Heller *et al. International Journal of STEM Education* (2026) 13:25
https://doi.org/10.1186/s40594-026-00619-w

International Journal of
STEM Education

REVIEW **Open Access**

A second-order meta-analysis of interventions and their effectiveness in STEMM education: effects on motivational, affective, and attitudinal outcomes

Fabian Heller[1*], Charlotte Popp[2], Albert Ziegler[2] and Heidrun Stoeger[1]

Abstract

Background The STEMM sector is increasingly vital, with a growing demand for skilled professionals. Beyond cognitive skills, non-cognitive factors such as motivation, affect, and attitudes have a significant influence on STEMM performance and aspirations. To gain an overview of the field of research, a second-order meta-analysis was conducted on the effectiveness of interventions on non-cognitive outcomes in STEMM education. A systematic search (2000–2023) identified 6664 review studies, from which 42 meta-analyses containing 64 meta-analytic effect sizes were included. These encompassed 1903 primary studies, covering both K-12 and higher education contexts. The interventions were categorized by the targeted outcome, intervention method, and targeted STEMM discipline.

Results Mathematics was the most frequently studied subject, and common interventions included applications of educational technology and student-centered instruction. The most studied outcomes were attitudes towards STEMM and affective factors, such as mathematics anxiety and satisfaction. Overall, interventions showed a significant positive effect on non-cognitive outcomes ($d = 0.42$, 95% CI [0.35, 0.50], $p < .001$), indicating improved motivation, attitudes, or reduced anxiety toward STEMM. Effectiveness did not differ significantly across outcome types or intervention methods, although mathematics and general STEM studies showed slightly smaller effects. We also found substantial variability in the meta-analytic effects. Notably, gaps in the literature were identified. Affective states such as boredom and enjoyment were rarely addressed, and out-of-school contexts were underrepresented.

Conclusion This highlights the need for more comprehensive and diverse meta-analyses. Nonetheless, student-centered teaching, games, and technology-based methods appear promising for enhancing non-cognitive outcomes in STEMM education.

Keywords STEM, STEMM, Meta-analysis, Motivation, Attitudes, Anxiety, K-12, Higher education

---

*Correspondence:
Fabian Heller
fabian.heller@ur.de
[1] Chair of School Research, School Development, and Evaluation, University of Regensburg, Regensburg, Germany
[2] Chair of Educational Psychology and Excellence Research, Friedrich-Alexander-University Erlangen-Nuremberg, Nuremberg, Germany

Introduction

Careers in STEMM fields—science, technology, engineering, mathematics, and medical sciences—are frequently associated with higher salaries (Anger & Plünnecke, 2022; U.S. Bureau of Labor Statistics, 2024) and better employment prospects (U.S. Bureau of Labor Statistics, 2022, 2024) than in other professional sectors. STEMM jobs are essential contributors to a country's economy and its ability to innovate (AAAS, 2025), and critical sectors like energy, health, and digitalization depend on a competent STEMM workforce. Nevertheless, many STEMM disciplines continue to experience a persistent shortage of qualified professionals (Anger et al., 2024). One contributing factor is the underrepresentation of particular demographic groups, including women and several ethnic minorities (National Science Board & National Science Foundation, 2024).

Numerous interventions have been developed and evaluated within STEMM education over the past decades to address, this issue, often implemented during students' school years or in higher education. Many of these interventions focused on cognitive domains, such as building competencies (see Chen et al., 2025), improving academic performance (see Ting et al., 2023), or addressing misconceptions (e.g., Leinonen et al., 2013). However, an increasing body of research emphasizes the relevance of non-cognitive factors in explaining low participation rates in STEMM fields (e.g., Sakellariou & Fang, 2021).

Among the most prominent non-cognitive variables are motivational factors, such as low interest in STEMM subjects (Lawner et al., 2019; Sakellariou & Fang, 2021); affective factors, such as math anxiety or boredom (Daker et al., 2021; Krannich et al., 2019; Tze et al., 2016); and attitudinal factors, such as stereotypes or self-concept (Gioannis, 2022; Krannich et al., 2019; Lazarides & Lauermann, 2019). An increasingly complex landscape of interventions has emerged, each aiming to positively influence one or more of these non-cognitive variables. While several meta-analyses have attempted to synthesize the effectiveness of such interventions (e.g., Higgins et al., 2019), not all adhere to scientific quality standards. Furthermore, they often focus on isolated non-cognitive variables, specific intervention types, or particular STEMM disciplines.

