Metacognition and Learning (2025) 20:7
https://doi.org/10.1007/s11409-024-09405-x

Are metacognition interventions in young children effective? Evidence from a series of meta-analyses

Janina Eberhart[1] · Franziska Schäfer[2] · Donna Bryce[1,3]

Received: 18 January 2024 / Accepted: 30 October 2024

Abstract

A metacognitive learner acts in a planful way, monitors their progress, flexibly adapts their strategies, and reflects on their learning. Unsurprisingly, a metacognitive approach to learning is an important predictor of children's academic performance and many attempts have been made to promote metacognition in young children. The current meta-analytic study evaluates the impact of such metacognition interventions on outcomes related to self-regulated learning and academic achievement in typically developing pre- and elementary school children. Structural, content-related, and methodological moderators were tested in this study including 349 effect sizes from 67 studies. An overall effectiveness of metacognition interventions was evidenced at immediate post-test $g = 0.48$ (95% CI [0.35, 0.61]), and at follow-up $g = 0.29$ (95% CI [0.17, 0.40]). Interestingly, metacognition interventions effectively enhanced children's self-efficacy only at follow-up, suggesting that the positive impact of these interventions can unfold over a protracted period. For the first time, children's executive functions were considered as outcome variables and results indicated a positive impact of metacognition interventions on these variables. One notable finding was that interventions that were delivered by teachers or task materials were more effective than interventions that were delivered by researchers for two outcomes related to self-regulated learning. This finding may reflect recent improvements in how teachers and researchers collaborate to develop intervention programs. The study supports and extends existing evidence that young learners benefit from metacognition interventions in myriad ways and provides novel insights relevant for pedagogical practice and theories of self-regulated learning.

Keywords Meta-analysis · Metacognition · Preschool · Elementary school · Metacognition interventions · Self-regulated learning

Introduction

After a midyear review, an elementary school teacher recognizes that many of her third-grade pupils are working less independently than she would expect for their age. Very few children plan their tasks or set goals for themselves, many rely solely on teacher feedback about their learning progress, and they appear helpless when faced with a challenge, not

Extended author information available on the last page of the article

Springer

knowing which strategies they should apply to solve their problem. The skills that this teacher has noticed many children are lacking in her classroom can be termed metacognitive skills: cognitive processes that enable learners to plan, reflect on and regulate their thinking and learning processes (Nelson & Narens, 1990). Metacognition contributes significantly to children's learning and achievement at school (Dent & Koenka, 2016; Ohtani & Hisasaka, 2018) and, fortunately for this teacher, intervention studies have shown that children's metacognitive skills can be enhanced (e.g., Carretti et al., 2014; Dresel & Haugwitz, 2008). As a result, metacognition has received increasing attention and researchers, practitioners and policy makers alike aim to understand how children can be best supported in developing this skill (Quigley et al., 2021). Promoting metacognitive skills in pre- or elementary school promises to be particularly beneficial as it may set children on a positive learning trajectory. Seminal meta-analyses in this field have made important contributions to our understanding of how metacognition and related constructs can be promoted (e.g., Dignath & Büttner, 2008; Dignath et al., 2008; Donker et al., 2014). Subsequently, various new intervention and training studies that have benefitted from these insights have been developed and implemented. In the present meta-analytic study, we evaluate the effectiveness of contemporary metacognition interventions for outcomes related to self-regulated learning and academic achievement in young learners. Further, we investigate whether the effectiveness of metacognition interventions is moderated by the age of the children involved, structural features and content-related characteristics of the interventions. Notably, which components of metacognition were targeted within interventions, as well as the instructional methods adopted in these interventions were examined. With these analyses, we aim to report the current state of the art regarding metacognition interventions, and to provide insights that are valuable both for researchers engaged in progressing theory in this field and practitioners who want to enhance metacognition in their classrooms.

Metacognition: A critical subdomain of self-regulated learning

In its original conceptualization, metacognition encompassed metacognitive knowledge and metacognitive experiences (Flavell, 1979). Flavell described metacognitive knowledge as the knowledge and beliefs that one has about one's own cognitive processes (e.g., a child may know that they typically struggle with reading comprehension). Since then, the definition of metacognitive knowledge has been broadened to include knowledge about tasks, strategies and persons (Whitebread et al., 2009). In turn, metacognitive experiences include experiences of understanding or not understanding specific concepts (e.g., realizing that a paragraph that has been read was not understood). Brown (1987) used a similar differentiation but suggested that knowledge about cognition is stable and stateable whereas the regulation of cognition is unstable and not stateable. While Flavell's (1979) concept of metacognitive experiences focused on the monitoring of cognitive processes, Nelson and Narens (1990) offered an information processing perspective of the processes that enable metacognitive learners to regulate their learning.[1] That is, they considered not only monitoring but also the adaptive processes that follow monitoring, namely metacognitive control. In this model a differentiation is made between the meta- and object-level whereby the object-level includes the task or cognitive activity, and the meta-level contains a mental

[1] While originally developed as a model of metamemory, this model has since been applied outside of the memory domain.

representation of this task or cognitive activity (Nelson & Narens, 1990). Metacognitive monitoring processes are thought to be taking place when information flows from the object- to the meta-level, serving to update the mental representation of one's own cognition. Based on information gleaned through monitoring processes, learners might adapt their behavior which is referred to as metacognitive control. While Nelson and Narens (1990) focus on the two metacognitive skills that are predominantly implemented while processing a cognitive or learning task, Veenman and Elshout (1999) and Whitebread and colleagues (2009) highlighted additional behaviors metacognitive learners engage in such as planning and self-reflection.

Whereas models related to metacognition have their origin in cognitive psychology (e.g., Flavell, 1979; Nelson & Narens, 1990), other models stemming from educational psychology describe similar cognitive processes in models of self-regulated learning (Kim et al., 2023). Indeed, a review of prominent self-regulated learning models concluded that metacognition is a common feature of these models (Panadero, 2017). For instance, Zimmerman's (2002) cyclical phases model includes monitoring and control processes during the performance phase and Boekaerts (1999) considers metacognitive strategies as an important component of self-regulated learning. Self-regulated learners are aware of their skills and can be characterized as proactive and resourceful in the face of challenges, whether these challenges are cognitive, metacognitive, or motivational (e.g., Zimmerman, 2002). Thus, self-regulated learning is a broader concept which includes various subdomains, amongst them metacognition. In the current paper we focus on interventions that aim to promote metacognition. However, depending on their theoretical grounding, some studies that aim to promote metacognition also address other aspects of self-regulated learning (e.g., cognitive or motivational aspects). Rather surprisingly, no previous meta-analysis has focused specifically on the metacognitive features of interventions and examined their impact on the other subdomains of self-regulated learning as we do here. Instead, some previous reviews and meta-analyses examined the effectiveness of self-regulated learning interventions (Dignath & Büttner, 2008; Dignath et al., 2008; Donker et al., 2014) and distinguished broadly between the different subdomains (e.g., metacognitive, cognitive, motivational). In the current study we focus on interventions that target metacognition and try to elucidate what features are important for their effectiveness.

The effectiveness of metacognition interventions on child outcomes

An increasing number of interventions aim to promote metacognition as this skill is associated with a variety of child outcomes (e.g., Ohtani & Hisasaka, 2018). Children who regularly plan and monitor their learning, and change their approach if necessary, are not only more likely to achieve more academically but they may also enjoy advantages in other domains of self-regulated learning (e.g., they may become aware of what they can do which may lead to increased self-efficacy). Thus, researchers have aimed to synthesize findings of individual intervention studies in meta-analyses and reviews (e.g., Dignath & Büttner, 2008; Dignath et al., 2008; Donker et al., 2014; Wang & Sperling, 2020). However, none of the existing meta-analyses focused specifically on metacognition interventions; instead, all were conducted in the scope of self-regulated learning and therefore also included interventions that target adjacent skills and learning processes such as cognitive strategies and motivation regulation. In the following, selected findings from such studies including young learners will be summarized, focusing on insights that contribute to our hypotheses regarding the effectiveness of promoting metacognition and the role of potential moderators. In general, we expect to find that interventions designed to promote metacognition in young learners effectively enhance various child outcomes, in line with previous related meta-analyses. In the following, we establish the relationship between metacognition and child outcomes related to academic achievement as well as self-regulated learning.

Outcomes related to academic achievement

Unsurprisingly, since metacognition is characterized by the active engagement in thinking about one's own thinking, metacognition is a predictor of learning and success at school (Dent & Koenka, 2016; Ohtani & Hisasaka, 2018). Two recent meta-analyses found small but significant associations between elementary school children's metacognitive processes and their academic performance (Dent & Koenka, 2016; Ohtani & Hisasaka, 2018), even after controlling for intelligence (Ohtani & Hisasaka, 2018). Furthermore, another set of meta-analyses explored if interventions that focus on self-regulated learning, including metacognitive skills, enhanced young learners' academic outcomes. These meta-analyses found that interventions that focused on enhancing metacognitive skills led to improvements in children's academic skills (Dignath & Büttner, 2008; Dignath et al., 2008; Donker et al., 2014). Subsequently, de Boer and colleagues (2013) focused on identifying the reasons behind the effectiveness of self-regulated learning interventions. They examined which (combination of) learning strategies most improved children's academic performance and found that interventions incorporating metacognitive aspects, such as metacognitive knowledge, planning, and predicting were more effective in improving students' performance than interventions lacking these strategies. In a more recent meta-analysis, Zheng (2016) synthesized the effectiveness of interventions utilizing self-regulated learning scaffolds within computer-based learning environments. Some of the self-regulated learning scaffolds were metacognitive, for instance those that encouraged students to monitor their learning processes. The findings indicated that these self-regulated learning scaffolds improved academic performance. Donker et al. (2014) found the largest effect of metacognition interventions on writing. Thus, metacognition interventions appear to have a positive impact on children's academic achievement, but it is less clear if these interventions are more beneficial for some academic outcomes than others.

Outcomes related to self-regulated learning

As stated previously, metacognition interventions may not only have an impact on academic outcomes, but they may also enhance other skills, behaviors or attitudes subsumed under the concept self-regulated learning. For instance, we would hope that interventions with activities that engage children in metacognitive processes such as planning, monitoring, control, and/or reflection, help children to improve in exactly these skills. That is, if children repeatedly engage in planning processes, they should become more sophisticated planners. Consistent with this, Dignath and colleagues (2008) and Dignath and Büttner (2008) focused on self-regulated learning interventions conducted in classroom environments and observed positive impacts on metacognitive strategies amongst others. Thus, it can be expected that metacognition interventions promote *metacognition*.

In addition to the expectation that metacognition interventions improve metacognition, improvements in motivational outcomes can also be hypothesized. Self-efficacy is the belief that one can successfully complete a task or achieve a goal (Bandura, 1977), and this may be positively influenced by metacognition interventions if children experience that they can take control of their learning and successfully master a task when they approach it more metacognitively. For instance, a study that helped elementary school students to plan their study time at home found that children developed higher levels of self-efficacy in comparison to children who did not receive this support (Lai & Hwang, 2016). Other motivational processes that drive a learner to begin and maintain their engagement with a learning activity may be affected by metacognition interventions for the same reasons. Such effects may include changes in goal orientation (increased mastery goals, increased intrinsic motivation), changes in task value and interest, and changes in motivation regulation (e.g., enhanced self-talk and persistence). Indeed, Theobald (2021) reported positive impacts of self-regulated learning training programs on older learners' motivational outcomes (amongst other outcomes). Thus, two other skills that may be promoted through metacognition interventions are *self-efficacy* and *motivational orientations*.

Metacognition interventions may promote metacognitive control by aiding children to independently select the appropriate cognitive strategy (e.g., rehearsal, elaboration) during learning. They may also prompt children to engage in task planning before embarking on a learning activity. One could then reasonably assume that children would become more adept at implementing sophisticated cognitive strategies on demand and have more opportunity to apply and refine various organizational and resource management skills. We name such skills, which include for example the organization of task material, knowing when to seek help, or managing one's time efficiently (e.g., Lai & Hwang, 2016; Stoeger et al., 2014), *learning strategies*.

Another possible outcome of metacognition interventions is that enhanced monitoring and regulation would lead to improvements in children's *executive functions*. Executive functions are higher-order cognitive skills thought to support self-regulation that include inhibitory control (the ability to inhibit a prepotent response in favor of a more appropriate response), updating of working memory (the ability to hold information in mind while manipulating it), and cognitive flexibility (the ability to flexibly shift between responses, tasks or strategies; Miyake et al., 2000; Müller & Kerns, 2015). Executive functions are linked to academic (Spiegel et al., 2021), social and behavioral outcomes (Stucke & Doebel, 2023) in young children. Further, a link between metacognition and executive functions has already been made both theoretically (e.g., Fernandez-Duque et al., 2000; Roebers, 2017) and empirically (e.g., Bryce et al., 2015; Kälin & Roebers, 2022), with some authors suggesting that executive functions "may be the building blocks that metacognitively sophisticated thinkers use in their achievement of complex tasks" (Fernandez-Duque et al., 2000, p. 291). Considering the impact of metacognition interventions on executive functions is a unique contribution of this meta-analytic study.

Supporting children to behave more metacognitively is hypothesized to benefit them in various ways, which makes it relevant to consider the impact of interventions on various child outcomes. As well as considering the impact of metacognition interventions on general outcomes (that is, averaging across all outcomes that were measured in the included studies), in the current study we also consider the impact on specific outcomes related to self-regulated learning and academic achievement.

Long-term outcomes

The longer-term effectiveness of psychological and educational interventions is often not assessed within individual studies, which means it is often not possible to evaluate within meta-analyses. For instance, Dignath et al. (2008) were not able to analyze the effect sizes for follow-up assessments as too few studies included these. Recent calls for the routine inclusion of longer-term follow-up assessments (Watts et al., 2019) and suggested methods for doing so (Llewellyn-Bennett et al., 2016) reflect a shift in values regarding this practice. Accordingly, in the current study we aimed to also evaluate the longer-term effectiveness of the metacognition interventions included if the dataset allowed for this.

The role of potential moderators

To better understand whether specific factors make metacognition interventions effective, in the current study we aimed to examine several potential moderators.

Age group

Firstly, we tested if children of a specific *age group* would benefit more from metacognition interventions than others. Dignath and Büttner (2008) found that elementary school children benefitted more from self-regulated learning interventions in terms of academic performance in mathematics and other subjects, and secondary school children had larger effects in terms of academic performance in reading/writing and strategy use. An often-used rhetoric suggests that young learners benefit from learning strategy training as they are still developing these skills and do not yet possess established learning routines (Dignath & Büttner, 2008). Older students with already established study skills might be less willing to adapt their skills (Hattie et al., 1996). A meta-analysis that included a narrow age group similar to the one included in the current study, elementary school children, found no effect of age when considering the impact of interventions on academic performance (Dignath et al., 2008). Thus, all age groups seem to benefit from self-regulated learning interventions. However, results are somewhat mixed when different outcomes are taken into consideration and only one meta-analysis (Dignath et al., 2008) focused on elementary school children.

Structural characteristics

Metacognition interventions may also vary in their effectiveness as a function of *structural characteristics*. Most previous meta-analyses focused on studies that implemented metacognition interventions in educational settings (e.g., Dignath & Büttner, 2008; Dignath et al., 2008, 2023) and excluded studies with interventions delivered in highly controlled settings. Thus, the influence of the study environment (naturalistic vs. highly controlled) is unclear. In turn, previous results indicated that interventions that were conducted by researchers rather than teachers were more effective (Dignath & Büttner, 2008; Dignath et al., 2008). Other aspects such as the duration of the intervention or group size in which the intervention was conducted indicated mixed results. Some meta-analyses and reviews found that interventions that had more sessions were more impactful (Dignath & Büttner, 2008), whereas others found no impact or mixed results of duration (Dignath et al., 2008; Wang & Sperling, 2020). Authors stated that it requires time for students to acquire and practice new skills which is related to the length of the intervention. However, duration can be conflated with other factors such as intensity (e.g., number of sessions per week) and the person who delivered the intervention. For example, for practical reasons it is more likely that teachers implement longer interventions which could mean that they are less intense but also that teachers encourage the application of metacognitive skills outside of the suggested intervention time. In the current meta-analytic study, we consider a range of structural characteristics of intervention studies with the aim of identifying features that lead to the greatest improvements in child outcomes. It is feasible that more recent developments in how researchers and teachers work together may have caused a shift in the design and effectiveness of interventions.

