RQ3, considering moderator influence via stepwise meta-regression analysis, we found that seven moderators—research design, sample level, sample size, number and mode of intervention components, and the existence of particular types of intervention components—explained a statistically significant portion of the heterogeneity of treatment effects between studies. We reflect on the results according to the typology of moderators, including research design, sample characteristics, and intervention characteristics, and we do so separately for each type under the assumption that other moderators remain fixed.

In terms of research design, the overall effect size of pre-post design was higher than that of pre-post-control design. This result was consistent with our original hypothesis, namely, that studies adopting a single-group design would yield higher mean treatment effect size than studies using experimental or quasi-experimental designs. The significance level was set at 0.10, consistent with Niu et al. (2013). As Lakens (2013) explained, the increased statistical power and sensitivity in within-participants designs allow for the detection of smaller, yet significant, treatment effects that might be missed in between-participants designs. However, the significance level being set at 0.10 suggests that, while our findings support the hypothesis, the evidence is not as robust as it could be. It is crucial to balance the increased power and the potential validity threats inherent in within-participants designs. Careful consideration of these trade-offs helps enhance the robustness and reliability of effect size estimates in intervention studies.

In terms of sample characteristics, the results indicated that the *sample level* significantly influenced treatment efficacy regardless of outliers. Specifically, the treatment efficacy of studies delivered to K-12 students was significantly higher than that to prospective teachers and in-service teachers. This result matched our hypothesis that lower grade level participants might benefit more from interventions targeting their mathematical problem posing competence. Higher grade level participants, who are more accustomed to the conventional teacher-led instruction and are relatively successful in learning mathematics in this way of teaching, are more likely to possess low motivation in posing problems (Silver, 1994) and thus benefit less from the interventions. Furthermore, Cai and Hwang (2020) delved into the nuances of problem posing from a pedagogical standpoint, highlighting the difference between students and teachers. For teachers, problem posing extends beyond merely generating problems based on given problem situations or modifying existing problems, which are the areas students are solely focusing on. Teachers might also consider activities like predicting problems students might pose, generating situations for students to pose problems, and posing problems for students to solve. This distinction underscores that the impact of a problem posing intervention could vary depending on the roles of the participants. As Voica and Pelcer (2009) noted, individuals without role-specific constraints, who focus purely on the mathematical aspects of problem posing, might have a better performance. To conclude, although we recognize that problem posing has different pedagogical implications for students and teachers, this distinction can shed some light on the significant variance in treatment efficacy driven by the sample level. It also motivates further investigation into the reasons behind this phenomenon and, in particular, how interventions might uniquely resonate with participants with or without role-specific constraints.

In addition, we found that studies implemented with small groups of students (less than 25) produced relatively larger effects than those implemented with large groups (more than 50). This result is consistent with our original hypothesis about the role of group size and has received support in the literature from at least two perspectives. From a statistical perspective, effect sizes based on small samples were found to be larger than effect sizes based on larger samples even when the actual magnitudes of the intervention effects were identical (Lipsey et al., 2012). From an instructional perspective, significant gains made by low-performing students were attributed in part to the number of hours spent in, as well as the intensity

of, the intervention (Torgesen, 2000), and this intensity was likely to be higher when the intervention was delivered to a smaller group. Relatedly, as mathematical problem posing is a relatively new activity compared to problem solving for many participants, participants are more likely to have a low performance at the beginning and thus require more assistance through small-group or individual instruction (Cirino et al., 2015; Powell et al., 2009).

In terms of intervention characteristics, we found that interventions that were designed with more than one intervention component or incorporated one particular type of intervention component were associated with significantly larger effects than those conducted with a single intervention component or incorporated mixed types of intervention components. These results suggest a complex role of design multiplicity of intervention components, appearing to favor multiplicity within but not across type of intervention components. A commonly held view is that interventions with more than one type of components are more effective than single-type-component interventions (Squires et al., 2014), since there are multiple barriers at different levels to changing participants' behaviors (Grimshaw et al., 2012). In the case of problem posing, participants may lack prior experience with problem posing activity including strategies of how to pose problems. Accordingly, it is reasonable to expect that multifaceted interventions that target several of these barriers simultaneously using a mixture of types of intervention components (e.g., activity-based practice combined with method-based assistance) may be more effective to address the barriers to a behavior. Yet, despite the face validity of this point and our results that support the opposite, evidence as to whether multifaceted interventions are truly more effective remains uncertain (Squires et al., 2014). As the field explores this matter more, it is useful to note our finding that interventions that employed the method-based assistance or environment-based support produced a higher effect size than interventions that applied no such type of assistance or support.

Finally, regarding another key moderator with respect to intervention characteristics, namely, intervention duration, there was not enough evidence that duration was a moderator since none of the coefficients of duration in the regression models were statistically significant. On one hand, interventions of relatively long duration might have offered participants additional opportunities to receive explicit modeling and practice to develop their skills, as well as opportunities for ongoing progress monitoring and feedback (Powell & Fuchs, 2015). On the other hand, interventions of relatively short

duration might have been implemented with higher fidelity (Stylianides & Stylianides, 2013). The way intervention duration was calculated in this review is also worth consideration. Specifically, we used intervention duration to refer to the length of time over which an intervention was implemented or spread across, rather than the length of time that participants actually experienced the intervention components. This way of calculating duration might not reflect accurately the intensity of the intervention as it might include time periods when participants did not receive an intervention treatment. This, in turn, could result in a dilution of the overall intervention effect, making it more difficult to detect a significant relationship between duration and treatment efficacy.

### Limitations and future meta-analyses

Despite our best efforts to identify relevant publications, we were unable to access several potential studies. We contacted authors to obtain these articles, but we received no response on some occasions. Also, we did not take account of publications presented at conferences due to concerns about inconsistent standards of peer review and the relatively short length of articles in conference proceedings that may not allow authors to present in sufficient detail their research designs and findings. Furthermore, although our statistical analysis did not indicate publication bias, the tendency of journals to publish studies with significant effects, combined with the large proportion of such studies in our meta-analysis, suggests that publication bias may still have influenced our findings. Finally, although the meta-analysis considered several kinds of potential moderators, the moderator analysis showed considerable between-study heterogeneity, suggesting other factors not accounted for in this analysis might affect the effect sizes. Future meta-analyses exploring additional potential moderators, such as the dependent measure type (researcher-developed assessment or standardized measures) and study quality rating, are needed for deepening our understanding of the factors that impact an intervention's effectiveness.