This study provides a systematic overview of meta-analyses examining interventions to improve non-cognitive outcomes in STEMM education, with a specific focus on K-12 and higher education students. We only included meta-analyses that adhere to established scientific standards for quality. Specifically, we give an overview of (a) interventions targeting the most frequently studied subcategories of non-cognitive outcomes—namely, motivational, affective, and attitudinal variables; (b) the various types of interventions employed; and (c) the specific STEMM subjects the interventions in the meta-analyses focus on. Using a second-order meta-analysis, we evaluate the overall effectiveness of these interventions and compare their effects across different outcome categories, intervention types, and STEMM disciplines.

Non-cognitive variables and their role in STEMM education

The term *non-cognitive variables* encompasses a wide range of constructs and multiple ways of categorizing them. In our study, we categorize the non-cognitive variables into motivational, affective (affective/emotional), and attitudinal. We chose this categorization for two main reasons. First, for our task of categorizing meta-analyses in STEMM education, broad categories are preferable to a more fine-grained categorization, as most meta-analyses don't report separate effects for the detailed variables that would underpin a finer categorization. Second, there is a theoretical distinction between the three categories we chose (Bohner & Dickel, 2011; Pekrun, 2006; Urhahne & Wijnia, 2023). The aim of the following sections is to define these three categories and present theories for the different constructs. We also describe the individual variables that comprise these broad categories and assess their relevance in the context of STEMM education in K-12 and higher education. We do not suggest that *non-cognitive variables* are independent of cognitive processes. For example, affective variables closely influence cognitive processes related to learning (Pekrun, 2006), and there is a close relation between motivational variables and cognitive use of learning strategies (Schunck et al., 2014). We use the term *non-cognitive* to distinguish motivational, affective, and attitudinal variables from those relating to cognitive competencies or academic performance.

Motivational variables

Motivation is a process of goal-directed activities (Schunk & Usher, 2012). Different facets of the motivational process influence which goal-directed behaviors are followed (Urhahne & Wijnia, 2023). Motivation in the educational context influences the extent to which positive learning behaviors (e.g., rehearsing at home) are carried out, thereby affecting learning and performance (Schunck et al., 2014). Multiple motivational theories, like the expectancy-value theory (e.g., Wigfield & Eccles, 2000), social cognitive theory (e.g., Bandura, 1986), attribution theory (Weiner, 1986), self-determination theory (e.g., Ryan & Deci, 2000), achievement goal theory (Kaplan & Maehr, 2007), or interest theory (e.g., Hidi & Renninger, 2006), refer to different facets of the motivational process and their importance in the educational context.

A wide range of motivational variables has been investigated in STEMM education research. Interest (e.g., Zheng et al., 2025), task values and costs (e.g., Wille et al., 2020), self-efficacy (e.g., Živković et al., 2023), as well as mastery and performance orientation (e.g., Mamaril et al., 2016) are among the many variables that have been investigated in STEMM education research. Interest refers to either the psychological state of engaging or the predisposition to reengage with objects, events, or ideas (Hidi & Renninger, 2006) and can be seen as a form of intrinsic motivation (Ryan & Deci, 2000). In the context of STEMM education, interest may relate to aspects of the STEMM classroom or lessons, or the concept of STEMM. Self-efficacy refers to the individual's belief in their ability to perform a behavior that leads to a desired outcome (Bandura, 1977). Self-efficacy is an important part of the motivational process that mediates between the self and an action (Urhahne & Wijnia, 2023) and has been shown to be positively related to performance (e.g., Živković et al., 2023) and educational choices (e.g., Cuder et al., 2024) in STEMM. Task values and costs play an important role in achievement-related choices, according to expectancy-value theory (Eccles & Wigfield, 2020; Wigfield & Eccles, 2000). A high value for a certain subject (or domain) relates (in combination with a high expectancy in that domain) positively to academic performance and educational choices in that domain, while a high perceived cost towards a certain domain is negatively related to performance and choices (e.g., Piesch et al., 2020; Wille et al., 2020). This has been shown for different value facets and costs (e.g., Wille et al., 2020). Mastery and performance orientation in a domain, as part of achievement goal theory (Kaplan & Maehr, 2007), influence learning and achievement behavior depending on how success and competence are defined by the individual. A higher mastery orientation was more strongly positively associated with career aspirations than a higher performance orientation (e.g., Mamaril et al., 2016). Motivation towards a STEMM subject or school in general, and interest and self-efficacy as part of motivational processes, play a role in STEMM-related career interest and choices as well as academic performance in K-12 and higher education (Amalina et al., 2025; Ferdinand et al., 2024; Jiang et al., 2020). Some findings hint at a possible causal role of motivational variables in STEMM education: Increased interest and utility value in biology, following a utility-value intervention, led to higher enrollment in a following biology and final course grades (Rosenzweig et al., 2019). Another utility-value intervention increased STEM persistence by bolstering students' motivation in STEM (Asher et al., 2023).