Content-related characteristics

In addition to structural characteristics, *content-related characteristics* may also play a role in the effectiveness of metacognition interventions. A characteristic of metacognition interventions is that they are typically implemented within an academic school subject (Roebers, 2017). For example, children may be asked to monitor their comprehension while reading a text or to create a plan when completing a writing task. As such, some metacognition interventions have an elementary focus on promoting metacognition, whereas others focus primarily on the academic school subject while also promoting metacognitive skills on the side. This variation of focus is a common feature of metacognition interventions but has not been addressed in any of the previous meta-analyses. One might speculate that metacognition interventions with a primary focus on metacognition may lead to larger improvements in certain outcomes (e.g., metacognition) than interventions with a secondary focus on metacognition. Consequently, this seems like a relevant factor to consider when assessing the effectiveness of metacognition interventions.

As mentioned previously, most meta-analyses and reviews have included studies with interventions that targeted self-regulated learning strategies more broadly (Dignath & Büttner, 2008; Dignath et al., 2008; Donker et al., 2014; Wang & Sperling, 2020). Metacognition itself, however, is a multifaceted concept including various components such as planning, goal-setting, monitoring, control, self-reflection and evaluation. Dignath et al. (2008) did code whether interventions targeted the components planning, monitoring and/or evaluation and their results indicated that a combination of planning and monitoring, or planning and evaluation was most effective (Dignath et al., 2008). Dignath et al. (2023) focused specifically on tools that support one component of metacognition, namely metacognitive monitoring, and found that using these tools was beneficial for academic achievement, self-regulated learning and motivation. In the current meta-analytic study, we aim to provide empirical evidence not only regarding the overall effectiveness of recent metacognition interventions, but also regarding whether targeting particular metacognition components (namely, planning, monitoring, control, and/or reflection) is most beneficial for children's learning. Insights regarding the impact of targeting specific components of metacognition would provide evidence for potential mechanisms and guide the development of more impactful interventions. Further, given that different disciplines place different emphases on the various components of metacognition, as reviewed previously, this analysis could be informative regarding the relevance of different theoretical models to application.

It is also important to understand whether specific instructional methods are beneficial for metacognition interventions. Intervention instructors (whether teachers or researchers) typically apply various instructional methods such as reminders or prompts, informed instruction, modelling, feedback, individual or peer activities in which children practice their metacognitive skills. In terms of peer activities, it has previously been shown that elementary school students did not benefit from group work (Dignath & Büttner, 2008; Dignath et al., 2008). However, it might be premature to dismiss group work and other peer activities such as peer tutoring as effective instructional methods to promote metacognition as it may depend on how these are implemented. Elementary school students in particular might need additional instructions and support in order for their metacognition to benefit from peer activities (Dignath & Büttner, 2008). Further, previous findings have indicated that explicit self-regulated learning strategy instruction, which includes informed instruction about the value of a metacognitive approach, leads to greater learning gains than implicit strategy instruction, in which skills are applied but not explicitly discussed (e.g., Kistner et al., 2010). Importantly, however, this evidence comes from secondary school contexts, and as such the effectiveness of different instructional methods for elementary school children remains an open question. Since the effectiveness of different instructional methods has not been explored in previous meta-analyses and this information is of great interest to practitioners, we aim to address this in the current study.

Methodological features

The final set of moderators we evaluated in the current study considered the impact of *methodological features* of intervention studies. For instance, the type of outcome measures employed could impact the size of the effects observed. Previous studies observed larger effects for so-called online measures of metacognition (i.e. during learning) than offline measures of metacognition (i.e. before or after the learning task; Dent & Koenka, 2016; Ohtani & Hisasaka, 2018). In terms of academic achievement, previous studies have found stronger effects when it is assessed via researcher-developed tests than standardized tests (Donker et al., 2014), and stronger associations between metacognition and standardized tests than school grades (Dent & Koenka, 2016). Further, one may expect that studies that adhered to best practice in terms of intervention design may observe smaller effects than those that did not. Under this perspective we tested whether the type of control group employed (passive or active) and the type of allocation to group (random or non-random) impacted effect sizes.

The current study

In summary, while the seminal meta-analyses reviewed so far have provided invaluable insights into the effectiveness of metacognition and self-regulated learning interventions and helped to inform and inspire myriad new intervention studies, we have also identified certain outstanding issues that we aimed to address in this study. First, there is no recent meta-analysis focusing on young children and including interventions conducted both in naturalistic (e.g., classrooms) and highly controlled (e.g., labs) environments. Furthermore, the impact of metacognitive interventions on related outcomes such as executive functions has not yet been considered. Although previous studies classified interventions regarding their content (e.g., cognitive, metacognitive, or motivational aspects of self-regulated learning), more specific analyses such as differentiating if training focused on metacognition as a primary or secondary focus, or which components of metacognition were targeted (e.g., planning, monitoring, control, or reflection) were rarely included. Finally, the impact of how metacognition interventions were implemented (instructional methods) has not yet been addressed. This meta-analytic study aims to fill these gaps in the literature and update the field on metacognition interventions (comprised of more than one session, to distinguish from experimental manipulations) published over the last 23 years. Specifically, the current meta-analytic study aims to address four research questions. The first research question regards the effectiveness of metacognition interventions, the last three address potential moderator effects.

Research question 1: Are metacognition interventions effective?

a. Are metacognition interventions that are targeted at typically developing children (up to 12 years of age) effective in enhancing child outcomes? A significantly positive effect size is hypothesized.
b. Does the effectiveness of metacognition interventions vary by the types of outcomes (Metacognition, Self-efficacy, Motivational orientations, Learning strategies, Executive functions, Language, Mathematics, Other academic outcomes)? As previous studies have reported positive impacts of metacognition or self-regulated learning interventions on almost all of these outcomes, positive effects are expected. The exception here is executive functions, which has not been included as an outcome measure in previous meta-analyses.
c. Is there evidence for the longer-term effectiveness of metacognition interventions? Given the paucity of previous reviews regarding longer-term effectiveness of similar interventions, no strong hypotheses are made.

Research question 2: Is the effectiveness of metacognition interventions impacted by the age of the children targeted? Given the mixed findings in studies with young children, no strong hypotheses were formulated.

Research question 3: Does the effectiveness of metacognition interventions vary by characteristics of the intervention?

a. What impact do different structural characteristics (e.g., environment, who delivered the intervention, group size, duration and intensity) have on the effectiveness of interventions? Based on previous findings, it was hypothesized that researchers may be more effective than teachers in delivering interventions. No hypotheses were formulated regarding the other structural characteristics.
b. Are metacognition interventions that have their primary focus on metacognition more effective than metacognition interventions that have their secondary focus on metacognition? Based on theoretical considerations, it was hypothesized that interventions with a primary focus on metacognition would show larger effects on metacognition outcomes and other outcomes related to self-regulated learning than interventions with a secondary focus on metacognition.
c. Is the effectiveness of metacognition interventions affected by the intervention content (i.e., which metacognition component/s was/were targeted, from planning, monitoring, control, evaluation)? Since such a fine-grained coding of intervention content has not previously been conducted, it was not possible to generate evidence-based hypotheses.
d. Is the effectiveness of metacognition interventions affected by the instructional methods used in the intervention (e.g., informed instruction, modelling, prompts, practice, feedback, peer activities)? Similar as for RQ3c, it was not possible to generate evidence-based hypotheses.

Research question 4: Do methodological moderators affect the effect sizes observed? Given previous findings, we expected stronger effects for online than offline measures of self-regulated learning outcomes, and stronger effects for researcher-developed tests than standardized measures of academic outcomes than others.

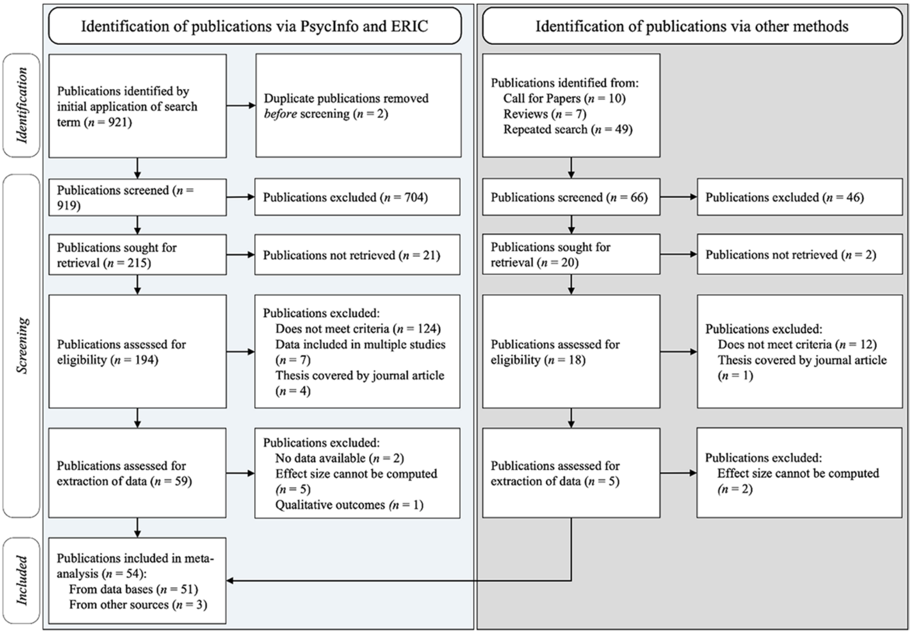

Fig. 1 PRISMA flowchart of identification and inclusion of publications. *Note. Publication* refers to journal articles, preprints, or theses

Methods

The goal of the current meta-analytic study[2] is to provide a comprehensive overview of interventions that aimed to promote metacognition in children. Studies reporting interventions that were conducted in controlled, lab-like environments as well as in naturalistic environments, such as children's classrooms, were included. The studies of the meta-analytic study tapped a variety of outcomes that can be grouped as those that relate to self-regulation / self-regulated learning (SRL-related outcomes) and those that relate to academic achievement.

Literature search

To address the research questions, comprehensive literature searches were conducted in relevant psychology and education databases (PsycINFO and ERIC) at different stages of the project with the most recent being in April 2024. The search terms were developed based on an existing meta-analysis by Dignath and Büttner (2008) and the search strategy applied is outlined in Supplementary Material Table S1. The search led to 919 publications with duplicates removed (see Fig. 1). Three researchers screened the titles and abstracts of the 919 publications and selected publications as potentially relevant if they were quantitative

[2] This study was preregistered on OSF, see https://osf.io/bjkxy/.

and included a metacognition intervention or training that aimed to enhance cognitive or academic skills of children under the age of 12 years. Title and abstract screening led to 215 potentially relevant publications. A more comprehensive coding process, based on retrieved full publications, led to 51 publications. Additional publications ($n=66$) were identified via other methods, namely a call for unpublished material (via various email listservs and Twitter) and the examination of reviews and reference lists. After checking the eligibility criteria of these additional publications, three fulfilled the criteria and were added to the final set of publications. Thus, 54 publications were included in the current meta-analytic study, some of which contributed more than one study resulting in a total of 67 studies. We use the term study to describe a comparison between experimental and control group. Some publications included multiple relevant group comparisons or reported results for multiple intervention programs.

Eligibility criteria

Heterogeneity in study design is a well-known problem in meta-analyses. To make studies more comparable and impose specific standards, several criteria were predetermined and had to be fulfilled for studies to be eligible for inclusion.

Basic information

Publications had to be published in an academic journal or as a dissertation after the year 2000 and had to be written in English. The year 2000 was chosen because previous meta-analyses excluded studies that were conducted in highly-controlled environments (Dignath & Büttner, 2018; Dignath et al., 2008) or included articles on metacognition interventions that were published before 2000 (Hattie et al., 1996). Here we focus on interventions that are aligned with more recent developments in teaching and pedagogical practice.

Type of intervention

Studies had to report an intervention that included aspects of metacognition such as planning, monitoring, control, or reflection. Studies had to be empirical and quantitative. Studies did not have to exclusively focus on metacognition in their interventions (see section Content-related characteristics of the interventions for elaboration). The intervention had to consist of more than one session, to separate interventions from experimental manipulations, and also include a control group. Thus, correlational studies, case studies, and qualitative studies were not eligible for inclusion. Interventions that focused on enhancing emotion regulation or pro-social behavior were excluded.

Sample characteristics

Empirical studies with more than ten participants in the intervention and control groups were eligible. Studies that specifically targeted children with disabilities (i.e., physical, psychological, intellectual or socioemotional impairments; World Health Organization, 2011) were not included because different methods might be employed in interventions targeting children with disabilities. As such, we refer to our findings as reflecting the effectiveness of interventions on typically developing children. Studies that focused on children from specific socioeconomic groups such as low-income backgrounds or specific ability levels such as gifted children were included, as long as they were participating in the regular school system. In terms of age, only studies in which the sample had a mean age of 12 years or younger were eligible for inclusion. If an article reported information on multiple age groups, only the relevant age groups were included (e.g., Williams et al., 2002).

Numerical information

Studies had to report numerical data that could be used to calculate effect sizes. Ideally, studies reported mean values, standard deviations, and sample sizes. If this information was not available, pre-test adjusted post-test scores, gain scores, reported *F*-values, *t*-test results, or effect sizes were required (see section below on effect size calculations). If this information was unavailable, studies had to be excluded (e.g., when the information was not provided in the publication and authors were unable to provide the original data).

Coding process

For the coding process, we retrieved publications and extracted relevant information for the eligible studies. Some of the information that we extracted, such as sample size or study design, did not involve subjectivity. However, there was some information that had to be coded and where some degree of subjectivity was involved, such as coding the focus of the intervention. In the cases where subjectivity was involved inter-rater reliability information is provided. The coding team comprised the three authors (with expertise in education, psychology, and metacognition) and a trained research assistant.

General information

Author information, publication year and type of publication (e.g., journal article, dissertation) were extracted. To describe the participants of the study, the mean age of the children, sample size, and the country and region in which the study was conducted were coded. If the mean age of participants was not reported in years and months, the age of participants was estimated from their school year using OECD data (https://gpseducation.oecd.org/Home) or the SchoolWix website (https://schoolwix.com/).

Study design

Whether interventions included an active or passive control group was coded. An active control group takes part in activities that are similar to the real intervention (e.g., similar levels of novelty), but do not include the key ingredients (e.g., components that target metacognition). Passive control groups, also referred to as business as usual, remain in their regular setting. If studies included an active and passive control group, the active control group was chosen as a comparison to the intervention group. It was also recorded if the assignment to these control groups was random. Interventions were coded as random if children or classrooms were assigned randomly to the intervention or control group and non-random if some self-selection was implicated in group assignment (e.g., teachers' willingness to take part in an intervention study).