### Conclusions

Although mathematical problem posing is a younger field of inquiry within mathematics education compared to its twin activity of mathematical problem solving, it has attracted increased research attention over the recent years and, gradually, an important theoretical and research foundation has been established in relation to both (e.g., Cai et al., 2015; Silver, 2023; Toh et al., 2023). Our findings in this review of interventions to improve participants' mathematical problem posing competence, including the mechanisms underlying the more

or less effective interventions and moderators' influence on intervention efficacy, help deepen theoretical understanding of this competence and how to promote it (Bronfendrenner, 1977; Snyder et al., 2019).

The findings provide researchers and practitioners with useful guidance about how to design and implement (more) effective interventions to enhance students' and teachers' mathematical problem posing competence and, through this, participants' other important skills that are believed to be associated with problem posing competence, notably, mathematical problem solving and creativity skills (Bonotto & Santo, 2015; Cai & Hwang, 2002). As we explained previously, the effectiveness of the interventions we reviewed differed across intervention designs. In particular, the number and mode of intervention components, along with the existence of certain types of intervention components, emerged as significant factors influencing the treatment efficacy. Researchers and practitioners who design new interventions can select and tailor appropriate intervention component dosage or types in order to optimize the treatment efficacy, while considering their particular aims, contextual factors, and participant needs. However, it is important to recognize that intervention implementation is a dynamic, context-specific process. Each layer of a context, whether at the micro (individual), meso (team or organization), or macro (system) level, can affect the intervention's effectiveness (Moullin et al., 2020). Thus, ongoing tailoring of intervention design, as well as formative and summative evaluations of factors at any of these levels, is necessary to comprehensively evaluate mechanisms of intervention success.

The demonstrated effectiveness and uncovered mechanisms of existing interventions in this study highlight the feasibility of integrating problem posing into real-world educational settings and mathematics curricula. Teachers can translate the identified intervention components into classroom practices in teaching mathematics both *for* and *through* problem posing (Silver, 2023). To support these practices, professional development programs to equip teachers with the necessary knowledge and skills are sorely needed, including how to integrate problem posing and problem solving activities in mathematics curricula and classrooms (Toh et al., 2023). Through the collaborative effort of researchers, practitioners, and policymakers, the theoretical and practical advancements in mathematical problem posing can be translated into tangible educational improvements.

### Appendix

See Table 2.

**Table 2** Summary of the studies included in the meta-analysis (N = 26)

Research ID	Journal/chapter	Country/district	Details of Intervention	Period	Samples <sup>a</sup>	Component
English (1997)	Educ. Stud. Math	Australia	<ul style="list-style-type: none"> <li>- Problem exploration and problem sorting</li> <li>- Analysis problem structure and model new problems</li> <li>- Create new problems from given problem components</li> <li>- Transform a given problem into a new problem</li> </ul>	10 weeks, 35 min per week	27 fifth graders	EPP PPA CPPS ILE SPP
Cai and Hwang (2021)	ZDM—Math. Educ	USA	<ul style="list-style-type: none"> <li>- Discuss the nature of PP</li> <li>- Engage in PP activities</li> <li>- Predict the problems students might pose</li> <li>- Discuss and modify lesson plans to integrate PP activities</li> <li>- Consider PP lessons designed by other teachers</li> <li>- Create and implement the PP teaching</li> </ul>	4–8 workshops, up to whole year	103 primary/secondary teachers	WPP PPA ILE
English (1998)	J. Res. Math. Educ	Australia	<ul style="list-style-type: none"> <li>PP activities incorporated in each session</li> <li>- Explain the nature of PP</li> <li>- Interpret the problem situation</li> <li>- Engage in PP activities</li> <li>- Share ideas on other problems that could be created</li> </ul>	Sixteen 45-min sessions, 2 per week	54 8.1 years old students	WPP PPA CPPS ILE
Leavy and Hourigan (2020)	J. Math. Teach. Educ	Ireland	<ul style="list-style-type: none"> <li>- Give lectures underpinning PP and problem solving</li> <li>- Engage in series of PP and problem solving activities</li> </ul>	3 weeks	415 prospective primary teachers	WPP PPA ILE
Li et al., (2020a, 2020b)	Educ. Stud. Math	China	<ul style="list-style-type: none"> <li>Four-cycle workshops</li> <li>- Overview of PP</li> <li>- Engage in PP activities</li> <li>- Predict the problem students might pose</li> <li>- Write and modify the lessons plan integrating PP</li> <li>- Discuss lesson plans taught by other teachers</li> </ul>	Four 3-day workshop	74 primary teachers	WPP WGP PPA ILE



**Table 2** (continued)

Research ID	Journal/chapter	Country/district	Details of Intervention	Period	Samples <sup>a</sup>	Component
Chen et al. (2015)	Chapter	China	<ul style="list-style-type: none"> <li>- Explore the concept of PP</li> <li>- Explore the assessment criteria of PP</li> <li>- Generate new problem from PP tasks</li> <li>- Transform a given problem into new problems</li> <li>- Practice on PP</li> </ul>	11 units, 90 min per unit, one unit per week	69 fourth graders	WPP WGP PPA EPP SPP CPPS
Cai et al. (2020)	Int. J. Educ. Res	USA	<ul style="list-style-type: none"> <li>- Discuss of the nature of PP</li> <li>- Engage in PP activities</li> <li>- Predict the problems students might pose</li> <li>- Discuss and modify their lesson plans to integrate PP activities</li> <li>- Discuss PP lessons designed by other teachers</li> </ul>	3-day workshop	50 primary teachers	WPP WGP PPA ILE
Ayvaz and Durmus (2021)	Think. Skills Creat	Turkey	<p>A 30-h action plan consisting of 18 problem posing activities</p> <ul style="list-style-type: none"> <li>- First action: 9 structured PP activities</li> <li>- Second action: 5 semi-structured PP activities</li> <li>- Third action: 4 free PP activities</li> </ul>	30 h	Six 7th grade gifted students	PPA
Silver et al. (1996)	J. Res. Math. Educ	USA	<p>Work by individual or pairs</p> <ul style="list-style-type: none"> <li>- Initially pose problems (before problem solving)</li> <li>- Pose problems (during problem solving)</li> <li>- Additionally posing problems (after problem solving)</li> </ul>	45 min	53 secondary teachers; 28 prospective teachers <sup>a</sup>	PPA ILE
Crespo and Sinclair (2008)	J. Math. Teach. Educ	USA	<ul style="list-style-type: none"> <li>- Explore the situation first VS pose spontaneously</li> <li>- Evaluate their own posed problems</li> <li>- Discuss what makes a "good" or "interesting" problems</li> </ul>	3 weeks	22 prospective teachers <sup>a</sup>	PPA CPPS EPP WGP ILE