Emotional/affective variables

Emotions can be "regarded as systems of interrelated psychological processes" (Pekrun, 2000, p. 144), containing emotion-specific cognitions, motivational tendencies, physiological processes, expressive behavior, and an affective component, which is the activation of subcortical brain systems that are subjectively experienced as emotional feelings (Pekrun, 2000). Theories of emotions and affects differ in the emotions that they distinguish (Lazarus, 1991). Typically, affects are characterized depending on their positive or negative valence and also whether they are activating or not (Camacho-Morles et al., 2021; Russell, 1980). In the educational context, the control-value theory of achievement emotions (Pekrun, 2000) describes emotions in an achievement or academic context, their antecedents, and effects (Pekrun, 2006). Pekrun (2000) offers a classification for the achievement emotions along the lines of emotions or affects in general. He classifies emotions by valence (positive, negative, neutral), contextual frame of reference (individual, social), and time reference (prospective, retrospective, concurrent). Achievement emotions can be related to the activity of learning (activity achievement emotions) or the learning outcome (outcome achievement emotions). Boredom is an example of an achievement emotion related to the activity of learning; shame about failure is an example of an outcome-related emotion (Pekrun, 2000). The activity achievement emotions, enjoyment, anger, and boredom, were all found to significantly relate to academic performance in a recent meta-analysis (Camacho-Morles et al., 2021); only frustration did not show a significant correlation. While higher enjoyment was connected with better performance, anger and boredom were associated with lower performance (Camacho-Morles et al., 2021). Tze et al. (2016) found stronger negative correlations between boredom and motivation, as well as boredom and study strategies/behaviors, than between boredom and achievement in their meta-analysis. The significant correlations were found across K-12 and higher education (Camacho-Morles et al., 2021; Tze et al., 2016) and also for the investigated STEMM subjects, namely mathematics, science, and chemistry (Camacho-Morles et al., 2021). For enjoyment, research found that happier students could memorize the content better (Hernik & Jaworska, 2018), and showed higher achievement even in later school stages (Morris et al., 2021). Anger was strongly negatively related to achievement, with motivation as a potential moderator (Camacho-Morles et al., 2021).

Even though research on the relationship between emotional sub-facets and academic achievement (Camacho-Morles et al., 2021) is growing, most research in the STEMM domain focuses on mathematics anxiety (Barroso et al., 2021; Namkung et al., 2019). Meta-analytic evidence found this "negative affect related to math" (Barroso et al., 2021, p. 135) to be significantly, negatively related to mathematics achievement from grade 1 to college (Barroso et al., 2021; Namkung et al., 2019). Mathematics anxiety has been found, in multiple studies, to be negatively related to career choices in STEMM (Cuder et al., 2024; Eidlin-Levy et al., 2023; Ferdinand et al., 2024). Other emotions, apart from mathematics anxiety, were also related to academic choices, retention, and career aspirations: In an analysis of PISA data, enjoyment of learning was positively associated with future career intentions in STEMM (Wang et al., 2021). Among second-grade students at a STEMM-focused university, enjoyment and boredom were directly related to dropout intentions (Respondek et al., 2017).