Structural characteristics of the interventions

Information on how the interventions were structured and delivered was extracted. This included the environment such as if the intervention was delivered in a highly controlled environment (e.g., a research lab or a separate room in a school or pre-school) or in a naturalistic environment (e.g., children's classrooms). Furthermore, we coded the mode of delivery, that is if the intervention was delivered predominantly by a teacher, a researcher, task materials (either via computer program or workbooks), or a combination of a teacher and task materials. Group size was also coded, namely if the intervention was delivered to children individually, in small groups (minimum two children), or in a whole classroom setting. Finally, the intensity of the intervention was recorded. Intervention intensity was operationalized as number of sessions, frequency, and duration of the intervention. Number of sessions was the absolute number of sessions that the intervention included overall. Frequency indicated how many sessions were delivered per week. Finally, the duration of the intervention comprised the number of weeks that an intervention lasted. Thus, the three intensity descriptors describe different aspects of intensity. Interventions can be short in duration with many sessions whereas other interventions can stretch over many weeks with fewer sessions.

Content-related characteristics of the interventions

Metacognition is a skill that has to be applied in the context of a task (e.g., monitoring comprehension while reading) and it is typically assessed in educational contexts such as reading, writing, or math (Roebers, 2017). Interventions vary in the extent to which children's metacognition is targeted in these interventions; specifically, whether interventions have their primary or secondary focus on metacognition (see Table 1). For instance, a study with a primary focus on metacognition is Fuchs et al. (2003), in which children in the intervention group engaged in various activities such as scoring their work, tracking their progress with child-friendly methods, and setting themselves goals before starting a new task. In contrast, an example of a study with a secondary focus on metacognition is Artuso et al. (2019), in which children in the intervention group practiced reading comprehension extensively and only engaged in metacognitive control when choosing the appropriate reading strategy. The intervention focus was coded as mutually exclusive for each study. At this stage some studies were identified that did not target metacognition at all or insufficient information was provided regarding the content of the intervention; these were excluded. Three coders were involved in this coding process. Cohen's Kappa between two of the coders was excellent (κ= .84; 89.5% agreement). Because there was lower interrater reliability between the third coder and the other two coders (κ = .42 and .53; percent agreement 62% and 70%), the coding of the third coder was subsequently recoded by one of the other two coders. Furthermore, the intervention content of the metacognitive part of the intervention (i.e. which components of metacognition were targeted) and the specific instructional methods employed in the metacognitive part of the intervention were coded (see Table 1 for details). This coding process was not mutually exclusive and multiple codes could be given for both metacognition intervention content and instructional methods. The metacognition intervention content and instructional methods of all studies were double coded by the three authors. Due to our conservative approach to double code the intervention content and instructional methods of all studies interrater reliability was

Table 1 Names and definitions of content-related characteristics coded in the meta-analytic study

Code	Definition
Intervention focus: Metacognition as a…	
Primary focus	The intervention focused only on training metacognition or targeted both an academic skill and metacognition to a similar extent
Secondary focus	The intervention targeted an academic skill for most of the time and aspects of metacognition were only addressed sometimes
Metacognition intervention content	
Planning	Before starting a task, children are encouraged to think about how they could approach a task or set goals for themselves. Does not include task instructions or the goals of the task being given to them by an adult or another peer
Monitoring	During a task, children are encouraged to engage in monitoring processes, make explicit monitoring judgements (prospective, concurrent or retrospective), or self-observe. Does not include monitoring a peer
Control	During a task, children are encouraged to select appropriate strategies, or otherwise implement changes in a task based on information gained through monitoring. Does not include simply practicing new strategies as instructed
Reflection	After a task, children engage in self-evaluation or reflection, or re-evaluate an earlier monitoring judgement
Instructional methods	
Informed instruction	Children are explicitly provided information about some aspect of metacognition; the child passively consumes this information. May include explicit teaching of metacognitive strategies and/or the conditions for using different strategies, may be independent of a learning task and include definitions of the components of metacognition. Does not include teaching academic skills
Modelling	Demonstration of a metacognitive activity or skill by an adult, a peer or a fictional character. For example, a teacher demonstrates monitoring during reading by thinking aloud. Important is that the process is demonstrated
Prompts	Children receive reminders about a metacognitive skill or strategy within the context of a task (e.g., "remember to plan your writing"), but are not necessarily compelled to apply the skill or strategy
Practice	Children practice applying a metacognitive skill or strategy. This is ensured by the task structure e.g., monitoring judgements are collected, planning worksheets are completed
Feedback	Children receive feedback on any aspect of their metacognition, e.g., their planning, goal setting, monitoring judgements, or strategy selection. May be provided by an adult, peer or computer program. The provision of grades or test scores was not considered as feedback
Peer activities	Children engage in metacognitive activities with peers. Could include peer feedback, peer tutoring, or cases in which one child ensures that another implements a metacognitive strategy or skill

Only one code could be given for intervention focus; multiple codes could be given for metacognition intervention content and for instructional methods. For the coding of metacognition intervention content, the modality of the activity was irrelevant; that is, each could be realized verbally, in written format, etc.

not calculated. Disagreements were resolved by discussion. Furthermore, it is important to note that some studies included multicomponent interventions targeting constructs such as executive functions or motivation in addition to metacognition.

Outcome measures and constructs

The intervention studies included in this meta-analysis employed a range of outcome measures to assess the effectiveness of their interventions. These outcome measures vary not only in the construct which they aim to assess but also in the type of outcome measure employed (for example, questionnaires, standardized tests, and think aloud protocols; please refer to Research Question 4: Methodological Features for an elaboration). This meta-analysis focuses on outcome measures that assess cognitive skills or cognitive processing rather than emotional processing. As such, any outcome measures that assessed emotional regulation, positive affect or negative affect (including anxiety) were not included. If an outcome measure could not be categorized as being either a positive or negative indicator of a construct it was excluded from the meta-analysis.

Eight outcome constructs that we coded and analyzed separately were defined in our pre-registration. In a deviation from our pre-registration, we additionally opted to code two aspects of motivation separately. Taking a similar approach to Theobald (2021), we coded on the one hand self-efficacy and self-concept and on the other hand goal orientation, intrinsic value and motivation regulation. While we acknowledge that the latter category is rather heterogenous and groups concepts from different theoretical traditions, we conceptualized these as reflecting motivational processes involved in initiating or maintaining engagement in a learning activity. For reporting, we group the outcome constructs according to whether they are related to, on the one hand, self-regulation and/or self-regulated learning (SRL-related outcomes) and on the other hand related to academic achievement (Academic outcomes). See Table 2 for the outcome constructs and their definitions.

Data analysis

Effect sizes were calculated to indicate the size of the difference between experimental and control groups at post-test while taking differences at pre-test into account. To calculate these effect sizes, we aimed to extract mean values, standard deviations, and sample sizes of the experimental and control groups at pre- and post-test. When this information was provided, the standardized mean difference between the groups at pre-test was subtracted from the standardized mean difference between the groups at post-test (following the approach by Lipsey & Wilson, 2001; and Wilson & Lipsey, 2000). If this information was not available, we noted pre-test adjusted post-test scores, gain scores, reported F-values, reported t-test results, or reported effect sizes and calculated the effect size accordingly. Formulas used to calculate effect sizes under different conditions are reported in full in Sect. 3.4.1 of our pre-registration (see https://osf.io/bjkxy/). When calculating effect sizes, where necessary outcome measures were reverse coded so that an increase from pre- to post-test indicates improvement on the outcome measure. We opted to apply the correction factor J to our calculated effect sizes to base our analyses on the more conservative Hedges' g.

As mentioned above, some publications reported data from more than one intervention program that met the criteria for inclusion in the meta-analysis (namely Ferreira

Table 2 Names and definitions of outcome constructs coded in the meta-analytic study

Outcome construct	Definition
SRL-related Outcomes	
Metacognition	Outcomes that measure a child's metacognitive knowledge, metacognitive monitoring, metacognitive strategies, or metacognitive control
Self-efficacy	Outcomes that measure a child's belief in their own abilities, either regarding learning in general or in a specific subject
Motivational orientations	Outcomes that measure a child's mastery-approach goals, intrinsic motivation, task value, interest, persistence or motivation regulation
Learning strategies	Measures that include how a child organizes learning materials, seeks help, manages their time, or uses sophisticated learning strategies
Executive functions	Outcomes that measure a child's executive functions, tapping at least one of: inhibitory control, updating of working memory, task switching/shifting/cognitive flexibility
Academic Outcomes	
Language	Measures of writing, reading (including comprehension, fluency, word reading), spelling, speaking, listening, vocabulary, knowledge of writing or reading strategies, in a child's native language
Mathematics	Outcomes that measure a child's mathematical reasoning, calculation skills, or other mathematical abilities
Other academic subjects	Measures of performance or knowledge in an academic subject other than Mathematics and Language. For example, English as a Foreign Language, Science
Other	Includes all other outcome measures that could not be coded in the categories above. For instance, short term memory, visuo-spatial abilities, non-verbal IQ, crystallized IQ

et al., 2017; Harris et al., 2023; Hoffmann, 2010; Huff & Nietfeld, 2009; Jacob et al., 2020; Salas et al., 2023; Schuster et al., 2020, 2023; Souvignier & Mokhlesgerami, 2006; Spörer & Schünemann, 2014). In such cases, each experimental group and an appropriate control group was entered as a separate study to the meta-analysis. In all but three cases (Harris et al., 2023; Salas et al., 2023; Schuster et al., 2020) this resulted in data from the same control group being entered multiple times into our meta-analysis. To account for dependencies in the data and prevent studies with shared control groups being too influential, before calculating the effect sizes the sample size of the repeated control group was adjusted by dividing it by the number of times it entered the meta-analysis (following the procedure of Dignath et al., 2008).

Many studies included multiple outcome measures of the same outcome construct. In our preregistered data analysis plans, we stated we would calculate effect sizes for each outcome measure individually, and then calculate a combined effect size for each outcome construct and each study; these aggregated effect sizes per outcome construct and their associated variances would be entered into data analyses. During the peer review process, we decided to deviate from our pre-registered plan and follow the suggestion of a reviewer to instead model the full nested structure of our dataset using multi-level linear mixed effects models. The advantages of this are threefold: 1) we are able to include more effect sizes while still taking the dependencies in the data into account, 2) we can more precisely pinpoint the source of heterogeneity in our data, and 3) there is some evidence that using multilevel models can allow for testing moderators more precisely, for example, because multiple effect sizes from the same sample and outcome are included that may differ on crucial moderators (Fernández-Castilla et al., 2020).

Where feasible, each meta-analysis was conducted both for the whole dataset (providing information on the effect sizes across various cognitive outcomes) and separately for each outcome construct (providing information on the effectiveness of metacognition interventions for each specific outcome construct). For the former, four-level models were fitted (using the ‘rma.mv’ function in *metafor*) in which the random effect of outcome measure (e.g., Think aloud; level 2) was nested within the random effect of outcome construct (e.g., Metacognition; level 3) which was nested within the random effect of Study (level 4), as described in Harrer et al. (2021). To approximate the variance–covariance matrix the ‘vcalc’ function of the *metafor* package was used, setting $\rho_{\text{outcome construct}}$ to 0.4 and setting $\rho_{\text{outcome measures}}$ to 0.7. For meta-analyses conducted separately for each outcome construct an analogous approach was taken whereby three-level models were fitted (omitting the random effect of outcome construct). In all cases, an overall effect size was calculated and homogeneity analysis indicated if true effect size differences remained in the data. In the case of multilevel models, heterogeneity can be attributed to different levels of the model and we achieved this for the four-level model using an adaptation of the variance decomposition for three-level models (Harrer et al., 2021 see R Script on OSF). Following Pigott (2012), we interpret I^2 values (the percentage of total variance attributable to each level) of 25% as small, 50% as medium, and 75% as large heterogeneity.

As well as estimating the mean effect size (to address Research Question (RQ) 1a and 1b), we examined the impact of various moderators on the mean effect size (to address RQs 1c to 4). Moderator analyses were conducted for the whole dataset using four-level models (providing information about the impact of moderators on the effect size for general cognitive outcomes) as well as for each outcome construct separately using three-level models where feasible. A priori we established that categorical moderator analyses would only be conducted where at least four effect sizes were available for each level of the moderator (following the procedure by Takacs & Kassai, 2019). Based on rules of thumb proposed by Borenstein et al. (2009) and Harrer et al. (2021), we only evaluated continuous moderators (namely, Age, intensity of the intervention) when there were ten studies available.[3]

Outliers and influential cases were assessed for each meta-analysis and where relevant addressed as reported in the preregistration.[4] When action was taken, this is reported in the Results sections. The risk of publication bias[5] was assessed for the analyses related to RQ1 and is reported at the end of that section.

[3] Note, this decision was not pre-registered.

[4] As the ‘influence’ function is not available for multilevel models in *metafor*, outliers from the multi-level random effects models were defined as ESs that were more than ± 3 SD from the mean ES.

[5] The method for assessing risk of publication bias deviated from our pre-registration since Duval and Tweedie’s (2000) Trim and Fill method is not possible for multilevel models. Instead Egger’s regression test was used whereby the model was modified to include the standard error of the effect sizes as a moderator, and a significant moderator effect indicates an asymmetrical funnel plot (Egger et al., 1997; Rodgers & Pustejovsky, 2021).

Table 3 Features of studies included in the dataset for meta-analyses

Study characteristics	*N* studies	% studies
Type of publication		
Published (journal articles)	58	87
Unpublished (e.g., theses)	9	13
World region[a]		
Europe	42	63
North America	16	24
Asia	2	3
Oceania	1	1.5
South America	1	1.5
Not reported	5	7
Age of sample		
0 – 4 years	0	0
5 – 8 years	14	21
9 – 12 years	53	79
Total sample size[b]		
20 – 50	20	30
50 – 100	16	24
100 +	31	46
Type of control group		
Active	36	54
Passive	31	46
Allocation to groups		
Random	48	72
Non-random	13	19
Not reported	6	9
Includes follow-up assessment values		
Yes	19	28
No	48	72

The total number of studies included was 67

[a] Only regions represented in our dataset are listed here. No studies were included from Africa or Central America and the Caribbean

[b] Total sample sizes less than 20 was not possible, as an inclusion criterion was that the experimental and control group each included at least 10 participants

Transparency and openness

In this manuscript we report our search terms, eligibility criteria, as well as the procedure for including and excluding publications for the meta-analytic study. We also provide a detailed description of the data analytical approach. Datasets including the raw data can be found online, as well as a data analysis script that allows the reader to conduct the analyses reported (see https://osf.io/bjkxy/). Data were analyzed using R, version 4.2.3 (R Core Team, 2023) and the package *metafor*, version 4.2–0 (Viechtbauer, 2010). This study was preregistered on OSF (the pre-registration can be found under the OSF link above).

Results

Descriptive statistics

In total, 349 effect sizes (ESs) could be calculated from 67 studies. A summary of the study characteristics is presented in Table 3. Of note is that the majority of studies were published in peer-reviewed outlets, conducted within Europe or North America and targeted children towards the end of elementary school. Most of the included studies were conducted in naturalistic classroom environments and the majority of the interventions targeted more than one metacognition component with more than one instructional method. In the following, analyses related to each RQ are presented, and a summary of all results is provided in Table 4.

Research question 1: The effectiveness of metacognition interventions

A forest plot containing every individual ES contained in the dataset is presented in Fig. 2. Figure 3 shows the overall average effect sizes and confidence intervals (Hedge's *g*, displayed as diamonds) for each outcome construct separately and for general outcomes. As can be seen, metacognition interventions do enhance child outcomes in young typically developing children, and they seem to improve certain outcomes more than others.

Research question 1a: General outcomes

To address RQ 1a, all ESs from all outcome constructs were analyzed together in a 4-level multilevel model as described previously. Six ESs ranging from $g = 3.22$ to $g = 4.21$ were identified as outliers (from Boykin et al., 2019; de Vreeze-Westgeest & Vogelaar, 2022; Iordanou, 2022; Salas et al. 2023; Tsiriotakis 2013) and winsorized to + 3 *SD* of the mean. After this, a positive mean effect size of $g = 0.48$, $SE = 0.07$, 95% CI [0.35, 0.61], $p < .001$, was estimated. Significant heterogeneity in the dataset was observed, $Q(348) = 6755.30$, $p < .001$, with $I^2_{Level\ 2} = 59.77\%$ (variance attributable to different outcome measures of the same outcome constructs), $I^2_{Level\ 4} = 33.36\%$ (variance attributable to different studies), and no variance attributable to different outcome constructs within studies (level 3).