**Table 2** (continued)

Research ID	Journal/chapter	Country/district	Details of Intervention	Period	Samples <sup>a</sup>	Component
Cankoy (2014)	Int. J. Sci. Math. Educ	Cyprus	<ul style="list-style-type: none"> <li>- Introduce a four-step problem solving procedure</li> <li>- Solve a set of text-book problems</li> <li>- Write problems for friends in free structured situations</li> <li>- Discuss about the quality of selected posed problems</li> </ul>	5 weeks	30 fifth graders	PPA ILE
Grundmeier (2015)	Chapter	USA	<ul style="list-style-type: none"> <li>- Engage in problem reformulation tasks</li> <li>- Engage in problem generation tasks</li> </ul>	1 h and 50 min	19 prospective teachers <sup>a</sup>	PPA
Putra et al. (2020)	Int. J. Instr	Indonesia	<ul style="list-style-type: none"> <li>- Experimental Class obtained scientific approach with what-if-not strategy</li> <li>- Control class obtained conventional Teaching</li> </ul>	Less than 1 day	68 ninth graders	SPP
Leavy and Hourigan (2022)	J. Math. Teach. Educ	Ireland	<ul style="list-style-type: none"> <li>- Overview of PP and problem solving</li> <li>- Co-design problems</li> <li>- Analysis of penpal responses</li> <li>- Circle week with problem design</li> <li>- Solve a selection of mathematical problems</li> <li>- Discuss the feedback on problem design</li> <li>- Feedback and reflection</li> </ul>	12 weeks	28 prospective primary teachers	WPP PPA EPP ILE
Lowrie (2002)	Math. Educ. Res. J	Australia	<ul style="list-style-type: none"> <li>Student teachers' support PP sessions</li> <li>- Pose a problem</li> <li>- Discuss approaches that would be required to solve the problem</li> <li>- Solve the problem</li> <li>- Reflect upon the manner in which they solved the problem</li> <li>- Gain insights into the children's perception of PP</li> <li>- Assist and encourage students to formulate problems</li> </ul>	1 h per week for 5 weeks	25 first graders; 28 third graders; 53 undergraduates	PPA EPP ILE

**Table 2** (continued)

Research ID	Journal/chapter	Country/district	Details of Intervention	Period	Samples <sup>a</sup>	Component
Kojima and Miwa (2008)	Int. J. Artif. Intell. Educ	Japan	<ul style="list-style-type: none"> <li>- Present example problem A and ask to solve it</li> <li>- Pose a problem from example problem A</li> <li>- Present with three cases problems and ask to solve it</li> <li>- Pose another problem from example problem A</li> </ul>	–	112 undergraduates	PPE
Kojima et al. (2013)	Int. J. Artif. Intell. Educ	Japan	<ul style="list-style-type: none"> <li>- Training and learning in a system operation</li> <li>- Present the base and example problems to each condition groups</li> <li>- Pose problem by using system</li> </ul>	45 min	40 undergraduates	PPE TPP
Kojima et al. (2015)	Res. Pract. Technol. Enhanc. Learn	Japan	<ul style="list-style-type: none"> <li>- Instruct how to pose a novel problem from a base by a learning task</li> <li>- Reproduce problems according to examples</li> <li>- Evaluate the case problem</li> </ul>	45 min	167 undergraduates	PPE TPP
Daher and Anabousy (2018)	Chapter	Israel	<ul style="list-style-type: none"> <li>Groups differed in using of technology (applet) and use of what-if-not strategy</li> <li>- Introduce the technology or what-if-not strategy</li> <li>- Pose problem by using technology or strategy</li> </ul>	–	79 prospective secondary teachers	TPP SPP
Cheng et al. (2014)	J. Comput. Educ	Taiwan	<ul style="list-style-type: none"> <li>- Experimental group's PP by annotated expression</li> <li>- Control group's PP by pure expressions</li> </ul>	3 weeks	29 fourth graders	PPA
Chang et al. (2012)	Comput. Educ	Taiwan	<ul style="list-style-type: none"> <li>- Try out the PP system</li> <li>- Pose problem (system based or paper based)</li> <li>- Verify and revise posed problem (system based or paper based)</li> <li>- Solve posed problem (system based or paper based)</li> <li>- Create new problems according to feedback (system based or paper based)</li> </ul>	160 min	45 fifth graders	TPP

**Table 2** (continued)

Research ID	Journal/chapter	Country/district	Details of Intervention	Period	Samples <sup>a</sup>	Component
Tavşanlı et al. (2018)	Pegem. J. Educ. Instr	Turkey	- PP activities with/without graphic organizers	4 weeks	38 third graders	TPP
Divrik et al. (2020)	Int. Electron. J. Elem. Educ	Turkey	- Inquiry-based learning method with PP activities practice -PP/PS homework guide form supported by metacognitive strategies	9 weeks	63 fourth graders	PPA SPP
Nakano et al. (2002)	Chapter	Japan	An Intelligent learning environment (ILE) for PP - Diagnoses the problems posed by the learner - Helps the learner to correct wrong problems - Leads the learner in the next step of PP	40 min	55 fourth graders	ILE
Kalmpourtzis (2019)	Br. J. Educ. Technol	Greece	PP incorporated in the game design activities	3 months	34 kindergarten students	PPA
Hsiao et al. (2013)	Turkish. Online. J. Educ. Technol	Taiwan	PP homework along with/without worked examples	6 weeks	107 undergraduates	PPE

<sup>a</sup> Information was not available regarding whether the participants were prospective primary or secondary teachers

### Acknowledgements

The authors are grateful to the Editor and five anonymous reviewers for their helpful comments and feedback on earlier versions of this paper. The authors would also like to thank Shengya Wang and Murou Li for their assistance with the literature search and data coding as research assistants on this project.

### Author contributions

All authors contributed equally to the preparation of this paper.

### Funding

This study was supported by a Chongqing Postdoctoral Fellowship (No. 2020379) to the first author, hosted at the University of Cambridge under the co-mentorship of the other authors.

### Data availability

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

### Declarations

### Competing interests

The authors declare no competing interests.