Attitudinal variables

Attitudes are a broad concept with a long history of being researched in psychology (Allport, 1935). In short, an attitude can be defined as "an evaluation of an object of thought" (Bohner & Dickel, 2011, p. 392), with these objects ranging from simple items to complex ideas. Different models of attitudes vary in whether they conceptualize attitudes as stable or temporary on-the-spot judgements (Bohner & Dickel, 2011). In the educational context, constructs such as stereotypes, self-concept, and general attitudes toward a subject or course are of vital interest to researchers (e.g., Gioannis, 2022; Osborne et al., 2003; Wu et al., 2021). Stereotypes can be defined as "an individual's set of beliefs about the characteristics or attributes of a group" (Judd & Park, 1993, p. 110). Gender stereotypes and their effects take a central role in research on stereotypes in education (Kollmayer et al., 2018), and can be conceptualized as explicitly and implicitly held beliefs of gender roles (Gioannis, 2022). One prevalent stereotype, with potential consequences for participation in STEMM, is that women are worse than men at mathematics (Song et al., 2017).

An individual's self-concept can be seen as their perception of themselves (Shavelson et al., 1976) or, in other words, beliefs and attitudes about oneself. These perceptions are influenced by the individual's interactions with the environment (Shavelson et al., 1976). A high academic self-concept in an educational domain, which is of relevance in the educational context, would refer to the belief that one is capable in this domain. Self-concept and self-efficacy share similarities. Still, we categorized self-efficacy as a motivational variable and self-concept as an attitudinal variable, as self-efficacy is more closely related to concrete, goal-driven actions, whereas self-concept is more reflective and less closely related to concrete actions.

In the context of STEMM education, research often focused more broadly on attitudes towards science, i.e., "the feelings, beliefs and values held about an object that may be the enterprise of science, school science, the impact of science on society or scientists themselves" (Osborne et al., 2003, p. 1053). A typical item to assess attitudes towards science (or, analogously, other STEMM subjects) would be "Science is fun" (Osborne et al., 2003). Motivational or affective aspects have sometimes been included in assessments of attitudes towards science (Osborne et al., 2003), which can make it difficult to assess the role of the different non-cognitive categories in STEMM education.

Attitudes play a significant role in STEMM education in various ways. Gioannis (2022) found direct associations between implicit gender-science stereotypes and STEMM major intentions in high-school students; gender math stereotypes were negatively associated with girls' career aspirations at the high-school level (Song et al., 2017). In addition to choices or intentions, higher gender-math stereotypes were, through the mediating effect of mathematics self-concept, negatively related to women's mathematical performance in college (Xie et al., 2023). Wu et al. (2021) found a reciprocal relationship between academic self-concept and achievement, also for the domain of mathematics.

The interplay between motivational, affective, and attitudinal variables

Importantly, motivational, affective, and attitudinal variables are not only related to outcomes such as STEMM performance or educational choices, but they are also interrelated and have additive or interactive effects on these outcomes. For example, in their meta-analysis, Li et al. (2021) found a moderate, negative correlation between students' math anxiety and motivation towards mathematics. Math anxiety and intrinsic motivation towards mathematics were additionally found to interact in their influence on performance in mathematics: Students with high intrinsic math motivation showed inverted-U relations between math anxiety and mathematics performance (lower performance for low and high math anxiety) while for students with low intrinsic math motivation generally negative association between math anxiety and mathematics performance were found (Wang et al., 2015). Self-efficacy and math anxiety were significantly negatively related in mathematics (Akin & Kurbanoglu, 2011; Shimizu, 2025), and self-efficacy was positively related to behavioral engagement in mathematics class (Shimizu, 2025). Additionally, self-efficacy was positively related to positive attitudes in mathematics, which, in turn, were negatively related to math anxiety (Akin & Kurbanoglu, 2011). Ferdinand et al. (2024) investigated the joint effects of math anxiety, self-efficacy, and interest on STEM career choices. They found that math anxiety and math interest had independent effects on STEM career choices, whereas math self-efficacy did not. Math interest remained predictive even after controlling for mathematical achievement. Tsai et al. (2023) found positive relationships between science attitudes and motivational variables in mathematics. Engineering gender stereotypes were related to more negative emotional states in engineering, while not significantly related to engineering interest or engineering self-efficacy (Ramos-Sandoval & Ramos-Diaz, 2024). Some studies.