An adapted Egger's regression test was conducted to evaluate the risk of publication bias in the dataset, and the significant moderator effect, $Q(1) = 72.78$, $p < .001$, provided evidence for an asymmetrical funnel plot. An annotated funnel plot can be found in the Supplementary Materials. It is of note that the asymmetry appears to be driven by the lack of negative or low ESs opposite the ESs contributed by two studies (Ferreira et al., 2017; Iordanou, 2022) and is not a more widespread problem of publication bias or selective reporting. Consequently, we ran a moderator analysis with Publication Type as a moderator. This indicated that publication type did not significantly moderate the effect size, $Q(1) = 2.44$, $p = .118$, $B = -0.32$, although it should be noted that the ESs from theses came from only nine studies. Indeed, the effect went in the direction one would expect – estimated mean ES was small and non-significant for theses ($g = 0.20$, $p = .335$, $k = 48$) and large and significant for published journal articles ($g = 0.52$, $p < .001$, $k = 301$).

Table 4 Overview of research questions and results for different types of outcomes

Research question	General Outcomes	Outcomes related to self-regulated learning					Outcomes related to academic achievement		
		Metacognition	Self-efficacy	Motivational orientations	Learning strategies	Executive functions	Language	Mathematics	Other academic subjects
1a/b: Are metacognition interventions effective?	✓	✓	×	×	✓	✓	✓	✓	✓
1c: Is there evidence for the long-term effectiveness?	✓	×	✓	×	×	–	✓	–	–
Moderator analyses									
2: Is effectiveness impacted by age?	×	×	×	✓	✓	×	×	×	×
3a: Environment	×	–	–	–	–	–	×	–	×
3a: Mode of delivery	×	✓	×	×	✓	–	×	×	×
3a: Group size	×	×	×	×	✓	–	×	–	×
3a: Intensity: No. of sessions	×	×	×	×	×	×	×	×	×
3a: Intensity: Frequency	×	×	×	×	×	×	×	×	×
3a: Intensity: Total duration	×	×	×	×	✓	×	×	×	–
3b: Focus	×	✓	×	×	×	×	×	✓	×
3c: Content focus – Planning	×	×	–	–	–	×	×	–	×
3c: Content focus – Monitoring	✓	✓	×	–	✓	–	×	–	–
3c: Content focus – Control	×	×	×	×	✓	×	×	×	×
3c: Content focus – Reflection	×	×	×	×	–	–	✓	×	–
3c: Content focus – No. of foci	×	×	×	×	✓	×	×	×	×
3d: Instructional method – Informed instruction	×	✓	–	–	–	×	×	×	×
3d: Instructional method – Modelling	×	×	×	×	×	–	×	–	×
3d: Instructional method – Prompts	×	×	×	×	×	×	×	–	×
3d: Instructional method – Practice	×	×	–	–	–	×	×	–	–
3d: Instructional method – Feedback	×	✓	×	×	×	–	✓	–	×
3d: Instructional method – Peer activities	×	×	–	×	–	–	×	–	–
3d: Instructional method – No. of methods	×	×	×	×	×	×	×	×	×

Table 4 (continued)

Research question	General Outcomes	Outcomes related to self-regulated learning					Outcomes related to academic achievement		
		Metacognition	Self-efficacy	Motivational orientations	Learning strategies	Executive functions	Language	Mathematics	Other academic subjects
4: Outcome type	✓	×	–	–	✓	–	✓	–	–
4: Type of control group	×	×	×	×	×	×	×	×	×
4: Assignment to groups	×	×	×	×	×	–	×	–	–

The checkmark (✓) indicates a significant result in the meta-analytical model conducted with respect to the research question, whereas the cross (×) indicates no significant result. The hyphen (–) indicates that no meta-analytical model could be conducted to answer the respective research question due to the low number of studies available

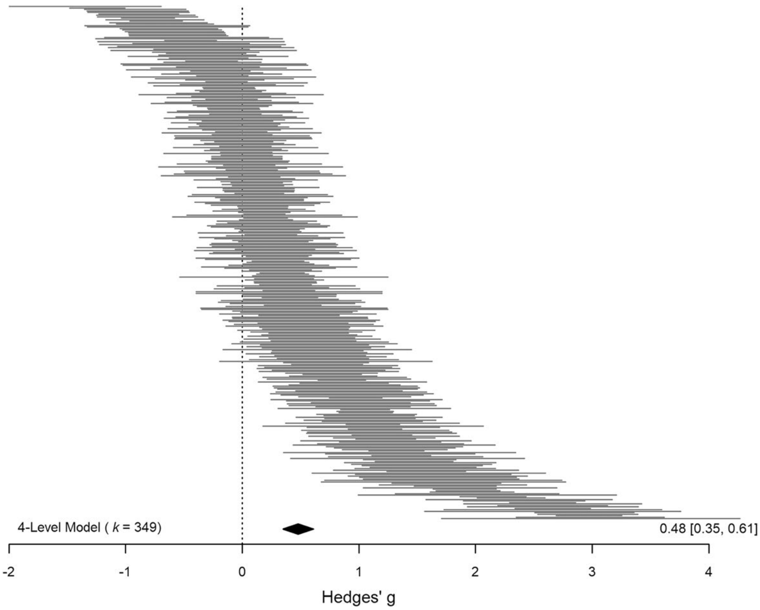

Fig. 2 Forest plot of every effect size in the dataset

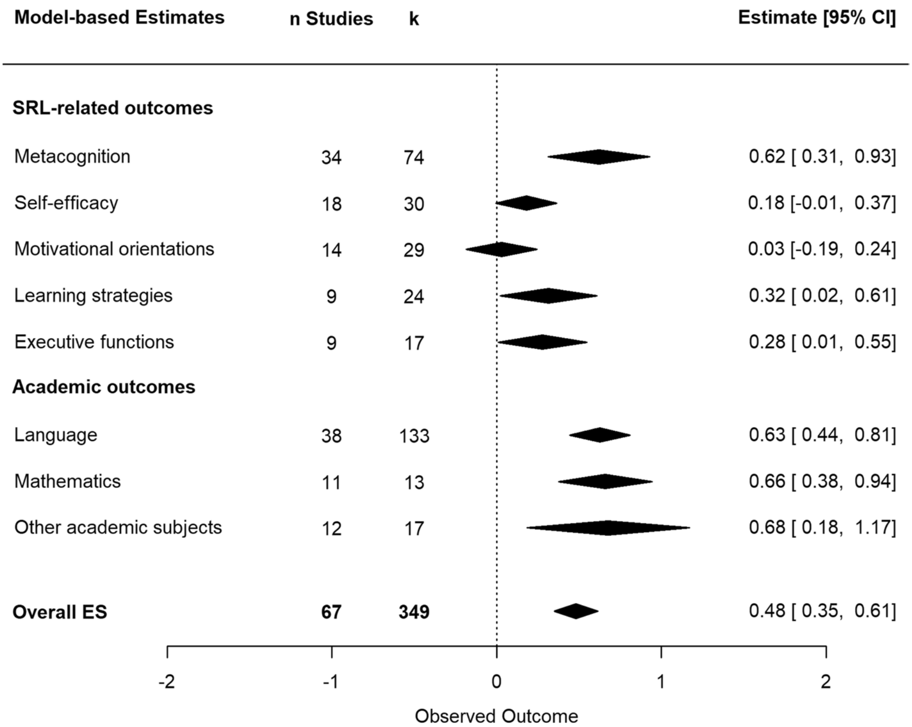

Fig. 3 Summary polygons for the meta-analyses conducted to address research questions 1a and b

Research question 1b: SRL-related outcomes

Of the five outcome constructs classed as being related to SRL, three (namely, Metacognition, Learning strategies and Executive functions) showed significant positive effects of metacognition interventions. Thirty-four studies measured Metacognition as an outcome variable, and after winsorizing one outlier ES (from de Vreeze-Westgeest & Vogelaar, 2022), an average effect size of $g = 0.62$, $SE = 0.16$, 95% CI [0.31, 0.93], $p < .001$ was estimated. There was significant heterogeneity in the data, $Q(73) = 1460.85$, $p < .001$, total $I^2 = 96.47\%$, with the majority of it (74%) being attributed to the level of studies. A positive impact of metacognition interventions was observed on children's Learning strategies. From nine studies an average effect size of $g = 0.32$, $SE = 0.15$, 95% CI [0.02, 0.61], $p = .036$ was estimated. Significant heterogeneity was observed, $Q(23) = 275.41$, $p < .001$, total $I^2 = 86.51\%$, with the majority of it (70%) being attributed to the level of outcome measures. Nine studies measured Executive functions as outcome variables, and from these an average effect size of $g = 0.28$, $SE = 0.14$, 95% CI [0.01, 0.55], $p = .043$ was estimated. There was significant heterogeneity in the data, $Q(16) = 158.49$, $p < .001$, total $I^2 = 75.35\%$, with all of this being attributed to the level of outcome measures. From the 18 studies that measured Self-efficacy as an outcome, a non-significant effect size of $g = 0.18$, $SE = 0.10$, 95% CI [-0.01, 0.37], $p = .058$ was estimated (one outlier ES from Perels et al., 2009 was winsorized first). Significant heterogeneity was observed, $Q(29) = 97.94$, $p < .001$, total $I^2 = 75.63\%$ of which 21% was attributed to outcome measures and 54% was attributed to studies. Fourteen studies measured Motivational orientations as an outcome, and after winsorizing one outlier ES (from Vandevelde et al., 2017) the estimated average effect size was close to zero, $g = 0.03$, $SE = 0.11$, 95% CI [-0.19, 0.24], $p = .791$. There was significant heterogeneity in the data, $Q(28) = 411.16$, $p < .001$, total $I^2 = 85.69\%$, the majority of which (65%) could be attributed to outcome measures. No ESs were identified as influential cases.

Research question 1b: Academic outcomes

Metacognition interventions appeared to have a positive impact on all three academic outcomes identified. Thirty-eight studies measured some aspect of Language, and after winsorizing three outlying ESs (from Boykin et al., 2019; Iordanou, 2022; Tsiriotakis, 2013) an estimated effect size of $g = 0.63$, $SE = 0.09$, 95% CI [0.44, 0.81], $p < .001$ indicated improvement following metacognition interventions. Significant heterogeneity was observed in the data, $Q(132) = 3458–59$, $p < 0.001$, total $I^2 = 93.98\%$ of which 79% was attributed to outcome measures. For the eleven studies that measured achievement in Mathematics, an overall significant improvement following metacognition interventions was observed, $g = 0.66$, $SE = 0.14$, 95% CI [0.38, 0.94], $p < .001$. There was significant heterogeneity in the data, $Q(12) = 129.64$, $p < .001$, total $I^2 = 84.22\%$, almost all of which was attributed to outcome measures (79%). The numerically largest average ES was estimated for the outcome construct Other academic subject (12 studies included), $g = 0.68$, $SE = 0.25$, 95% CI [0.18, 1.17], $p = .007$. There was significant heterogeneity observed, $Q(16) = 285.57$, $p < .001$, total $I^2 = 94.04\%$, almost all of which was attributed to outcome measures (88%). No ESs were identified as influential cases.

Research question 1c: The longer-term effectiveness of metacognition interventions

Of the 67 studies in the dataset, 19 reported follow-up assessment values that allowed us to calculate ESs for at least some outcomes. Follow-up assessments took place on average 15 weeks after the post-test assessment (range: 3 – 48 weeks, $SD = 11.71$). This subset of 19 studies and 97 ESs was employed for these analyses.

General outcomes

When ESs reflecting change from pre-test to follow-up assessment, all outcomes were analyzed together in a 4-level model and 97 ESs from 19 studies were included. After winsorizing one outlying ES (from Glaser & Brunstein, 2007), a positive mean effect size of $g = 0.29$, $SE = 0.06$, 95% CI [0.17, 0.40], $p < .001$ was estimated. Significant heterogeneity in the dataset was observed, $Q(96) = 1031.47$, $p < .001$, with $I^2_{Level\ 2} = 69.71\%$ (variance attributable to outcome measures), no variance attributable to different constructs within studies, and $I^2_{Level\ 4} = 11.88\%$ (variance attributable to studies). The adapted Egger's test provided evidence of an asymmetric funnel plot, possibly indicating selective reporting or publication bias in this dataset, $Q(1) = 7.18$, $p = .007$.

SRL-related outcomes

Four SRL-related outcomes were represented in high enough numbers to be analyzed separately, namely Metacognition, Self-efficacy, Motivational orientations and Learning strategies. For three of these the estimated ES was not significant at follow-up (Metacognition: $g = 0.12$, $SE = 0.10$, 95% CI [-0.08, 0.31], $p = .243$; Motivational orientations: $g = 0.04$, $SE = 0.09$, 95% CI [-0.13, 0.21], $p = .617$; Learning strategies: $g = 0.23$, $SE = 0.12$, 95% CI [-0.001, 0.45], $p = .051$). Six studies (7 ESs) measured Self-efficacy in a follow-up assessment and when these were analyzed, a significant average effect size of $g = 0.24$, $SE = 0.12$, 95% CI [0.01, 0.48], $p = .044$ was estimated. Significant heterogeneity was observed in the data, $Q(6) = 18.88$, $p = .004$, all of which could be attributed to the level of studies (total $I^2 = 74.66\%$). It should be noted that one ES from Brunstein and Glaser's (2011) outcome measure "writing self-efficacy" was identified as an influential case and the estimated effect size was no longer significant when this was dropped from the model ($p = .325$).

Academic outcomes

Of the academic outcomes, it was only possible to estimate the average ES for follow-up assessments of Language, as 14 studies reported this (contributing 41 ESs). One outlier ES was winsorized (from Glaser & Brunstein, 2007), after which a significant average ES was estimated, $g = 0.45$, $SE = 0.08$, 95% CI [0.29, 0.60], $p < .001$. There was significant heterogeneity in the data, $Q(40) = 269.32$, $p < .001$, with all of it being attributed to the level of outcome measures, $I^2_{Level2} = 73.79\%$.

Moderator analyses

Having evaluated the effectiveness of metacognition interventions on immediate and follow-up assessments, now we turn our attention to potential moderators of this effect. All moderator analyses were conducted with the complete dataset and ESs calculated from immediate post-test scores. Moderator analyses were conducted both for general outcomes and for each outcome separately where feasible. For outcome constructs where outliers were identified in the analyses conducted to address RQ1, moderator analyses were conducted using the outlier-corrected datasets and outliers will not be mentioned in each section. For all moderator analyses, selected findings are described in the text, full results including 95% CI for all levels of each moderator can be found in Supplementary Materials and a summary of these findings can also be found in Table 4. It should be noted that the various moderators only had an enhancing effect on the impact of interventions; there were almost no negative estimated effect sizes in these data indicating that in general interventions bring about positive change and certain features may be associated with null effects whilst others may be associated with positive effects.

Research question 2: The impact of age on the effectiveness of metacognition interventions

Fifty-eight of the 67 studies (contributing 324 ESs) included in the dataset were eligible to be included in an analysis to determine whether the age of the participating children is associated with the effectiveness of metacognition interventions. Studies which included participants from more than two school grades or more than two chronological ages (e.g., 9- to 11-year-olds) were not included in this analysis. Mean ages ranged from 5.3 to 11.7 years, with the majority of studies including children aged 9 to 12 years of age.