Received: 15 November 2023 Accepted: 10 September 2024

Published online: 28 September 2024

### References

Abramovich, S., & Cho, E. K. (2015). Using digital technology for mathematical problem posing. In F. M. Singer, N. Ellerton, & J. Cai (Eds.), *Mathematical*

*problem posing: From research to effective practice* (pp. 72–89). Springer. [https://doi.org/10.1007/978-1-4614-6258-3\\_4](https://doi.org/10.1007/978-1-4614-6258-3_4)

- Abu-Elwan, R. (2007). The use of Webquest to enhance the mathematical problem-posing skills of pre-service teachers. *International Journal for Technology in Mathematics Education*, 14(1), 31–39.
- Aydin, H., & Monaghan, J. (2018). Encouraging students' problem posing through importing visual images into mathematical software. *Teaching Mathematics and Its Applications*, 37, 141–154. <https://doi.org/10.1093/teamat/hrx005>
- Ayvaz, U., & Durmus, S. (2021). Fostering mathematical creativity with problem posing activities: An action research with gifted students. *Thinking Skills and Creativity*, 40, 100846. <https://doi.org/10.1016/j.tsc.2021.100846>
- Balduzzi, S., Rucker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence-Based Mental Health*, 22(4), 153–160. <https://doi.org/10.1136/ebmental-2019-300117>
- Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., Wallet, P. A., Fiset, M., & Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research*, 74(3), 379–439. <https://doi.org/10.3102/00346543074003379>
- Bevan, D., Williams, A. M., & Capraro, M. M. (2019). Strike a pose: The impact of problem-posing on elementary students' mathematical attitudes and achievement. In N. Jarmila & M. Hana (Eds.), *International symposium elementary mathematics teaching* (pp. 80–88). Prague.
- Bicer, A. (2021). A systematic literature review: Discipline-specific and general instructional practices fostering the mathematical creativity of students. *International Journal of Education in Mathematics, Science and Technology*, 9(2), 252–281. <https://doi.org/10.46328/ij-emst.1254>
- Bicer, A., Lee, Y., Perihan, C., Capraro, M. M., & Capraro, R. M. (2020). Considering mathematical creative self-efficacy with problem posing as a measure of mathematical creativity. *Educational Studies in Mathematics*, 105(3), 457–485. <https://doi.org/10.1007/s10649-020-09995-8>

- Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift Für Psychologie*, 223, 3–13.
- Boller, K., Tarrant, K., & Schaack, D. D. (2014). *Early Care and Education Quality Improvement: A Typology of Intervention Approaches*. OPRE Research Report # 2014-36. Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Service.
- Bonotto, C. (2010). Realistic mathematical modeling and problem posing. In R. Lesh, P. Galbraith, C. Haines, & A. Hurford (Eds.), *Modeling students' mathematical modeling competencies* (pp. 399–408). Boston: Springer. [https://doi.org/10.1007/978-1-4419-0561-1\\_34](https://doi.org/10.1007/978-1-4419-0561-1_34)
- Bonotto, C., & Santo, L. D. (2015). On the relationship between problem posing, problem solving, and creativity in primary school. In F. M. Singer, N. Ellerton, & J. Cai (Eds.), *Mathematical problem posing: From research to effective practice* (pp. 103–124). New York: Springer. [https://doi.org/10.1007/978-1-4614-6258-3\\_5](https://doi.org/10.1007/978-1-4614-6258-3_5)
- Borde, R., Smith, J. J., Sutherland, R., Nathan, N., & Lubans, D. R. (2017). Methodological considerations and impact of school-based interventions on objectively measured physical activity in adolescents: A systematic review and meta-analysis. *Obesity Reviews*, 18(4), 476–490. <https://doi.org/10.1111/obr.12517>
- Bronfenbrenner, U. (1977). Toward an experimental ecology by human development. *American Psychology*, 32, 513–531. <https://doi.org/10.1037/0003-066X.32.7.513>
- Brown, S. I., & Walter, M. I. (1993). Problem posing in mathematics education. In S. I. Brown & M. I. Walter (Eds.), *Problem posing: Reflections and application* (pp. 16–27). Erlbaum.
- Cai, J. (2022). What research says about teaching mathematics through problem posing. *Education & Didactique*, 16, 31–50.
- Cai, J., Chen, T., Li, X., Xu, R., Zhang, S., Hu, Y., Zhang, L., & Song, N. (2020). Exploring the impact of a problem-posing workshop on elementary school mathematics teachers' conceptions on problem posing and lesson design. *International Journal of Educational Research*, 102, 101404. <https://doi.org/10.1016/j.ijer.2019.02.004>
- Cai, J., & Hwang, S. (2002). Generalized and generative thinking in U.S. and Chinese students' mathematical problem solving and problem posing. *Journal of Mathematical Behavior*, 21(4), 401–421. [https://doi.org/10.1016/S0732-3123\(02\)00142-6](https://doi.org/10.1016/S0732-3123(02)00142-6)
- Cai, J., & Hwang, S. (2020). Learning to teach through mathematical problem posing: Theoretical considerations, methodology, and directions for future research. *International Journal of Educational Research*, 102, 101391. <https://doi.org/10.1016/j.ijer.2019.01.001>
- Cai, J., & Hwang, S. (2021). Teachers as redesigners of curriculum to teach mathematics through problem posing: Conceptualization and initial findings of a problem posing project. *ZDM*, 53(6), 1403–1416. <https://doi.org/10.1007/s11858-021-01252-3>
- Cai, J., Hwang, S., Jiang, C., & Silber, S. (2015). Problem posing research in mathematics education: Some answered and unanswered questions. In F. M. Singer, N. Ellerton, & J. Cai (Eds.), *Mathematical problem posing: From research to effective practice* (pp. 3–34). Springer. [https://doi.org/10.1007/978-1-4614-6258-3\\_1](https://doi.org/10.1007/978-1-4614-6258-3_1)
- Cai, J., & Leikin, R. (2020). Affect in mathematical problem posing: Conceptualization, advances, and future directions for research. *Educational Studies in Mathematics*, 105, 287–301. <https://doi.org/10.1007/s10649-020-10008-x>
- Cai, J., & Rott, B. (2024). On understanding mathematical problem-posing processes. *ZDM*, 56(1), 61–71. <https://doi.org/10.1007/s11858-023-01536-w>
- Cankoy, O. (2014). Interlocked problem posing and children's problem posing performance in free structured situation. *International Journal of Science and Mathematics Education*, 12, 219–238. <https://doi.org/10.1007/s10763-013-9433-9>
- Carrara, E., Pfeffer, I., Zusman, O., Leibovici, L., & Paul, M. (2018). Determinants of inappropriate empirical antibiotic treatment: Systematic review and meta-analysis. *International Journal of Antimicrobial Agents*, 51(4), 548–553. <https://doi.org/10.1016/j.ijantimicag.2017.12.013>
- Chang, K., Wu, L., Weng, S., & Sung, Y. (2012). Embedding game-based problem solving into problem posing system for mathematics learning. *Computers & Education*, 58, 775–786. <https://doi.org/10.1016/j.compedu.2011.10.002>
- Chauhan, S. (2017). A meta-analysis of the impact of technology on learning effectiveness of elementary students. *Computers & Education*, 105, 14–30. <https://doi.org/10.1016/j.compedu.2016.11.005>
- Chen, L., Van Dooren, W., & Verschaffel, L. (2015). Enhancing the development of Chinese fifth-graders' problem-posing and problem-solving abilities, beliefs, and attitudes: A design experiment. In F. M. Singer, N. Ellerton, & J. Cai (Eds.), *Mathematical problem posing: From research to effective practice* (pp. 309–329). Springer. [https://doi.org/10.1007/978-1-4614-6258-3\\_15](https://doi.org/10.1007/978-1-4614-6258-3_15)
- Cheng, H. N. H., Weng, Y., & Chan, T. (2014). Computer-supported problem posing by annotated expressions: Content-first design and evaluation. *Journal of Computer Education*, 1(4), 271–294. <https://doi.org/10.1007/s40692-014-0019-5>
- Chinese Ministry of Education. (2022). *Quanrizhi yiwu jiaoyu shuxue kecheng biaoazhun [Mathematics Curriculum Standard of compulsory education (2022 version)]*. People's Education Press.
- Cirino, P. T., Fuchs, L. S., Elias, J. T., Powell, S. R., & Schumacher, R. F. (2015). Cognitive and mathematical profiles for different forms of learning difficulties. *Journal of Learning Disabilities*, 48(2), 156–175. <https://doi.org/10.1177/0022219413494239>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.
- Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, 1(3), 98–101. <https://doi.org/10.1111/1467-8721.ep10768783>
- Contreras, J. (2007). Unraveling the mystery of the origin of mathematical problems: Using a problem-posing framework with prospective mathematics teachers. *The Mathematics Educator*, 17(2), 15–23.
- Courtney, S. A., Caniglia, J., & Singh, R. (2014). Investigating the impact of field trips on teachers' mathematical problem posing. *Journal of Experiential Education*, 37(2), 144–159. <https://doi.org/10.1177/1053825913498369>
- Crespo, S. (2003). Learning to pose mathematical problems: Exploring changes in preservice teachers' practices. *Educational Studies in Mathematics*, 52, 243–270. <https://doi.org/10.1023/A:1024364304664>
- Crespo, S., & Sinclair, N. (2008). What makes a problem mathematically interesting? Inviting prospective teachers to pose better problems. *Journal of Mathematics Teacher Education*, 11(5), 395–415. <https://doi.org/10.1023/A:1024364304664>
- Daher, W., & Anabousy, A. (2018). Flexibility of pre-services teachers in problem posing in different environments. In F. M. Singer (Ed.) *Mathematical Creativity and Mathematical Giftedness*, ICME-13 Monographs. [https://doi.org/10.1007/978-3-319-73156-8\\_9](https://doi.org/10.1007/978-3-319-73156-8_9)
- Damschroder, L. J., & Hagedorn, H. J. (2011). A guiding framework and approach for implementation research in substance use disorders treatment. *Psychology of Addictive Behaviors*, 25, 194–205. <https://doi.org/10.1037/a0022284>
- Danusso, L., Testa, I., & Vicentini, M. (2010). Improving prospective teachers' knowledge about scientific models and modelling: Design and evaluation of a teacher education intervention. *International Journal of Science Education*, 32(7), 871–905. <https://doi.org/10.1080/09500690902833221>
- Darhim, T., Darhim, D., Juandi, D., Gardenia, N., & Kandaga, T. (2021). High school students' attitudes towards mathematics lessons using the PQ4R strategy and problem posing mathematical problems. *Journal Pendidikan Matematika*, 6(2), 171–180. <https://doi.org/10.22236/KALAMATIKA.vol6no2.2021pp171-180>
- Depaepe, F., Verschaffel, L., & Kelchtermans, G. (2013). Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education*, 34, 12–25. <https://doi.org/10.1016/j.tate.2013.03.001>
- Divrik, R., Pilten, P., & Tas, A. M. (2020). Effect of inquiry-based learning method supported by metacognitive strategies on fourth-grade students' problem-solving and problem-posing skills: A mixed methods research. *International Electronic Journal of Elementary Education*, 13(2), 287–308. <https://doi.org/10.26822/iejee.2021.191>
- 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>
- Einstein, A., & Infeld, L. (1938). *The evolution of physics*. Simon & Schuster.
- Ellerton, N. F. (2013). Engaging pre-service middle-school teacher-education students in mathematical problem posing: Development of an active