Interventions in STEMM education focusing on motivational, affective, and attitudinal variables

As discussed above, all three non-cognitive variables play a significant role in STEMM education, influencing outcomes such as performance, choices, and study or career intentions. Accordingly, numerous interventions aim to modify these non-cognitive variables. Many of these interventions are effective (e.g., Hong, 2010; O'Keefe et al., 2023; Shin et al., 2023; van den Berg et al., 2019). However, interventions vary in effectiveness and may be ineffective or even cause negative effects (e.g., Bai et al., 2012; Brathwaite, 2010; Kebritchi et al., 2010). In the following sections, we will describe a few examples for each of the three categories of non-cognitive variables.

Various types of interventions, ranging from long-term out-of-school programs (Mohr-Schroeder et al., 2014) and online mentoring programs (Stoeger et al., 2013) to shorter interventions such as growth mindset interventions (Burnette et al., 2020), showed positive effects on students' motivation, e.g., their interest in STEMM fields. In addition to numerous primary studies, there are a few meta-analyses of interventions focusing on motivational outcomes (e.g., Fadda et al., 2022; Young et al., 2017). Fadda et al. (2022) found a small effect of digital games on students' motivation in mathematics, and Young et al. (2017) found a small-to-medium effect of out-of-school time on students' interest in STEM.

The types or approaches of interventions focusing on motivational outcomes vary, as do the STEMM fields in which they are implemented. For example, there are lab-based role models (Cheryan et al., 2013) as well as a game-based intervention that aims to increase students' STEMM interest (Hunt et al., 2025). Although there are some indications that the same type of intervention has varying degrees of effectiveness (e.g., for the effects of games on motivation; Bai et al., 2012; Kebritchi et al., 2010), there has been little research to date on whether different types of interventions influence the same motivational outcomes differently.

Interventions aimed at altering emotional and affective outcomes in STEMM education primarily focused on reducing mathematics anxiety. Positive effects were found, among others, for intensive one-on-one cognitive tutoring (Supekar et al., 2015), a cognitive reappraisal intervention for community college students (Jamieson et al., 2016), and emotional and math strategy training (Passolunghi et al., 2020). Multiple interventions assessed the effects of games on mathematics anxiety, with unclear results (e.g., Rocha & Dondio, 2021; Vanbecelaere et al., 2020). Interventions less frequently targeted other affects than anxiety. Exceptions include physical activity interventions that enhance primary school children's and preschoolers' enjoyment of mathematics (Liang et al., 2023; van den Berg et al., 2019), and a virtual reality intervention that improves nursing students' satisfaction (Jung & Park, 2022). In addition to primary studies, there are also meta-analyses of intervention studies focusing on affect (e.g., Bicer et al., 2020; Dondio et al., 2023). Bicer et al. (2020) found a substantial reduction in students' math anxiety following Cognitive Behavioral Therapy interventions. Educational games, on the other hand, showed only a small, non-significant reduction in math anxiety (Dondio et al., 2023).

Attitudes as outcomes in STEMM education were also targeted and investigated in various interventions. Gender stereotypes, as one example of attitudes, were reduced after a multiple-session identity threat intervention (Zhao et al., 2018) and a growth-mindset intervention in a science museum (Law et al., 2021). Women's implicit attitudes toward STEM were improved by exposing them to female STEM experts (Stout et al., 2011). Another intervention at the institutional level of schooling found that girls in single-sex classes had better physics-related self-concepts than those in coeducational classes (Kessels & Hannover, 2008). No difference was found for boys. Yildirim (2018) found a positive impact of out-of-school learning environments (e.g., science museums) on attitudes towards the science course; a four-week science through sports curriculum also significantly increased attitudes towards science (Galoyan et al., 2022). For interventions targeting attitudinal variables, there were also multiple meta-analyses investigating different types of interventions (e.g., Higgins et al., 2019; Pahlke et al., 2014). Higgins et al. (2019) found a medium effect on students' attitudes toward the application of educational technology in mathematics education. Pahlke et al. (2014) reported small, nonsignificant effects of single-sex schooling on mathematics attitudes from controlled studies. Uncontrolled studies have shown small, positive effects on girls' attitudes towards mathematics and science. The science attitudes of boys were negatively impacted following single-sex schooling.