Age was not a significant moderator when analyzed within the 4-level model with all ESs included. Age of the sample was also entered as a continuous moderator in separate models for each separate outcome construct. Age emerged as a significant moderator of the ESs of Motivational orientations, $Q(1)=6.21$, $p=.013$, $B=0.11$, and Learning strategy outcome measures, $Q(1)=4.06$, $p=.044$, $B=-0.43$. Interestingly, the effectiveness of metacognition interventions on Motivational orientations increased with age, whereas it decreased with age for Learning strategies. These moderator effects should, however, be interpreted with caution, since there was very little variation in the age of participants in these datasets. The Motivational orientations result may have been strongly influenced by two studies with young (preschool) participants and negative ESs (both originating from Jacob et al., 2020), and the Learning strategies result may have been driven by the five ESs contributed by Lai and Hwang (2016) which had the youngest participants in this grouping (9.5 years) and the largest ESs. Age was not a significant moderator for any of the academic outcomes.

Research question 3a: Structural characteristics of metacognition interventions

The studies in this dataset varied substantially in terms of their structural characteristics. Various structural features were entered as moderators in separate meta-analyses in order to assess whether the environment in which the intervention took place (naturalistic vs. controlled), how the intervention was delivered (by teachers, researchers or task materials), the size of the group in which the intervention took place (individual training, small groups or whole class), or the intensity of the intervention impacted the effectiveness of the intervention.

Environment As the majority of studies (62 out of 67) were conducted in naturalistic environments, this moderator could only be analyzed within the 4-level model on general outcomes. Keeping this unequal distribution in mind, the results indicated that the environment in which the intervention took place was not a significant moderator.

Mode of delivery For these analyses, studies in which the intervention was predominantly delivered via task materials or task materials in combination with a teacher were grouped into one level. Mode of delivery emerged as a significant moderator of ES on Metacognition outcomes, $Q(2) = 16.49$, $p < .001$. The average ES was highest when interventions were delivered by task materials ($g = 1.22$, $p < .001$, $k = 10$ from six studies), followed by teachers ($g = 0.91$, $p < .001$, $k = 25$ from 17 studies) and researchers ($g = -0.12$, $p = .413$, $k = 39$ from 11 studies). Mode of delivery also emerged as a significant moderator of ES on Learning strategy outcomes, $Q(2) = 6.59$, $p = .037$ and showed the same pattern: the largest average ES was estimated when interventions were delivered by task materials ($g = 0.91$, $p = .004$, $k = 5$ from one study), followed by teachers ($g = 0.36$, $p = .051$, $k = 14$ from six studies) and then researchers ($g = -0.06$, $p = .765$, $k = 5$ from two studies). How the intervention was delivered was not a significant moderator for any of the academic outcomes.

Group size The size of group in which the intervention was conducted was a significant moderator for Learning strategy outcomes, $Q(1) = 4.28$, $p = .039$. This analysis indicated that the ES was larger when interventions were conducted in whole class groups ($g = 0.52$, $p = .008$, $k = 17$ from five studies) than when they were conducted in small groups ($g = -0.06$, $p = .765$, $k = 5$ from two studies). Otherwise, group size was not a significant moderator of the ES.

Intensity of intervention To assess whether intensity was a significant moderator of the impact of metacognition interventions, the number of sessions, frequency of sessions and total duration of the intervention were entered as continuous moderators in separate meta-analyses. Of all these moderator analyses, only one indicated that a measure of intensity was a significant moderator. The total duration in weeks was a significant moderator of the ES for Learning strategies outcomes, $Q(1) = 5.22$, $p = .022$, $B = -0.12$. This indicates a slight decrease in the impact of metacognition interventions on Learning strategies as the duration of the intervention increases. In this analysis the durations of interventions ranged from 3 to 10 weeks, and nine studies contributed 24 ESs.

Research questions 3b-d: Content-related characteristics of metacognition interventions

Research question 3b: Intervention focus In these analyses, we evaluate if studies that have a primary focus on metacognition lead to larger effect sizes in comparison to interventions that have a secondary focus on metacognition. For the outcome construct Metacognition, Focus was a significant moderator of the ES, $Q(1) = 5.84$, $p = .016$, $B = 0.71$, whereby smaller ESs were observed when metacognition was the primary focus ($g = 0.23$, $p = .121$, $k = 33$ from 16 studies) than the secondary focus ($g = 0.97$, $p < .001$, $k = 41$ from 18 studies). For one academic outcome, Mathematics, Focus was a significant moderator, $Q(1) = 4.12$, $p = .042$, $B = -0.56$. In this case, ESs were larger when metacognition was the primary focus ($g = 0.84$, $p < .001$, $k = 9$ from seven studies) than when it was the secondary focus ($g = 0.23$, $p = .047$, $k = 4$ from four studies).

Table 5 Content-related characteristics of studies included in the dataset for meta-analyses

Study characteristics	N studies	% studies
Intervention focus		
Primary focus on metacognition	29	43
Secondary focus on metacognition	38	57
Metacognition intervention content		
Planning	52	78
Monitoring	55	82
Control	25	37
Reflection	46	69
Number of components targeted		
One	7	10
Two	22	33
Three	25	37
Four	13	19
Instructional methods[a]		
Informed instruction	44	70
Modelling	36	57
Prompts	40	63
Practice	56	89
Feedback	20	32
Peer activities	10	16
Number of methods employed		
One	4	6
Two	18	29
Three	11	17
Four	19	30
Five	9	14
Six	2	3

The total number of studies included was 67. Only one code could be given for intervention focus; multiple codes could be given for metacognition intervention content and also for instructional methods

[a] For four studies it was not possible to code the instructional methods used in the intervention because insufficient information was provided

Research question 3c: Metacognition intervention content The majority of metacognition interventions targeted multiple components of metacognition (e.g., both monitoring and control). As can be seen in Table 5, the most often targeted component of metacognition was on-task metacognitive monitoring, followed by planning (which also included goal setting). Separate meta-analyses were conducted for each metacognition component included in the metacognitive part of the intervention (i.e., planning, monitoring, control, reflection). An intervention that targeted multiple components of metacognition entered multiple meta-analyses to test for an effect of including each metacognition component. In each of these meta-analyses, the inclusion of this particular component was entered as a categorical moderator. As such, it was investigated whether the inclusion of e.g., planning in an intervention impacted the size of the ES. The inclusion of Monitoring significantly impacted the estimated ES for general child outcomes, $Q(1) = 4.23$, $p = .040$, $B = -0.34$. This result indicates that when Monitoring was included in the metacognition intervention, the average ES was lower ($g = 0.41$, $p < .001$, $k = 277$ from 55 studies) than when it was not included ($g = 0.80$, $p < .001$, $k = 72$ from 12 studies). The moderator analyses for separate outcome constructs also resulted in two cases in which the inclusion of Monitoring in the intervention content significantly reduced the ES, namely for Metacognition outcomes ($Q(1) = 7.30$, $p = .007$, $B = -0.97$; Monitoring included: $g = 0.41$, $p = .005$, $k = 65$ from 27 studies; Monitoring not included $g = 1.45$, $p < .001$, $k = 9$ from seven studies) and for Learning strategies ($Q(1) = 4.03$, $p = .045$, $B = -0.74$; Monitoring included: $g = 0.18$, $p = .204$, $k = 19$ from eight studies; Monitoring not included: $g = 0.91$, $p = .004$, $k = 5$ from one study). The inclusion of Control was a significant moderator of the ES for Learning strategy outcomes, $Q(1) = 4.47$, $p = .035$, $B = -0.56$, whereby ESs were larger when Control was not included ($g = 0.58$, $p = .006$, $k = 9$ from four studies) than when it was included ($g = 0.07$, $p = .682$, $k = 15$ from five studies). Including Reflection in an intervention significantly moderated the estimated ES for the outcome construct Language, $Q(1) = 5.43$, $p = .020$, $B = 0.42$, whereby ESs were larger when Reflection was included ($g = 0.82$, $p < .001$, $k = 88$ from 25 studies) than when it was not included ($g = 0.34$, $p < .001$, $k = 45$ from 13 studies).

Further analyses tested for the effect of the number of components included in the intervention as a continuous moderator; that is, whether interventions that targeted more components of metacognition were more/less effective than those that targeted fewer. Number of components emerged as a significant moderator for the outcome construct Learning strategies, $Q(1) = 5.77$, $p = .016$, $B = –0.41$. This result indicated that as the number of metacognition components targeted in the intervention increased, the ES decreased. However, this finding should be interpreted with caution due to the limited variability in the number of components targeted by interventions and the unequal distribution of these across studies – most studies targeted three or four components, and only the study by Lai and Hwang (2016) targeted two components.

Research question 3d: Instructional methods A range of instructional methods were employed in the interventions (see Table 5). Seventy percent of intervention studies delivered Informed instruction about some aspect of metacognition and 89% ensured that children practiced the metacognitive activity that was targeted. Again, the majority of studies employed more than one instructional method, which reflects how varied these interventions were in terms of delivery. Analogously to the analyses of the intervention content, the impact of employing specific instructional methods on the effectiveness of the intervention was analyzed in a series of meta-analyses. Two such analyses indicated that the inclusion of an instructional method was a significant moderator of the ES for the outcome construct Metacognition. The inclusion of Informed instruction as an instructional method significantly decreased the ES for Metacognition outcomes, $Q(1) = 6.55$, $p = .011$, $B = -0.87$, whereby the average ES was lower when Informed instruction was used ($g = 0.34$, $p = .021$, $k = 59$ from 22 studies) than when it was not ($g = 1.22$, $p = .003$, $k = 8$ from eight studies). The inclusion of Feedback on metacognitive behaviors significantly increased the ES for Metacognition outcomes, $Q(1) = 4.26$, $p = .039$, $B = 0.75$, as the average ES was higher when feedback was included ($g = 1.20$, $p = .011$, $k = 26$ from seven studies) than when it was not ($g = 0.38$, $p = .010$, $k = 41$ from 23 studies). Likewise, the inclusion of Feedback on metacognitive behaviors significantly increased the ES for Language outcomes ($Q(1) = 4.58$, $p = .032$, $B = 0.39$; Feedback included: $g = 0.89$, $p < .001$, $k = 51$ from 12 studies; Feedback not included: $g = 0.48$, $p < .001$, $k = 80$ from 25 studies). Further analyses investigated the number of different instructional methods employed as a continuous moderator. These analyses tested whether interventions that included more types of instructional methods to enhance metacognition were more/less effective than those that used fewer. The number of methods did not emerge as a significant moderator in any analyses.

Research question 4: Methodological features

Type of measure Although not pre-registered, the type of outcome measure was coded as a possible methodological moderator of effect sizes. This was done differently for SRL-related outcomes and academic outcomes, closely following the approach taken by Dent and Koenka (2016). SRL-related outcome measures could be an offline questionnaire (e.g., subscales of the Motivated Strategies for Learning Questionnaire), an online behavioral measure (e.g., metacognitive monitoring judgments provided during a task), an online speech measure (e.g., coding of think aloud protocols), or another type of online measure (e.g., performance measures from a working memory updating task). To analyse Type of Measure as a moderator, these were grouped as offline ($k = 112$, from 31 studies) and online ($k = 63$, from 28 studies) measures. Academic achievement measures could be a researcher-developed study test (e.g., accuracy on a spelling task designed by the research team), a standardized test (e.g., the Arithmetic test from the WISC-III), a school grade (e.g., grade on class examination), or other (e.g., questionnaire about writing knowledge). Only the first two categories were included in the moderator analyses due to only one study employing school grades and the measures included in the category Other being very heterogeneous.

When examining the types of measures used to assess SRL-related outcomes, the Type of Measure adopted was not a significant moderator when all outcome constructs were analysed in the 4-level model. Interestingly, the types of measures were not evenly distributed across the various SRL-related outcome constructs. While all measures of Executive functions were online, the majority of measures of Self-efficacy and Motivational orientations were offline reflecting the dominance of computerized direct assessment in the former, and the dominance of questionnaires to assess the latter two constructs. Only the outcome constructs Metacognition and Learning strategies had sufficient numbers of both online and offline measurement types to allow for a moderator analysis. Type of Measure was only a significant moderator of the ES for Learning strategy outcomes, $Q(1) = 4.03$, $p = .045$. $B = 0.45$, whereby a larger ES was estimated for offline ($g = 0.44$, $p = .003$, $k = 20$ from eight studies) than online measurements ($g = -0.11$, $p = .665$, $k = 4$ from two studies).

The majority of academic outcomes were assessed using researcher-developed study tests ($k = 115$ from 39 studies), with 37 ESs (from 20 studies) being the result of standardized tests. When analysing all academic outcomes together, the Type of Measure was a significant moderator, $Q(1) = 24.37$, $p < .001$. This finding indicates that ESs are greater when they result from researcher-developed tests ($g = 0.88$, $p < .001$, $k = 115$) than from a standardized test ($g = 0.12$, $p = .015$, $k = 37$). Only the construct Language had sufficient studies that employed either researcher-developed tests or standardized tests to allow us to conduct moderator analysis and the results mirrored those including all academic outcomes: the Type of Measure was a significant moderator, $Q(1) = 22.58$, $p < .001$, $B = 0.77$, with a larger ES for researcher-developed tests ($g = 0.93$, $p < .001$, $k = 93$ from 24 studies) than standardized tests ($g = 0.13$, $p = .015$, $k = 34$ from 19 studies).

Type of control group One may hypothesize that the type of control group implemented in an intervention study would have an impact on the apparent effectiveness of the intervention, with the expectation that active control groups may show smaller effects than passive control groups. In the current dataset, 36 studies (contributing 154 ESs) employed an active control group and 31 studies (contributing 195 ESs) employed a passive control group. Interestingly, moderator analyses indicated that the type of control group did not moderate the size of the effect.

Allocation to group A similar null effect was observed when analysing the impact of how participants were allocated to experimental or control groups. In the current dataset, 48 studies (contributing 247 ESs) allocated participants (or whole classes) randomly to groups, whereas 13 studies (contributing 71 ESs) allocated non-randomly. Six studies did not provide sufficient information about the procedure to code this and their data were not included in these analyses. This factor did not emerge as a significant moderator of the effect size.

Discussion

The aim of the present meta-analytic study was to examine the effectiveness of metacognition interventions on improving typically developing children's outcomes related to self-regulated learning and academic achievement. We explored children's age as well as structural, content-related characteristics as moderators. Structural characteristics included the environment in which the intervention was conducted, mode of delivery, group size, and intensity of the intervention. The content-related characteristics considered whether metacognition was a primary or secondary focus in the intervention, what components of metacognition the intervention targeted, and the type of instructional methods that were used for delivery. Further, we considered three methodological features that may be expected to impact the size of the effect in intervention studies – the types of outcome measures, the type of control group selected and the method of allocating participants to groups. The meta-analytic study contains 349 effect sizes from 67 studies and the effectiveness of the metacognition interventions was evaluated both for all outcomes, and separately for different outcome constructs.

Heterogeneity of effects for General outcomes was high indicating a high amount of unexplained variance, predominantly on the level of outcome measures and studies. The assessment of publication bias indicated a risk of publication bias or selective reporting in the dataset (elaborated upon in Limitations). Moderator analyses were employed to answer our research questions and to explain the high levels of heterogeneity in the dataset. However, moderator analyses contributed only minimally to reducing heterogeneity and it should be noted that moderators only had the effect of enhancing the effect size; there was almost no evidence of significantly negative effects in these analyses. Nevertheless, this meta-analytic study provided novel insights into the features and effectiveness of the most recent attempts to promote children's metacognition, which are summarized in Table 4 and in the following.