- learning framework. *Educational Studies in Mathematics*, 83(1), 87–101. <https://doi.org/10.1007/s10649-012-9449-z>
- Ellerton, N. F., & Clarkson, P. C. (1996). Language factors in mathematics teaching and learning. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 987–1033). Kluwer. [https://doi.org/10.1007/978-94-009-1465-0\\_27](https://doi.org/10.1007/978-94-009-1465-0_27)
- English, L. D. (1997). The development of fifth-grade children's problem-posing abilities. *Educational Studies in Mathematics*, 34(3), 183–217. <https://doi.org/10.1023/A:1002963618035>
- English, L. D. (1998). Children's problem posing within formal and informal contexts. *Journal for Research in Mathematics Education*, 29(1), 83–106. <https://doi.org/10.5951/jresmetheduc.29.1.0083>
- Frank, J. L. (2020). School-based practices for the 21st century: Noncognitive factors in students learning and psychosocial outcomes. *Policy Insight from the Behavioral and Brain Sciences*, 7(1), 44–51. <https://doi.org/10.1177/2372732219898703>
- Garzón, J., Baldiris, S., Gutiérrez, J., & Pavón, J. (2020). How do pedagogical approaches affect the impact of augmented reality on education? A meta-analysis and research synthesis. *Educational Research Review*, 31, 100334. <https://doi.org/10.1016/j.edurev.2020.100334>
- Grimshaw, J. M., Eccles, M. P., Lavis, J. N., Hill, S. J., & Squires, J. E. (2012). Knowledge translation of research findings. *Implementation Science*, 7(1), 1–17. <https://doi.org/10.1186/1748-5908-7-50>
- Grundmeier, T. A. (2015). Developing the problem posing abilities of prospective elementary and middle school teachers. In F. M. Singer, N. Ellerton, & J. Cai (Eds.), *Mathematical problem posing: From research to effective practice* (pp. 412–417). New York: Springer. [https://doi.org/10.1007/978-1-4614-6258-3\\_20](https://doi.org/10.1007/978-1-4614-6258-3_20)
- Hadamard, J. (1945). *The psychology of invention in the mathematical field*. Dover.
- Harden, A., & Thomas, J. (2005). Methodological issues in combining diverse study types in systematic reviews. *International Journal of Social Research Methodology*, 8(3), 257–271. <https://doi.org/10.1080/13645570500155078>
- Harwell, M., & Maeda, Y. (2008). Deficiencies of reporting in meta-analyses and some remedies. *Journal of Experimental Education*, 76(4), 403–430. <https://doi.org/10.3200/JEXE.76.4.403-430>
- Hawkins, D. (2000). *The roots of literacy*. University Press of Colorado.
- Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin*, 92(2), 490–499. <https://doi.org/10.1037/0033-2909.92.2.490>
- Holmlund, H., & Silva, O. (2014). Targeting noncognitive skills to improve cognitive outcomes: Evidence from a remedial education intervention. *Journal of Human Capital*, 8(2), 126–160. <https://doi.org/10.1086/676460>
- Hsiao, J. Y., Hung, C. L., Lan, Y. F., & Jeng, Y. C. (2013). Integrating worked examples into problem posing in a web-based learning environment. *Turkish Online Journal of Educational Technology-TOJET*, 12(2), 166–176.
- Joaquin, M. N. B. (2023). Problem posing among preservice and inservice mathematics teachers. In T. L. Toh, M. Santos-Trigo, P. H. Chua, N. A. Abdullah, & D. Zhang (Eds.), *Problem posing and problem solving in mathematics education: International research and practice trends* (pp. 173–187). Springer.
- Kalmpourtzis, G. (2019). Connecting game design with problem posing skills in early childhood. *British Journal of Educational Technology*, 50(2), 846–860. <https://doi.org/10.1111/bjet.12607>
- Kilpatrick, J. (1987). Problem formulating: Where do good problems come from? In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 123–147). Erlbaum.
- Klinshtern, M., Koichu, B., & Berman, A. (2015). What do high school teachers mean by saying "I pose my own problem"? In F. M. Singer, N. Ellerton, & J. Cai (Eds.), *Mathematical problem posing: From research to effective practice* (pp. 449–467). Springer. [https://doi.org/10.1007/978-1-4614-6258-3\\_22](https://doi.org/10.1007/978-1-4614-6258-3_22)
- Kojima, K., & Miwa, K. (2008). A system that facilitates diverse thinking in problem posing. *International Journal of Artificial Intelligence in Education*, 18(3), 209–236.
- Kojima, K., Miwa, K., & Matsui, T. (2013). Supporting mathematical problem posing with a system for learning generation processes through examples. *International Journal of Artificial Intelligence in Education*, 22, 161–190. <https://doi.org/10.3233/JAI-130035>
- Kojima, K., Miwa, K., & Matsui, T. (2015). Experimental study of learning support through examples in mathematical problem posing. *Research and Practice in Technology Enhanced Learning*, 10, 1–18. <https://doi.org/10.1007/s41039-015-0001-5>
- Kontorovich, I. (2020). Problem-posing triggers or where do mathematics competition problems come from? *Educational Studies in Mathematics*, 105, 389–406. <https://doi.org/10.1007/s10649-020-09964-1>
- Kopparla, M., Bicer, A., Vela, K., Lee, Y., Bevan, D., Kwon, H., Caldwell, C., Capraro, M. M., & Capraro, R. M. (2019). The effects of problem-posing intervention types on elementary students' problem-solving. *Educational Studies*, 45(6), 708–725. <https://doi.org/10.1080/03055698.2018.1509785>
- Kul, Ü., & Çelik, S. (2020). A meta-analysis of the impact of problem posing strategies on student's learning of mathematics. *Revista Românească Pentru Educație Multidimensională*, 12(3), 341–368. <https://doi.org/10.18662/rrem/12.3/325>
- Kwon, H., & Capraro, M. M. (2018). The effects of using manipulatives on students' learning in problem posing: The instructors' perspectives. *Journal of Mathematics Education*, 11(2), 35–47. <https://doi.org/10.26711/007577152790026>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 863.
- Lavy, I., & Bershadsky, I. (2003). Problem posing via "what if not?" strategy in solid geometry—a case study. *Journal of Mathematical Behavior*, 22(4), 369–387. <https://doi.org/10.1016/j.jmathb.2003.09.007>
- Leavy, A., & Hourigan, M. (2020). Posing mathematically worthwhile problems: Developing the problem posing skills of prospective teachers. *Journal of Mathematics Teacher Education*, 23, 341–361. <https://doi.org/10.1007/s10857-018-09425-w>
- Leavy, A., & Hourigan, M. (2022). Balancing competing demands: Enhancing the mathematical problem posing skills of prospective teachers through a mathematical letter writing initiative. *Journal of Mathematics Teacher Education*, 25(3), 293–320. <https://doi.org/10.1007/s10857-021-09490-8>
- Li, X., Dusseldorp, E., Su, X., & Meulman, J. J. (2020a). Multiple moderator meta-analysis using the R-package Meta-CART. *Behavior Research Methods*, 52(6), 2657–2673. <https://doi.org/10.3758/s13428-020-01360-0>
- Li, X., Song, N., Hwang, S., & Cai, J. (2020b). Learning to teach mathematics through problem posing: Teachers' beliefs and performance on problem posing. *Educational Studies in Mathematics*, 105, 325–347. <https://doi.org/10.1007/s10649-020-09981-0>
- Liljedahl, P., & Cai, J. (2021). Empirical research on problem solving and problem posing: a look at the state of the art. *ZDM*, 53, 723–735. <https://doi.org/10.1007/s11858-021-01291-w>
- Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S., & Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms*. National Center for Special Education Research.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE Publications.
- Liu, Y., & Pásztor, A. (2022). Effects of problem-based learning instructional intervention on critical thinking in higher education: A meta-analysis. *Thinking Skills and Creativity*, 45, 101069. <https://doi.org/10.1016/j.tsc.2022.101069>
- Lo, C. K., Hew, K. F., & Chen, G. (2017). Toward a set of design principles for mathematics flipped classrooms: A synthesis of research in mathematics education. *Educational Research Review*, 22, 50–73. <https://doi.org/10.1016/j.edurev.2017.08.002>
- Lowrie, T. (2002). Young children posing problems: The influence of teacher intervention on the type of problems children pose. *Mathematics Education Research Journal*, 14(2), 87–98. <https://doi.org/10.1007/BF03217355>
- Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: a model comparison perspective*. Lawrence Erlbaum Associate.
- Milinković, J. (2015). Conceptualizing problem posing via transformation. In F. M. Singer, N. El-lerton, & J. Cai (Eds.), *Mathematical problem posing: From research to effective practice* (pp. 47–70). Springer. [https://doi.org/10.1007/978-1-4614-6258-3\\_3](https://doi.org/10.1007/978-1-4614-6258-3_3)
- Ministry of Education of Italy. (2007). *Indicazioni per il curricolo*. Ministero della Pubblica Istruzione.

- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7(1), 105. <https://doi.org/10.1037/1082-989X.7.1.105>
- Moullin, J. C., Diskson, K. S., Stadnick, N. A., Albers, B., Nilsen, P., Broder-Figert, S., Mukasa, B., & Aarons, G. A. (2020). Ten recommendations for using implementation frameworks in research and practice. *Implementation Science Communications*, 1, 1–12. <https://doi.org/10.1186/s43058-020-00023-7>
- Myers, J. A., Witzel, B. S., Powell, S. R., Li, H., Pigott, T. D., Xin, Y. P., & Hughes, E. M. (2022). A meta-analysis of mathematics word-problem solving interventions for elementary students who evidence mathematics difficulties. *Review of Educational Research*, 92(5), 695–742. <https://doi.org/10.3102/00346543211070049>
- Nakano, A., Hirashima, T., & Takeuchi, A. (2002). An evaluation of intelligent learning environment for problem posing. In S. A. Cerri, G. Gouardères, & F. Paraguaçu (Eds.), *Intelligent Tutoring Systems. ITS2002. Lecture notes in computer science*, Springer. [https://doi.org/10.1007/3-540-47987-2\\_86](https://doi.org/10.1007/3-540-47987-2_86)
- National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Author.
- Niu, L., Behar-Horenstein, L. S., & Garvan, C. W. (2013). Do instructional interventions influence college students' critical thinking skills? A meta-analysis. *Educational Research Review*, 9, 114–128. <https://doi.org/10.1016/j.edurev.2012.12.002>
- Öçal, M. F., Kar, T., Güler, G., & Ipek, A. S. (2020). Comparison of prospective mathematics teachers' problem posing abilities in paper-pencil test and on dynamic geometry environment in terms of creativity. *REDIMAT—Journal of Research in Mathematics Education*, 9(3), 243–272. <https://doi.org/10.17583/redimat.2020.3879>
- Otun, W. I., & Njoku, O. G. (2020). Developing pre-service mathematics teachers' mathematical problem solving-posing skills through solve-reflect-pose strategy in Lagos state, Nigeria. *Journal of Educational Research in Developing Areas*, 1(2), 140–152. <https://doi.org/10.47434/JEREDA.1.2.2020.140>
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372, n160. <https://doi.org/10.1136/bmj.n160>
- Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*, 61(10), 991–996. <https://doi.org/10.1016/j.jclinepi.2007.11.010>
- Powell, S. R., & Fuchs, L. S. (2015). Intensive intervention in mathematics. *Learning Disabilities Research & Practice*, 30(4), 182–192. <https://doi.org/10.1111/ldrp.12087>
- Powell, S. R., Fuchs, L. S., Fuchs, D., Cirino, P. T., & Fletcher, J. M. (2009). Do word-problem features differentially affect problem difficulty as a function of students' mathematics difficulty with and without reading difficulty? *Journal of Learning Disabilities*, 42, 99–110. <https://doi.org/10.1177/0022219408326211>
- Pressley, M., Graham, S., & Harris, K. (2006). The state of educational intervention research as viewed through the lens of literacy intervention. *British Journal of Educational Psychology*, 76(1), 1–19. <https://doi.org/10.1348/000709905X66035>
- Priest, D. J. (2009). *A problem-posing intervention in the development of problem-solving competence of underachieving, middle-year students* (Ph.D. thesis, Queensland University of Technology). <https://eprints.qut.edu.au/31740/>
- Putra, H. D., Herman, T., & Sumarmo, U. (2020). The impact of scientific approach and what-if-not strategy utilization towards students' mathematical problem posing ability. *International Journal of Instruction*, 13(1), 669–684. <https://doi.org/10.29333/iji.2020.13143a>
- Rosli, R., Capraro, M. M., & Capraro, R. M. (2014). The effect of problem posing on student mathematical learning: A meta-analysis. *International Education Studies*, 7(13), 227–241. <https://doi.org/10.5539/ies.v7n13.p227>
- Schindler, M., & Bakker, A. (2020). Affective field during collaborative problem posing and problem solving: A case study. *Educational Studies in Mathematics*, 105(3), 303–324. <https://doi.org/10.1007/s10649-020-09973-0>
- Sheridan, S. M., Smith, T. E., Moorman Kim, E., Beretvas, S. N., & Park, S. (2019). A meta-analysis of family-school interventions and children's social-emotional functioning: Moderators and components of efficacy. *Review of Educational Research*, 89(2), 296–332. <https://doi.org/10.3102/0034654318825437>
- Silver, E. A. (1994). On mathematical problem posing. *For the Learning of Mathematics*, 14(1), 19–28.
- Silver, E. A. (2023). Conclusion: Mathematics problem posing and problem solving: Some reflections on recent advances and new opportunities. In T. L. Toh, M. Santos-Trigo, P. H. Chua, N. A. Abdullah, & D. Zhang (Eds.), *Problem posing and problem solving in mathematics education: International research and practice trends* (pp. 247–259). Springer.
- Silver, E. A., & Cai, J. (1996). An analysis of arithmetic problem posing by middle school students. *Journal for Research in Mathematics Education*, 27(5), 521–539. <https://doi.org/10.5951/jresmetheduc.27.5.0521>
- Silver, E. A., & Cai, J. (2005). Assessing students' mathematical problem posing. *Teaching Children Mathematics*, 12(3), 129–135. <https://doi.org/10.5951/TCM.12.3.0129>
- Silver, E. A., Mamona-Downs, J., Leung, S. S., & Kenney, P. A. (1996). Posing mathematical problems: An exploratory study. *Journal for Research in Mathematics Education*, 27(3), 293–309. <https://doi.org/10.5951/jresmetheduc.27.3.0293>
- Snyder, K. E., Fong, C. J., Painter, J. K., Pittard, C. M., Barr, S. M., & Patall, E. A. (2019). Interventions for academically underachieving students: A systematic review and meta-analysis. *Educational Research Review*, 28, 100294. <https://doi.org/10.1016/j.edurev.2019.100294>
- Song, F., Hooper, L., & Loke, Y. (2013). Publication bias: What is it? How do we measure it? How do we avoid it? *Open Access Journal of Clinical Trials*, 5, 71–81. <https://doi.org/10.2147/OAJCT.S34419>
- Squires, J. E., Sullivan, K., Eccles, M. P., Worswick, J., & Grimshaw, J. M. (2014). Are multifaceted interventions more effective than single-component interventions in changing health-care professionals' behaviors? An overview of systematic reviews. *Implementation Science*, 9(1), 1–22. <https://doi.org/10.1186/s13012-014-0152-6>
- Stern, J. M., & Simes, R. J. (1997). Publication bias: Evidence of delayed publication in a cohort study of clinical research projects. *BMJ*, 315(7109), 640–645. <https://doi.org/10.1136/bmj.315.7109.640>
- Stevens, E. A., Rodgers, M. A., & Powell, S. R. (2018). Mathematics interventions for upper elementary and secondary students: A meta-analysis of research. *Remedial and Special Education*, 39(6), 327–340. <https://doi.org/10.1177/0741932517731887>
- Stoyanova, E., & Ellerton, N. F. (1996). A framework for research into students' problem posing in school mathematics. In P. Clarkson (Ed.), *Technology in Mathematics Education* (pp. 518–525). Mathematics Education Research Group of Australasia.
- Strauss, A., & Corbin, J. (2008). *Basics of qualitative research: Grounded theory procedures and techniques* (3rd ed.). Sage.
- Stylianides, A. J., & Stylianides, G. J. (2013). Seeking research-grounded solutions to problems of practice: Classroom-based interventions in mathematics education. *ZDM—Mathematics Education*, 45, 333–341. <https://doi.org/10.1007/s11858-013-0501-y>
- Stylianides, G. J., Stylianides, A. J., & Moutsios-Rentzos, A. (2024). Proof and proving in school and university mathematics education research: a systematic review. *ZDM—Mathematics Education*, 56(1), 47–59. <https://doi.org/10.1007/s11858-023-01518-y>
- Sutton, A. (2009). Publication bias. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 435–452). New York: Sage.
- Tavşanlı, Ö. F., Kozaklı, T., & Kaldırım, A. (2018). The effect of graphic organizers on the problem posing skills of 3rd grade elementary school students. *PEGEM Journal of Education and Instruction*, 8(2), 377–406. <https://doi.org/10.14527/pegegog.2018.016>
- Tipton, E., & Pustejovsky, J. E. (2015). Small-sample adjustments for tests of moderators and model fit using robust variance estimation in meta-regression. *Journal of Educational and Behavioral Statistics*, 40(6), 604–634. <https://doi.org/10.3102/1076998615606099>