Current study

Trying to grasp the research area of interventions with motivational, affective, or attitudinal outcomes is challenging. Although several meta-analyses have assessed the effectiveness of interventions on motivational, affective, or attitudinal outcomes, it is difficult to obtain a comprehensive overview. On the one hand, meta-analyses clarify the research area, integrate effects from small-scale studies and contradictory results, and explain heterogeneity in effects (Deeks et al., 2019; Gurevitch et al., 2018). On the other hand, not all meta-analyses meet quality standards, making appraisal of the results difficult (Ioannidis, 2016; Tamim et al., 2021). Additionally, meta-analyses often focus on a single non-cognitive outcome or even sub-facets of these outcomes, making comparisons difficult. To add to that, meta-analyses that focus solely on a specific intervention or a single STEMM subject further complicate the understanding of the research area. Finally, meta-analyses report not only positive effects of varying magnitudes but also null and negative effects.

To address the cluttered state of this research area and provide a structured overview of the field, we conducted a second-order meta-analysis of interventions on motivational, affective, and attitudinal outcomes in STEMM education. This approach enables us to structure the field of research, identify gaps and hotspots, and integrate and compare the effectiveness of different intervention types in different STEMM areas. Furthermore, by using meta-analyses as the basis of our analysis and assessing the quality of the included meta-analyses, we aim to obtain robust estimates of the effectiveness of the different interventions.

Research questions

In our study, we conducted a second-order meta-analysis of interventions with non-cognitive outcomes in STEMM education to first investigate the number and thematic scope of such interventions as well as differences in intervention type, STEMM domain, and targeted outcome. Second, we addressed the following research questions identified in the literature review.

> RQ1: Are the interventions effective overall?
> RQ2: Does the effectiveness differ across outcome types?
> RQ3: Does the effectiveness differ across intervention types?
> RQ4: Does the effectiveness differ across STEMM subjects?

Exploratory Analyses: For those interventions where there are enough effect sizes, does the effectiveness differ for the interactions between outcome and intervention types, intervention types and STEMM subjects, as well as outcomes and STEMM subjects?

Method

Literature search and inclusion criteria

The literature search was conducted as part of a broader review of STEMM education literature. A team of five researchers (including this paper's first and second authors) conducted a general search on second-order publications in STEMM education. Although the final inclusion criteria for this review limited the data to meta-analyses of interventions with non-cognitive outcomes for students, the initial search was broader.

The search was conducted across four databases: ERIC, PsycINFO, the Web of Science Core Collection, and Psyndex.[1] ERIC was included as the primary database for educational research because it also contains grey literature, such as research reports, conference papers, and dissertations. Web of Science contains a broad array of multidisciplinary research across the STEMM fields. PsycINFO and Psyndex were included because they comprehensively index psychological research from diverse sources and should aid in obtaining research on non-cognitive constructs. We believe this selection of four databases gives a comprehensive selection of the topic. Using a search term tailored to the respective databases, the titles and abstracts of articles were searched to identify relevant second-order publications in STEMM education. The complete search string for each database contained search words for STEMM (e.g., "mathematics", "science", "STEM", "STEMM"), words for education (e.g., "education", "teaching", "program"), and specific words for second-order publications (e.g., "meta-analysis", "systematic review", "scoping review", "synthesis"). The search string used across all databases is provided in Supplemental Material S1. To obtain a broad overview, including unpublished dissertations, we did not limit the search to any document type or article language. We filtered for articles published between January 1, 2000, and October 31, 2023, spanning more than 20 years of second-order research in STEMM education. 6664 articles were identified, with 5103 articles remaining after the automatic removal of duplicates. The PRISMA flowchart in Fig. 1 provides an overview of the search and inclusion process.

As a first step, the five researchers conducted initial inclusion coding of the dataset. Changing pairs of two people coded each of the 5103 articles by screening the titles and abstracts of every article to determine whether it was a second-order publication on STEMM education. Disagreements between the two coders were resolved via discussion. All coder pairs showed high reliability for in-/exclusion at this process step (Cohen's kappas between 0.81 and 0.95).

In a second step, the five researchers further categorized the remaining 1578 articles, all of which were second-order publications on STEMM education. Again, coder pairs were constructed, and two people categorized each article via this finer coding procedure. In this coding procedure, the title and abstract of each article

---

[1] ERIC was accessed via ProQuest, PsycINFO via the EBSCO host, and Psyndex via PubPsych. The Web of Science Core Collection was searched directly. For Web of Science, the Filters "Education Educational Research", "Education Scientific Disciplines", "Psychology Multidisciplinary", "Psychology Educational", "Psychology", and "Education Special" were used.