Metacognition interventions had a positive impact on child outcomes

The current meta-analytic study confirmed that metacognition interventions are effective for typically developing children revealing a weighted average effect of $g = 0.48$ across SRL-related (Metacognition, Self-efficacy, Motivational orientations, Learning strategies, Executive functions) and academic outcomes (Language, Mathematics, Other academic subjects). Within the SRL-related outcomes, metacognition interventions had the numerically largest impact on Metacognition, followed by Learning strategies and Executive functions. We did not find evidence that metacognition interventions had a significant effect on Self-efficacy and Motivational orientations.

No previous meta-analysis differentiated between Metacognition, Learning strategies and Executive functions. Typically, interventions were explored for their effectiveness on broader outcome categories such as self-regulated learning or cognitive and metacognitive strategies (Dignath & Büttner, 2008; Dignath et al., 2008, 2023). The present findings are nevertheless consistent with these previous meta-analyses (Dignath & Büttner, 2008; Dignath et al., 2008, 2023) that also found positive impacts of metacognition interventions on outcomes related to self-regulated learning. The inclusion of Executive functions as a separate outcome construct is a novel aspect of the current meta-analytic study. Children's executive functions have been considered as precursors to their metacognitive skills (Roebers, 2017). The current results indicate that metacognition interventions may also enhance executive functions, suggesting that the influence could be bidirectional.

In contrast to other meta-analyses (e.g., Dignath & Büttner, 2008; Dignath et al., 2008, 2023), the current meta-analytic study did not provide evidence that metacognition interventions lead to improved motivation, at least in the immediate post-test assessment (see section below on the follow-up assessment). A possible explanation for this could be that in the current study motivation was subdivided into Self-efficacy and Motivational orientations, and at least the latter grouping was rather heterogenous. Young learners tend to overestimate their performance (Destan & Roebers, 2015) and in becoming more metacognitive, their evaluations of their own skills might also become more realistic. Consequently, they might become more aware of certain weaknesses, which may in the short-term lead to lower levels of confidence in their own skills and lower self-efficacy. Similarly, as children experience that learning is effortful and progress can be slow, their intrinsic motivation might stagnate, resulting in no immediate change in motivational orientations.

A positive effect of metacognition interventions was found on the academic outcome categories Language, Mathematics, and Other academic subjects. The numerically largest effect was observed for Other academic subjects followed by Mathematics, and finally Language. This finding is comparable to the findings of previous meta-analyses on the impact of self-regulated learning interventions (Dignath et al., 2008; Dignath & Büttner, 2008; but see Donker et al., 2014) where the largest effects were observed for Mathematics. A possible explanation for why larger effects were found for Other academic subjects and Mathematics than Language outcomes is that teachers and children may less intuitively use metacognitive strategies in subjects such as math, science, and foreign language learning. In subjects related to reading or writing, metacognitive strategies might be part of the regular instruction at school already and there may be less room for improvement in these subjects. Thus, children's performance in math and other academic subjects might benefit more from metacognition interventions. Metacognition interventions targeting these domains appear to be deserving of more attention.

Sustained positive effects were found for self-efficacy and language at follow-up

As well as examining the immediate impact of metacognition interventions, the current study considered the longer-term effectiveness of interventions. This indicated that the effectiveness of metacognition interventions sustained to long-term changes, $g = 0.29$. Given the small number of studies that included a follow-up assessment, only five outcome constructs could be analyzed separately: Metacognition, Self-efficacy, Motivational orientations, Learning strategies and Language. Interestingly, metacognition interventions seemed to have a significant impact on Self-efficacy outcomes at follow-up even though there was no immediate effect of metacognition interventions on Self-efficacy. A possible explanation could be that while metacognition interventions lead to a temporarily more critical representation of one's own skills, with extended practice applying metacognitive strategies, the learner may observe improvements in their performance and become more confident. However, this should be interpreted with caution due to a study being identified as an influential case. Metacognition interventions appear to pay off immediately and result in sustained improvements in Language. More studies with longer-term follow-ups are needed to evaluate whether this is a general pattern that is also true for other academic outcome constructs.

Young children of different age groups did not benefit differently from metacognition interventions

Previous meta-analyses that included wide age ranges (children, adolescents, adults) found that self-regulated learning interventions had a larger impact on children in comparison to adolescents or adults (Dignath & Büttner, 2008; Hattie et al., 1996). Dignath et al. (2008) included a comparable age range as in the current study (elementary school children) and did not find that grade level was a significant moderator of effect sizes, except for strategy use where children in lower grades benefitted more. In the current meta-analysis, age of the sample had very little impact on overall child outcomes. Across all outcome constructs, age was only a significant moderator for two – Motivational orientations and Learning strategies – with opposing impacts of age. However, these findings should be interpreted with caution due to limited variation in the age ranges sampled in studies included in this meta-analytic study. On balance it seems that in pre- and elementary school children benefit similarly from metacognition interventions regardless of age.

The structural characteristics mode of delivery, group size and duration of intervention may enhance the effectiveness of interventions for some outcomes

In the current study, mode of delivery appeared as a significant moderator for Metacognition and Learning strategy outcomes. That is, effectiveness was highest when the intervention was delivered by task materials or by a combination of task materials and teachers, followed by teachers and finally researchers. Previous meta-analyses that focused on SRL indicated that interventions are more effective if conducted by researchers (Dignath & Büttner, 2008; Dignath et al., 2008). Thus, this finding is surprising and contrasts with previous results. A possible explanation could be that the mode of delivery is conflated with intensity. That is, longer, more diffuse interventions might be delivered by teachers whereas shorter and more concentrated interventions might be delivered by researchers. Furthermore, teachers may also incorporate aspects of the intervention throughout the school day which might increase its effectiveness. Closer examination of the dataset did not provide support for this explanation, and the fact that intensity of the intervention was not a significant moderator for metacognition outcomes and that duration was a negative moderator of Learning strategies also makes this explanation unlikely. Instead, our explanation is that more recently designed interventions are better aligned with current developments in teacher education, with teachers having more agency when delivering the intervention, leading to more buy-in and therefore positive effects. Indeed, there is evidence that teachers can vary in their degree of buy-in to new programs they should implement (Bitan-Friedlander et al., 2004) and that this can influence the effectiveness of such programs (Andrzejewski et al., 2016). For the outcome construct Learning strategies there was also an indication that interventions delivered to the whole class were more effective than those delivered in small groups, although the unequal distribution of effect sizes across studies should be kept in mind when interpreting this finding. In terms of intensity of the intervention, as mentioned above intervention impact decreased as the duration of the intervention increased for the outcome construct Learning strategies. This suggests that for metacognition interventions included in this dataset, it was not necessarily the case that "more is better" for enhancing Learning strategies.

A primary focus on metacognition was only beneficial for mathematics

Metacognition interventions are typically applied in the context of educational contexts such as reading, writing, or math (Roebers, 2017). Aspects of metacognition can be heavily featured and prompted within tasks (primary focus) or the focus can be on the academic subject and aspects of metacognition are only implemented on the side of the academic tasks (secondary focus). Approximately half of the studies in the present dataset had metacognition as their primary focus and half of the studies had metacognition as a secondary focus. Considering the different outcome constructs, the moderator focus (primary vs. secondary focus) was significant for Metacognition and Mathematics. Surprisingly, and contrary to our hypothesis, a primary focus on metacognition resulted in a smaller impact of interventions on metacognition outcomes. One would expect that an intervention that prominently features metacognitive skills and repeatedly reminds children of metacognitive strategies would in fact improve metacognition more than metacognition interventions that focused more on other aspects and only promoted metacognitive skills sometimes. Perhaps this finding reflects the fact that metacognitive skills in this age group of children are best acquired when closely related to a specific learning activity. For Mathematics outcomes, interventions with a primary focus on metacognition were more effective than those with metacognition as a secondary focus. It should be acknowledged that coders had to rely on the description of the intervention that was provided in the publications when classifying interventions as having a primary or secondary focus on metacognition; these descriptions varied in the level of detail provided which may have limited the predictive power of this moderator. Certainly, this seems like a potentially promising feature for future reviews and meta-analyses to consider.

Intervention content had limited impact on the effectiveness of interventions

By coding which components of metacognition were targeted in metacognition interventions, we were able to establish that in the current dataset the majority of studies targeted multiple metacognition components and very few studies targeted only one component of metacognition. The most commonly targeted component was metacognitive monitoring, followed by planning, reflection, and metacognitive control. The current set of intervention studies targeted various combinations of metacognition components and the number of studies for each combination was too small for a meaningful comparison. Thus, studies that included a specific metacognition component (e.g., planning) were compared to studies that did not include this component.

The inclusion or exclusion of three metacognition components significantly affected the effect sizes for specific outcome constructs. First, interventions that targeted metacognitive monitoring resulted in lower effect sizes for general outcomes, Metacognition and Learning strategy outcomes than those that did not target monitoring. This finding, however, should be interpreted with caution as the majority of studies (82%) targeted monitoring in their interventions. These findings also stand in contrast to Dignath et al. (2008) who found that monitoring alone had the largest effect size on reading and writing. Similarly, in that study, interventions that focused on monitoring and planning had large effect sizes on the outcome variables 'other performance' and 'strategy use' (Dignath et al., 2008). In our study, a similar pattern as for monitoring was observed for the inclusion of the metacognition component control – this was associated with smaller effect sizes for Learning strategies compared to studies that did not target metacognitive control. One may speculate that these findings occur because when the intervention targets on-task metacognitive skills monitoring and control more, less of the intervention can target the metacognitive behaviors that occur during before or after a learning task (namely, planning and reflection). Perhaps such preparatory behaviors are more specifically associated with outcome measures we categorized as Learning strategies (e.g., time management). In contrast, interventions that targeted the metacognition component reflection more effectively enhanced Language outcomes than those that did not. Interestingly, and in contrast to Dignath et al. (2008), we did not find significant effects for planning. Dignath et al. (2008) found that planning alone led to a high effect size for reading and writing and planning alone and in combination with monitoring led to a high effect size for 'other performance'.

In the related field of self-regulated learning, Wang and Sperling's (2020) review showed that most interventions targeted a combination of strategies including cognitive, metacognitive, and motivational strategies which is consistent with our findings indicating that interventions targeted various metacognition components. Prior evidence indicated that studies that used a combination of strategies were more effective than studies that only included a single strategy (Dignath et al., 2008; Wang & Sperling, 2020). This, however, conflicts with our findings as the number of metacognition components targeted mostly did not emerge as a significant moderator of the effectiveness of interventions, and when it did it was a negative moderator of the effect size for Learning strategies.

One possible reason why intervention content (i.e. which metacognition components were targeted) did not more consistently emerge as a significant moderator in our meta-analyses might be related to the heterogeneity of the studies that we included. Other meta-analyses and reviews focused on a specific aspect of metacognitive monitoring interventions (e.g., tools to foster monitoring; Dignath et al., 2023) or explored effects for selected child outcomes such as academic performance (Donker et al., 2014) or math (Wang & Sperling, 2020). Furthermore, other meta-analyses included only studies that were conducted in the classroom (Dignath & Büttner, 2008; Dignath et al., 2008), whereas our meta-analytic study included studies that were conducted in highly controlled as well as naturalistic classroom environments.

Providing feedback on children's metacognitive behaviors is an impactful instructional method

Another content-related characteristic that was analyzed in this meta-analytic study was the instructional methods applied in different interventions. Like the coding of targeted metacognition components, this was often challenging to code based on the information provided in publications. Indeed, some studies could not be included in these analyses due to a lack of information regarding instructional methods employed. Sharing a detailed intervention protocol is crucial and should be common practice in the spirit of transparency and replicability. Most studies for which instructional method could be coded indeed applied a variety of instructional methods. Of note, in the current dataset, metacognition interventions that included informed instruction were less effective in promoting outcomes related to Metacognition in comparison to interventions that did not include informed instruction. In contrast, metacognition interventions that provided feedback on the metacognitive behaviors of the children were more effective in enhancing Metacognition and Language outcomes than those that did not incorporate feedback. However, it needs to be noted that all the studies that included informed instruction or feedback also included other instructional methods such as practicing the application of the skill. Thus, given the dataset analyzed, a combination of instructional methods might be particularly beneficial. Still, there was no evidence that including more instructional methods was more effective than including fewer instructional methods. To our knowledge, no previous meta-analysis has examined if the applied instructional methods impact the effectiveness of metacognition interventions. This is surprising, as it is crucially important for pedagogical practice.

The type of outcome measure adopted impacts the apparent effectiveness of interventions

To understand if methodological features moderated our effects, we evaluated the role of types of measures, control group design, and group allocation. While the methodological features control group (active vs. passive control group) and group allocation (random vs. non-random allocation) did not emerge as significant moderators, the way in which outcomes were measured did. Offline measures of Learning strategies resulted in larger effect sizes than online measures of Learning strategies, but the unequal number of effect sizes contributing to each level should be considered before interpreting this finding too strongly. With regards to academic outcomes, significantly larger effect sizes were observed for studies that used researcher-developed tests than those that used standardized tests. This finding is in line with a previous meta-analysis and suggests that researcher-developed tests are more sensitive than standardized tests to capture intervention-related changes (Donker et al., 2014). However, the use of researcher-developed tests also raises the question of objectivity and comparability (Donker et al., 2014).

Limitations and future directions

The findings and interpretations of our meta-analytic study are limited in certain respects. Most studies of the present meta-analytic study included participants from Europe or North America. Thus, learning experiences of most of the world's populations are not represented and the results need to be considered under the caveat of the restricted sample. Furthermore, our sample excluded studies that specifically targeted children with disabilities, and we did not code for sample characteristics such as socioeconomic background, ethnicity, or achievement level of the sample. Consequently, it is unknown whether the present findings are equally applicable to all groups of children. Thus, future meta-analyses should capture more detailed information about the sample so that results might be analyzed taking specific groups into account.

The dataset used in the present study includes metacognition interventions that were conducted in naturalistic classroom and controlled environments and included classroom-based as well as computerized-interventions. As such, the meta-analytic study is representative of the metacognition interventions that are currently available in the field. However, the heterogeneity of the metacognition interventions was also high, which may have limited our findings and left a lot of variance across effect sizes unexplained.

There was also a lot of heterogeneity in terms of the content and instructional methods of the metacognition interventions. Interventions were often implemented as comprehensive approaches that included a combination of different components of metacognition and instructional methods across different school subjects making it challenging to determine features that were responsible for improved child outcomes. Indeed, some interventions targeted not only components of metacognition but included also other cognitive or motivational constructs. Future meta-analyses should systematically test features of metacognition interventions against one another in the same school subject. With such an approach it could be determined if some features are more effective or if they need to be applied in combination with one another. Such findings would be highly relevant for pedagogical practice as demonstrated in the next section. More generally, we hope that the patterns evidenced in our moderator analyses may help to inform future reviews and meta-analyses in this field about potentially interesting factors to consider as well as to help inform the development of effective interventions. For this, the interested reader should consult the size of the effects for different moderator levels and their CIs to weigh up the evidence this meta-analysis provides about a specific feature rather relying solely on the significance of moderator tests.

The analyses we conducted to assess publication bias and selective reporting produced inconsistent results. Whereas the standard measure, an adapted version of Egger's regression test, was significant indicating significant funnel plot asymmetry, a follow-up moderation analysis indicated no significant differences between effect sizes stemming from published and unpublished data. One reason for these perhaps contradictory findings might be due to a limitation in power for testing this moderation as the number of effect sizes from unpublished data was rather small. However, it should also be noted that the Egger's regression test is limited in that it does not exclusively test for publication bias or selective reporting (Page et al., 2021; Rodgers & Pustejovsky, 2021). There are several possible reasons that effect sizes and their precision might be correlated, which would result in a significant Egger's test, such as high between-study heterogeneity, low training fidelity, or even chance (Harrer et al., 2021; Page et al., 2021; Rodgers & Pustejovsky., 2021). Because of this, and given that the targeted moderation analysis for differences by publication status indicated no publication bias, we choose not to interpret the significant result of the Egger's regression test as indicating a strong publication bias.