- Toh, T. L., Santos-Trigo, M., Chua, P. H., Abdullah, N. A., & Zhang, D. (Eds.). (2023). *Problem posing and problem solving in mathematics education: International research and practice trends*. Springer. <https://doi.org/10.1007/978-981-99-7205-0>
- Torgesen, J. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice*, 15, 53–65. [https://doi.org/10.1207/SLDRP1501\\_6](https://doi.org/10.1207/SLDRP1501_6)
- Van Harpen, X. Y., & Presmeg, N. C. (2013). An investigation of relationships between students' mathematical problem posing abilities and their mathematical content knowledge. *Educational Studies in Mathematics*, 83(1), 117–132. <https://doi.org/10.1007/s10649-012-9456-0>
- 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>
- Voica, C., & Pelczer, I. (2009). Problem posing by novice and experts: Comparison between students and teachers. In V. Durand-Guerrier, S. Soury-Lavergne, & F. Arzarello (Eds.), *Proceedings of the Sixth Congress of the European Society for Research in Mathematics Education* (pp. 2356–2365). Lyon, France.
- Wang, M., Walkinton, C., & Rouse, A. (2022). A meta-analysis on the effects of problem-posing in mathematics education on performance and dispositions. *Investigations in Mathematics Learning*, 14(4), 265–287. <https://doi.org/10.1080/19477503.2022.2105104>
- Xia, X., Lü, C., Wang, B., & Song, Y. (2007). Experimental research on mathematics teaching of "situated creation and problem-based instruction" in Chinese primary and secondary schools. *Frontiers of Education in China*, 2(3), 366–377. <https://doi.org/10.1007/s11516-007-0030-y>
- Zhang, L., Stylianides, A. J., & Stylianides, G. J. (2022b). Problematizing the notion of problem posing expertise. In J. Hodgen, E. Geraniou, G. Bolondi & F. Ferretti. (Eds.), *Proceedings of the Twelfth Congress of European Research in Mathematics Education (CERME12)* (pp. 4058–4065). Free University of Bozen-Bolzano and ERME.
- Zhang, L., Cai, J., Song, N., Zhang, H., Chen, T., Zhang, Z., & Guo, F. (2022a). Mathematical problem posing of elementary school students: The impact of task format and its relationship to problem solving. *ZDM—Mathematics Education*, 54(3), 497–512. <https://doi.org/10.1007/s11858-021-01324-4>
- Zhang, L., Stylianides, A. J., & Stylianides, G. J. (2023). Identifying competent problem posers and exploring their characteristics. *Journal of Mathematical Behavior*, 72, 101086. <https://doi.org/10.1016/j.jmathb.2023.101086>
- Zheng, L., Bhagat, K. K., Zhen, Y., & Zhang, X. (2020). The effectiveness of the flipped classroom on students' learning achievement and learning motivation. *Journal of Educational Technology & Society*, 23(1), 1–15.

## Publisher's Note

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