Implications for pedagogical practice

The present meta-analytic study adds to existing evidence that promoting metacognition is beneficial for outcomes related to self-regulated learning and academic achievement. In contrast to previous meta-analyses, we found that children's metacognition can be best enhanced by teachers and/or task materials. Thus, researchers should not shy away from developing metacognition interventions for teachers. Classroom observations in German elementary schools showed that teachers spent little time teaching metacognitive strategies (Dignath & Büttner, 2018). As suggested by Dignath and Büttner (2018) this could imply that teachers either do not know how to teach metacognitive strategies, they might think that these strategies are not important, or they might not be aware that it is necessary to teach these strategies. Our findings demonstrate that teachers can promote metacognition effectively, but they might need support in how to teach these skills.

In the present meta-analysis, metacognition interventions that had a primary or secondary focus were equally beneficial for child outcomes, with the exception of two outcome constructs. For Mathematics it was more beneficial if metacognition was the primary focus of the intervention, for Metacognition it was more beneficial if metacognition was the secondary focus. The latter finding can be interpreted positively as it may suggest that metacognition strategies can be implemented alongside other activities, and they do not need to be the primary focus. Similarly, we did not find consistent evidence that a specific intervention content (planning, monitoring, control, reflection) or instructional method (with the exception of feedback) was more beneficial than others. The provision of feedback on children's metacognition emerged as a significant moderator of the effectiveness of interventions, suggesting young learners can benefit from external support to refine their newly acquired metacognitive skills, behaviors and strategies. Overall, the pattern of results suggests that there are various pathways for metacognition interventions to enhance child outcomes. As most of the interventions are multi-component approaches addressing a variety of content and instructional methods, it can be difficult to discern if there are active ingredients that might drive any observed changes. Furthermore, the implementation of a comprehensive intervention is difficult to scale due to time constraints and the efforts involved. In fact, recent developments in intervention research are moving towards a more dynamic and flexible so-called 'common elements' approach in which activities that are common across effective studies are identified (e.g., Clarke et al., 2022). These activities are then presented to teachers who can choose the activities that are applicable to their context and needs. This approach is easier to implement, less top-down, and more democratic as it provides teachers agency in deciding what the children in their classroom need most. Efforts in a conceptually related field, self-regulation, provide successful examples (e.g., Barnes et al., 2021; Raggio Colagrossi et al., 2023). This could be a promising future direction for the field of metacognition to explore.

Implications for theories of self-regulated learning

Select findings from this meta-analytic study on metacognition interventions also have the potential to inform the broader theories of self-regulated learning, specifically regarding how the different subdomains of self-regulated learning relate to one other. This issue is typically assessed in correlational studies and predominantly using questionnaires to assess the three main components of self-regulated learning (i.e., cognitive, motivational and metacognitive components). Here, based on contemporary intervention studies that included direct assessments as well as questionnaire methods, it could be established that interventions that target metacognition also improve executive functions (belonging to the cognitive subdomain) and that training metacognition only improves self-efficacy (belonging to the motivational subdomain) in a delayed fashion. The former finding is notable, as typically the opposite direction of causation is assumed in the relationship between metacognitive and executive skills (Bryce et al., 2015; Fernandez-Duque et al., 2000; Wacker & Roebers, 2022). While this interpretation must be made with caution, as some of these intervention studies (Carretti et al., 2014; Cornoldi et al., 2015; Partanen et al., 2015; Pozuelos et al., 2019) additionally included some executive function training, this finding suggests that the mechanisms by which executive functions and metacognition are related is deserving of renewed consideration. Regarding the relationship between metacognitive and motivational components of self-regulated learning, while motivation is undoubtedly required for children to engage in effortful metacognitive approaches to learning, the current findings also suggest that with time the benefits of metacognition can feed forward to motivation and result in children who have more self-belief in their own abilities. In the ideal case, this may result in a positive feedback loop whereby children enjoy boosts in motivation that allow them to continue behaving metacognitively. A meta-analysis examining the impact of SRL interventions on university students' outcomes also found a positive impact on motivational outcomes, namely on intrinsic motivation and interest as well as self-efficacy (Theobald, 2021). That study found that teacher-feedback was particularly important for the promotion of motivational outcomes suggesting that teachers can support the promotion of motivational outcomes, at least in older learners.

Conclusions

The current meta-analytic study adds to the body of evidence that young learners can benefit from metacognition interventions. The evidence provided by this study shows that teachers like the one in our example, who noticed that many of her third graders were lacking a metacognitive approach to learning, can effectively support children in developing these skills. If she takes steps to promote metacognition in her classroom, she can expect pupils to show gains both in academic subjects (such as reading and writing, math, science, second-language learning) and overarching skills related to self-regulated learning (such as metacognition, learning strategies, and executive functions). Although she might observe no immediate effects on the self-efficacy of her students, the evidence suggests students' self-efficacy will improve later as they notice that their newly enhanced metacognitive skills improve their performance and allow them to take control of their learning. Based on the current results, our teacher can be creative with how she integrates the promotion of metacognition and academic tasks, as there was no consistent evidence across all outcome constructs that a primary focus on metacognition is necessary for positive effects. Likewise, she can feel empowered to employ a range of instructional methods such as modelling a metacognitive approach to learning tasks, prompting students to monitor and control their learning, and providing feedback on students' metacognitive behaviors. Importantly, the current results are consistent with claims made by others in the literature that promoting metacognition in young learners can be fun and engaging, low-cost, and integrated with real learning tasks already implemented in classroom teaching (Perry et al., 2019).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11409-024-09405-x.

Acknowledgements The authors would like to thank Celine Ina Spannagl for her support during the coding process.

Authors' contributions The study was conceptualized, designed and supervised by DB and JE; literature search and coding of studies was conducted primarily by JE and FS with support from DB; data analysis was led by DB with support from JE; data validation was completed by all three authors; visualizations were produced by DB and FS; JE and DB contributed equally to the writing of the manuscript with substantial contributions from FS.

Funding Open Access funding enabled and organized by Projekt DEAL. This study was supported by the Deutsche Forschungsgemeinschaft (BR 6057/3–1).

Data availability Please refer to this link for the preprint of the manuscript: https://osf.io/preprints/psyarxiv/475br.

Please refer to this link for our preregistration, the datasets and the data analysis script: https://osf.io/bjkxy.

Declarations

Ethics approval As this manuscript reports a meta-analytic study, there are no ethical issues and ethical approval is not applicable.

Competing interests None of the authors have any competing interests.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Andrzejewski, C. E., Davis, H. A., Shalter Bruening, P., & Poirier, R. R. (2016). Can a self-regulated strategy intervention close the achievement gap? Exploring a classroom-based intervention in 9th grade earth science. *Learning and Individual Differences, 49*, 85–99. https://doi.org/10.1016/j.lindif.2016.05.013

*Arrimada, M., Torrance, M., & Fidalgo, R. (2019). Effects of teaching planning strategies to first-grade writers. *British Journal of Educational Psychology, 89*(4), 670–688. https://doi.org/10.1111/bjep.12251

*Artuso, C., Carretti, B., & Palladino, P. (2019). Short-term training on working memory updating and metacognition in primary school: The effect on reading comprehension. *School Psychology International, 40*(6), 641–657. https://doi.org/10.1177/0143034319881671

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

Barnes, S. P., Bailey, R., & Jones, S. M. (2021). Evaluating the impact of a targeted approach designed to build executive function skills: A randomized trial of brain games. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.655246

Bitan-Friedlander, N., Dreyfus, A., & Milgrom, Z. (2004). Types of "teachers in training": The reactions of primary school science teachers when confronted with the task of implementing an innovation. *Teaching and Teacher Education, 20*(6), 607–619. https://doi.org/10.1016/j.tate.2004.06.007

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research, 31*(6), 445–457. https://doi.org/10.1016/S0883-0355(99)00014-2

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley & Sons.

*Boykin, A., Evmenova, A. S., Regan, K., & Mastropieri, M. (2019). The impact of a computer-based graphic organizer with embedded self-regulated learning strategies on the argumentative writing of students in inclusive cross-curricula settings. *Computers & Education, 137*, 78–90. https://doi.org/10.1016/j.compedu.2019.03.008

Brown, A. L. (1987). Metacognition, executive control, self-regulation and other more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp. 65–116). Erlbaum.

*Brunstein, J. C., & Glaser, C. (2011). Testing a path-analytic mediation model of how self-regulated writing strategies improve fourth graders' composition skills: A randomized controlled trial. *Journal of Educational Psychology, 103*(4), 922–938. https://doi.org/10.1037/a0024622

Bryce, D., Whitebread, D., & Szűcs, D. (2015). The relationships among executive functions, metacognitive skills and educational achievement in 5 and 7 year-old children. *Metacognition and Learning, 10*(2), 181–198. https://doi.org/10.1007/s11409-014-9120-4

*Bulu, S. T., & Pedersen, S. (2010). Scaffolding middle school students' content knowledge and ill-structured problem solving in a problem-based hypermedia learning environment. *Educational Technology Research and Development, 58*(5), 507–529. https://doi.org/10.1007/s11423-010-9150-9

*Camacho, A., Alves, R. A., Silva, M., Ferreira, P., Correia, N., & Daniel, J. R. (2023). The impact of combining SRSD instruction with a brief growth mindset intervention on sixth graders' writing motivation and performance. *Contemporary Educational Psychology, 72*, 1–15. https://doi.org/10.1016/j.cedpsych.2022.102127

*Carretti, B., Caldarola, N., Tencati, C., & Cornoldi, C. (2014). Improving reading comprehension in reading and listening settings: The effect of two training programmes focusing on metacognition and working memory. *The British Journal of Educational Psychology, 84*(Pt 2), 194–210. https://doi.org/10.1111/bjep.12022

*Carroll, A. N. (2012). *The effects of training in self-regulated learning and achievement orientations in lower socioeconomic elementary students* [Dissertation, Louisiana Tech University]. https://digitalcommons.latech.edu/dissertations/334

Clarke, A., Baker, S., Ghiara, V., Burridge, H., Davie, P., Eberhart, J., Fischer, F., & Jackson, A. (2022). *Common elements: An innovative approach to improving children's outcomes in early childhood education*. Early Intervention Foundation. https://www.eif.org.uk/report/common-elements-an-innovative-approach-to-improving-childrens-outcomes-in-early-childhood-education

*Connor, C. M., Phillips, B. M., Kim, Y.-S. G., Lonigan, C. J., Kaschak, M. P., Crowe, E., Dombek, J., & Otaiba, S. A. (2018). Examining the efficacy of targeted component interventions on language and literacy for third and fourth graders who are at risk of comprehension difficulties. *Scientific Studies of Reading : The Official Journal of the Society for the Scientific Study of Reading, 22*(6), 462–484. https://doi.org/10.1080/10888438.2018.1481409

*Cordewener, K. A. H., Hasselman, F., Verhoeven, L., & Bosman, A. M. T. (2018). The role of instruction for spelling performance and spelling consciousness. *The Journal of Experimental Education, 86*(2), 135–153. https://doi.org/10.1080/00220973.2017.1315711

*Cornoldi, C., Carretti, B., Drusi, S., & Tencati, C. (2015). Improving problem solving in primary school students: The effect of a training programme focusing on metacognition and working memory. *The British Journal of Educational Psychology, 85*(3), 424–439. https://doi.org/10.1111/bjep.12083

*Cosentino, C. L. (2017). *The effects of self-regulation strategies on reading comprehension, motivation for learning, and self-efficacy with struggling readers*. https://api.semanticscholar.org/CorpusID:149052974

de Boer, H., Donker-Bergstra, A., Kostons, D. D. N. M., Korpershoek, H., & Werf, M. (2013). *Effective Strategies for Self-regulated Learning*.

*de Vreeze-Westgeest, M. G. J., & Vogelaar, B. (2022). Cognitive training in the domain of mathematics for potentially gifted children in primary school. *Education Sciences, 12*(2), 1–13. https://doi.org/10.3390/educsci12020127

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review, 28*(3), 425–474. https://doi.org/10.1007/s10648-015-9320-8

Destan, N., & Roebers, C. M. (2015). What are the metacognitive costs of young children's overconfidence? *Metacognition and Learning, 10*(3), 347–374. https://doi.org/10.1007/s11409-014-9133-z

Dignath, C., Buettner, G., & Langfeldt, H. P. (2008). How can primary school students learn self-regulated learning strategies most effectively?. A meta-analysis on self-regulation training programmes. *Educational Research Review, 3*(2), 101–129. https://doi.org/10.1016/j.edurev.2008.02.003

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning, 3*(3), 231–264. https://doi.org/10.1007/s11409-008-9029-x

Dignath, C., & Büttner, G. (2018). Teachers' direct and indirect promotion of self-regulated learning in primary and secondary school mathematics classes – insights from video-based classroom observations and teacher interviews. *Metacognition and Learning, 13*(2), 127–157. https://doi.org/10.1007/s11409-018-9181-x

Dignath, C., van Ewijk, R., Perels, F., & Fabriz, S. (2023). Let learners monitor the learning content and their learning behavior! A meta-analysis on the effectiveness of tools to foster monitoring. *Educational Psychology Review, 35*(2), 62. https://doi.org/10.1007/s10648-023-09718-4

Donker, A. S., de Boer, H., Kostons, D., Dignath van Ewijk, C. C., & van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review, 11*, 1–26. https://doi.org/10.1016/j.edurev.2013.11.002

*Dresel, M., & Haugwitz, M. (2008). A computer-based approach to fostering motivation and self-regulated learning. *The Journal of Experimental Education, 77*(1), 3–20. https://doi.org/10.3200/JEXE.77.1.3-20

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

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

*Eyupoglu, T. F. (2023). *The impact of metacognitive strategy training in a game-based learning environment for fifth-grade science* [Doctoral Dissertation, North Carolina State University]. https://repository.lib.ncsu.edu/bitstream/1840.20/40275/1/etd.pdf

Fernández-Castilla, B., Jamshidi, L., Declercq, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2020). The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. *Behavior Research Methods, 52*(5), 2031–2052. https://doi.org/10.3758/s13428-020-01373-9

Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive Attention and Metacognitive Regulation. *Consciousness and Cognition, 9*(2), 288–307. https://doi.org/10.1006/ccog.2000.0447

*Ferreira, P. C., Simão, A. M. V., & da Silva, A. L. (2015). Does training in how to regulate one's learning affect how students report self-regulated learning in diary tasks? *Metacognition and Learning, 10*(2), 199–230. https://doi.org/10.1007/s11409-014-9121-3

*Ferreira, P. C., Simão, A. M. V., & da Silva, A. L. (2017). How and with what accuracy do children report self-regulated learning in contemporary EFL instructional settings? *European Journal of Psychology of Education, 32*(4), 589–615. https://doi.org/10.1007/s10212-016-0313-x

*Finlayson, K., & McCrudden, M. T. (2022). Teacher-implemented self-regulated strategy development instruction for story writing with year 2 students in Aotearoa/New Zealand: A mixed methods study. *Teaching and Teacher Education, 119*, 1–11. https://doi.org/10.1016/j.tate.2022.103846

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

*Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., & Schroeter, K. (2003). Enhancing third-grade students' mathematical problem solving with self-regulated learning strategies. *Journal of Educational Psychology, 95*(2), 306–315. https://doi.org/10.1037/0022-0663.95.2.306

*Gaskill, P. J. (2002). *Effects of a goal-setting strategy on second graders' self-efficacy for a listening task* [Dissertation]. The Ohio State University.

*Glaser, C., & Brunstein, J. C. (2007). Improving fourth-grade students' composition skills: Effects of strategy instruction and self-regulation procedures. *Journal of Educational Psychology, 99*(2), 297–310. https://doi.org/10.1037/0022-0663.99.2.297

Harrer, M., Cuijpers, P., Furukawa, T., & Ebert, D. (2021). *Doing meta-analysis in R: A hands-on guide*. Chapman & Hall/CRC Press.

*Harris, K. R., Kim, Y.-S., Yim, S., Camping, A., & Graham, S. (2023). Yes, they can: Developing transcription skills and oral language in tandem with SRSD instruction on close reading of science text to write informative essays at grades 1 and 2. *Contemporary Educational Psychology, 73*, 1–14. https://doi.org/10.1016/j.cedpsych.2023.102150

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66*(2), 99–136. https://doi.org/10.3102/00346543066002099

*Hoffmann, K. F. (2010). *The impact of graphic organizer and metacognitive monitoring instruction on expository science text comprehension in fifth grade students* [Doctoral dissertation, North Carolina State University]. North Carolina State University.

*Huff, J. D., & Nietfeld, J. L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning, 4*(2), 161–176. https://doi.org/10.1007/s11409-009-9042-8

*Iordanou, K. (2022). Supporting strategic and meta-strategic development of argument skill: The role of reflection. *Metacognition and Learning, 17*(2), 399–425. https://doi.org/10.1007/s11409-021-09289-1

*Jacob, L., Benick, M., Dörrenbächer, S., & Perels, F. (2020). Promoting self-regulated learning in preschoolers. *Journal of Childhood, Education & Society*, *1*(2), 116–140. https://doi.org/10.37291/2717638X.20201237

*Jacobse, A. E., & Harskamp, E. G. (2009). Student-controlled metacognitive training for solving word problems in primary school mathematics. *Educational Research and Evaluation, 15*(5), 447–463. https://doi.org/10.1080/13803610903444519

Kälin, S., & Roebers, C. M. (2022). Longitudinal associations between executive functions and metacognitive monitoring in 5- to 8-year-olds. *Metacognition and Learning, 17*, 1079–1095. https://doi.org/10.1007/s11409-022-09306-x

Kim, Y., Zepeda, C. D., & Butler, A. C. (2023). An interdisciplinary review of self-regulation of learning: Bridging cognitive and educational psychology perspectives. *Educational Psychology Review*, *35*(3), Article 92. https://doi.org/10.1007/s10648-023-09800-x

Kistner, S., Rakoczy, K., Otto, B., Dignath-van Ewijk, C., Büttner, G., & Klieme, E. (2010). Promotion of self-regulated learning in classrooms: Investigating frequency, quality, and consequences for student performance. *Metacognition and Learning, 5*(2), 157–171. https://doi.org/10.1007/s11409-010-9055-3

*Kruit, P. M., Oostdam, R. J., van den Berg, E., & Schuitema, J. A. (2018). Effects of explicit instruction on the acquisition of students' science inquiry skills in grades 5 and 6 of primary education. *International Journal of Science Education, 40*(4), 421–441. https://doi.org/10.1080/09500693.2018.1428777

Lai, C.-L., & Hwang, G.-J. (2016). A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. *Computers & Education, 100*, 126–140. https://doi.org/10.1016/j.compedu.2016.05.006

*Lancaster, J. M. (2018). *Teaching comprehension monitoring strategies within a metacognitive framework: An analysis of the effects on fourth graders' reading achievement and calibration.* [Dissertation, North Carolina State University]. http://www.lib.ncsu.edu/resolver/1840.20/35389

Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-Analysis* (Vol. 49). SAGE Publications.

Llewellyn-Bennett, R., Bowman, L., & Bulbulia, R. (2016). Post-trial follow-up methodology in large randomized controlled trials: A systematic review protocol. *Systematic Reviews, 5*(1), 214–214. https://doi.org/10.1186/s13643-016-0393-3

*Lubliner, S., & Smetana, L. (2005). The effects of comprehensive vocabulary instruction on title I students' metacognitive word-learning skills and reading comprehension. *Journal of Literacy Research, 37*(2), 163–200. https://doi.org/10.1207/s15548430jlr3702_3

*Martelletti, D. M., Luzuriaga, M., & Furman, M. (2023). 'What makes you say so?' Metacognition improves the sustained learning of inferential reading skills in English as a second language. *Trends in Neuroscience and Education, 33*, 100213. https://doi.org/10.1016/j.tine.2023.100213

* Mason, L. H. (2004). Explicit self-regulated strategy development versus reciprocal questioning: Effects on expository reading comprehension among struggling readers. *Journal of Educational Psychology, 96*(2), 283–296. https://doi.org/10.1037/0022-0663.96.2.283

*McKeown, D., Wijekumar, K., Owens, J., Harris, K., Graham, S., Lei, P., & FitzPatrick, E. (2023). Professional development for evidence-based SRSD writing instruction: Elevating fourth grade outcomes. *Contemporary Educational Psychology, 73*, 1–14. https://doi.org/10.1016/j.cedpsych.2023.102152

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Müller, U., & Kerns, K. A. (2015). The development of executive function. In R. M. Lerner (Ed.), *Handbook of child psychology and developmental science* (7th ed., pp. 1–42). John Wiley & Sons Inc.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation, 26*, 125–173. https://doi.org/10.1016/S0079-7421(08)60053-5

*Obergriesser, S., & Stoeger, H. (2015). The role of emotions, motivation, and learning behavior in underachievement and results of an intervention. *High Ability Studies, 26*(1), 167–190. https://doi.org/10.1080/13598139.2015.1043003

Ohtani, K., & Hisasaka, T. (2018). Beyond intelligence: A meta-analytic review of the relationship among metacognition, intelligence, and academic performance. *Metacognition and Learning, 13*(2), 179–212. https://doi.org/10.1007/s11409-018-9183-8

Page, M. J., Sterne, J. A. C., Higgins, J. P. T., & Egger, M. (2021). Investigating and dealing with publication bias and other reporting biases in meta-analyses of health research: A review. *Research Synthesis Methods, 12*(2), 248–259. https://doi.org/10.1002/jrsm.1468

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00422

*Partanen, P., Jansson, B., Lisspers, J., & Sundin, Ö. (2015). Metacognitive strategy training adds to the effects of working memory training in children with special educational needs. *International Journal of Psychological Studies, 7*(3), 130–140. https://doi.org/10.5539/ijps.v7n3p130

*Perels, F., Dignath, C., & Schmitz, B. (2009). Is it possible to improve mathematical achievement by means of self-regulation strategies? Evaluation of an intervention in regular math classes. *European Journal of Psychology of Education, 24*(1), 17–31. https://doi.org/10.1007/BF03173472

Perry, J., Lundie, D., & Golder, G. (2019). Metacognition in schools: What does the literature suggest about the effectiveness of teaching metacognition in schools? *Educational Review, 71*(4), 483–500. https://doi.org/10.1080/00131911.2018.1441127

Pigott, T. (2012). Advances in meta-analysis. *Springer*. https://doi.org/10.1007/978-1-4614-2278-5

*Pozuelos, J. P., Combita, L. M., Abundis, A., Paz-Alonso, P. M., Conejero, Á., Guerra, S., & Rueda, M. R. (2019). Metacognitive scaffolding boosts cognitive and neural benefits following executive attention training in children. *Developmental Science, 22*(2), e12756. https://doi.org/10.1111/desc.12756

Quigley, A., Muijs, D., & Stringer, E. (2021). *Metacognition and self-regulated learning. Guidance report*. Education Endowment Foundation.

R Core Team. (2023). *R: A language and environment for statistical computing. R Foundation for Statistical Computing* [Computer software]. https://www.R-project.org/

Raggio Colagrossi, A. L., de Magalhães-Barbosa, M. C., McCoy, D. C., Barnes, S. P., Temko, S., Bailey, R., Jones, S. M., Bianchi, L. M., da Cunha, A. J. L. A., & Prata-Barbosa, A. (2023). Adaptation and efficacy of a social-emotional learning intervention (SEL kernels) in early childhood settings in southeastern Brazil: A quasi-experimental study. *Early Education and Development*, 1–18. https://doi.org/10.1080/10409289.2023.2219583

Rodgers, M. A., & Pustejovsky, J. E. (2021). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods, 26*(2), 141–160. https://doi.org/10.1037/met0000300

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review, 45*, 31–51. https://doi.org/10.1016/j.dr.2017.04.001

*Rosário, P., Núñez, J. C., Vallejo, G., Cunha, J., Azevedo, R., Pereira, R., Nunes, A. R., Fuentes, S., & Moreira, T. (2016). Promoting Gypsy children school engagement: A story-tool project to enhance self-regulated learning. *Indigenous Issues in Education and Research: Looking Forward, 47*, 84–94. https://doi.org/10.1016/j.cedpsych.2015.11.005

*Salas, N., Pascual, M., Birello, M., & Cross, A. (2023). Embedding explicit linguistic instruction in an SRSD writing intervention. *Written Communication, 40*(3), 857–891. https://doi.org/10.1177/07410883231169516

*Schünemann, N., Spörer, N., & Brunstein, J. C. (2013). Integrating self-regulation in whole-class reciprocal teaching: A moderator–mediator analysis of incremental effects on fifth graders' reading comprehension. *Contemporary Educational Psychology, 38*(4), 289–305. https://doi.org/10.1016/j.cedpsych.2013.06.002

*Schuster, C., Stebner, F., Geukes, S., Jansen, M., Leutner, D., & Wirth, J. (2023). The effects of direct and indirect training in metacognitive learning strategies on near and far transfer in self-regulated learning. *Learning and Instruction, 83*, 1–12. https://doi.org/10.1016/j.learninstruc.2022.101708

*Schuster, C., Stebner, F., Leutner, D., & Wirth, J. (2020). Transfer of metacognitive skills in self-regulated learning: An experimental training study. *Metacognition and Learning, 15*(3), 455–477. https://doi.org/10.1007/s11409-020-09237-5

*Souvignier, E., & Mokhlesgerami, J. (2006). Using self-regulation as a framework for implementing strategy instruction to foster reading comprehension. *Learning and Instruction, 16*(1), 57–71. https://doi.org/10.1016/j.learninstruc.2005.12.006

Spiegel, J. A., Goodrich, J. M., Morris, B. M., Osborne, C. M., & Lonigan, C. J. (2021). Relations between executive functions and academic outcomes in elementary school children: A meta-analysis. *Psychological Bulletin, 147*(4), 329–351. https://doi.org/10.1037/bul0000322

*Spörer, N., & Schünemann, N. (2014). Improvements of self-regulation procedures for fifth graders' reading competence: Analyzing effects on reading comprehension, reading strategy performance, and motivation for reading. *Learning and Instruction, 33*, 147–157. https://doi.org/10.1016/j.learninstruc.2014.05.002

*Stoeger, H., Sontag, C., & Ziegler, A. (2014). Impact of a teacher-led intervention on preference for self-regulated learning, finding main ideas in expository texts, and reading comprehension. *Journal of Educational Psychology, 106*(3), 799–814. https://doi.org/10.1037/a0036035

*Stoeger, H., & Ziegler, A. (2005). Evaluation of an elementary classroom self-regulated learning program for gifted mathematics underachievers. *International Education Journal, 6*(2), 261–271.

*Stoeger, H., & Ziegler, A. (2008). Evaluation of a classroom based training to improve self-regulation in time management tasks during homework activities with fourth graders. *Metacognition and Learning, 3*(3), 207–230. https://doi.org/10.1007/s11409-008-9027-z

Stucke, N., & Doebel, S. (2023). Early childhood executive function predicts concurrent and later social and behavioral outcomes: A review and meta-analysis. OSF. https://doi.org/10.31234/osf.io/s59ev

Takacs, Z. K., & Kassai, R. (2019). The efficacy of different interventions to foster children's executive function skills: A series of meta-analyses. *Psychological Bulletin, 145*(7), 653–697. https://doi.org/10.1037/bul0000195

*Teong, S. K. (2003). The effect of metacognitive training on mathematical word-problem solving. *Journal of Computer Assisted Learning, 19*(1), 46–55. https://doi.org/10.1046/j.0266-4909.2003.00005.x

Theobald, M. (2021). Self-regulated learning training programs enhance university students' academic performance, self-regulated learning strategies, and motivation: A meta-analysis. *Contemporary Educational Psychology, 66*, 101976. https://doi.org/10.1016/j.cedpsych.2021.101976

*Tok, Ş. (2013). Effects of the know-want-learn strategy on students' mathematics achievement, anxiety and metacognitive skills. *Metacognition and Learning, 8*(2), 193–212. https://doi.org/10.1007/s11409-013-9101-z

*Tsiriotakis, I. of K. (2013). *Writing difficulties and feelings of anxiety during the acquisition of english as a foreign language* [Dissertation, University of Crete]. http://hdl.handle.net/10442/hedi/33100

*Vandevelde, S., Van Keer, H., & Merchie, E. (2017). The challenge of promoting self-regulated learning among primary school children with a low socioeconomic and immigrant background. *The Journal of Educational Research, 110*(2), 113–139. https://doi.org/10.1080/00220671.2014.999363

Veenman, M., & Elshout, J. J. (1999). Changes in the relation between cognitive and metacognitive skills during the acquisition of expertise. *European Journal of Psychology of Education, 14*(4), 509–523. https://doi.org/10.1007/BF03172976

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

Wacker, S., & Roebers, C. M. (2022). Stop and think: Additional time supports monitoring processes in young children. *PLoS ONE, 17*(9), e0274460. https://doi.org/10.1371/journal.pone.0274460

Wang, Y., & Sperling, R. A. (2020). Characteristics of effective self-regulated learning interventions in mathematics classrooms: A systematic review. *Frontiers in Education*, *5*. https://doi.org/10.3389/feduc.2020.00058

Watts, T. W., Bailey, D. H., & Li, C. (2019). Aiming further: Addressing the need for high-quality longitudinal research in education. *Journal of Research on Educational Effectiveness, 12*(4), 648–658. https://doi.org/10.1080/19345747.2019.1644692

Whitebread, D., Coltman, P., Pino Pasternak, D., Sangster, C., Grau, V., Bingham, S., Almeqdad, Q., & Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition and Learning, 4*(1), 63–85. https://doi.org/10.1007/s11409-008-9033-1

*Williams, W. M., Blythe, T., White, N., Li, J., Gardner, H., & Sternberg, R. J. (2002). Practical intelligence for school: Developing metacognitive sources of achievement in adolescence. *Developmental Review, 22*(2), 162–210. https://doi.org/10.1006/drev.2002.0544

Wilson, S. J., & Lipsey, M. W. (2000). Wilderness challenge programs for delinquent youth: A meta-analysis of outcome evaluations. *Evaluation and Program Planning, 23*(1), 1–12. https://doi.org/10.1016/S0149-7189(99)00040-3

World Health Organization. (2011). *World report on disability*. https://www.who.int/disabilities/world_report/2011/en/

*Wright, J., & Jacobs, B. (2003). Teaching phonological awareness and metacognitive strategies to children with reading difficulties: A comparison of the two instructional methods. *Educational Psychology, 23*(1), 17–45. https://doi.org/10.1080/01443410303217

Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: A meta-analysis. *Asia Pacific Education Review, 17*(2), 187–202. https://doi.org/10.1007/s12564-016-9426-9

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Authors and Affiliations

Janina Eberhart[1] · **Franziska Schäfer**[2] · **Donna Bryce**[1,3]

✉ Janina Eberhart
janina.eberhart@uni-tuebingen.de

Franziska Schäfer
franziska.schaefer1@tu-darmstadt.de

Donna Bryce
donna.bryce@uni-a.de

1 Department of Psychology, University of Tübingen, Schleichstraße 4, 72076 Tübingen, Germany

2 Department of Psychology, Technical University of Darmstadt, Alexanderstr. 10, 64283 Darmstadt, Germany

3 Department of Psychology, University of Augsburg, Universitätsstraße 10, 86159 Augsburg, Germany