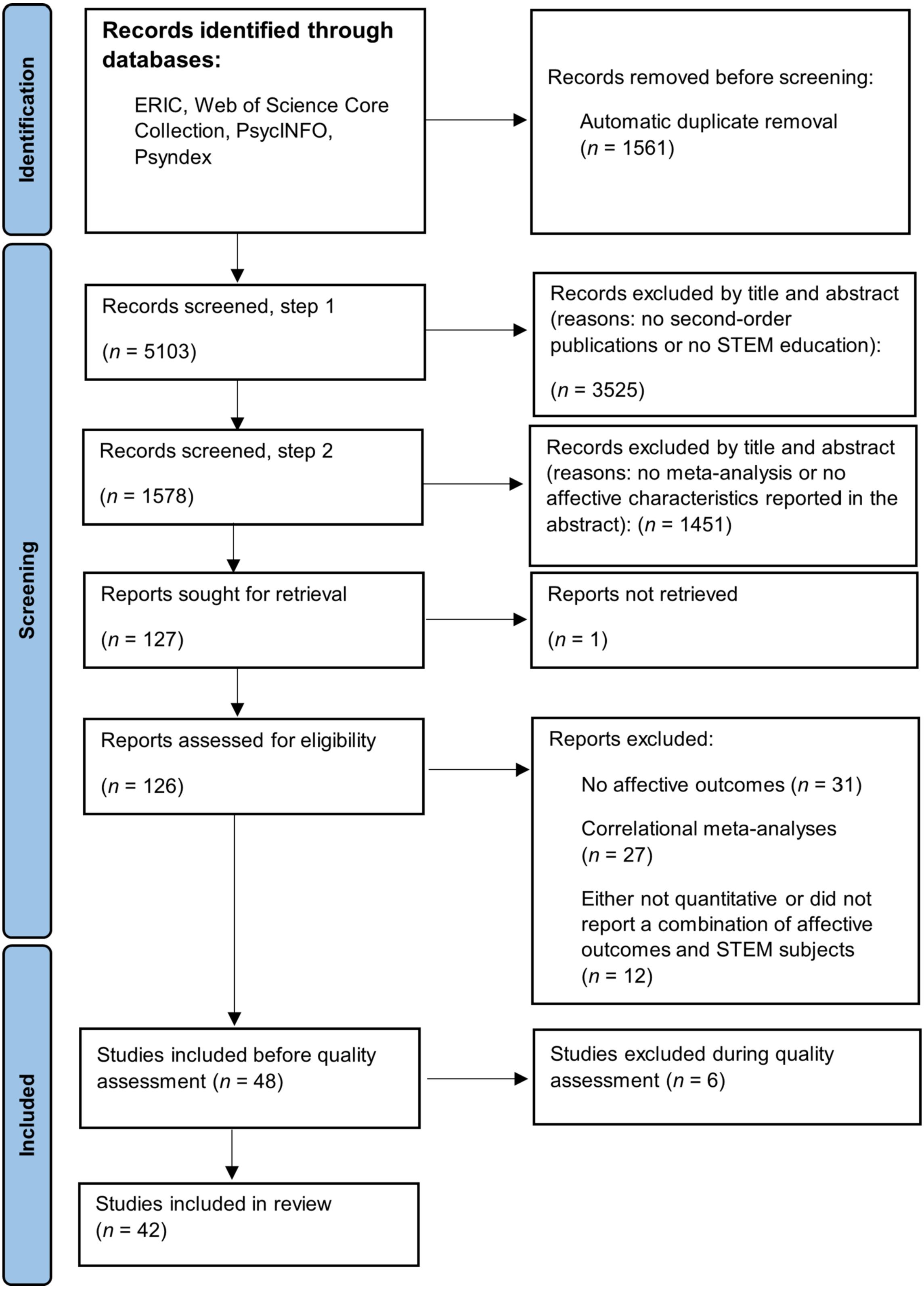

Fig. 1 PRISMA flow chart

were consulted to tag categories: whether the article was a meta-analysis, the exact STEMM disciplines it referred to, whether it was researching students, and whether it mentioned motivational, affective, or attitudinal variables. Cohen's Kappa for those three non-cognitive categories showed good agreement (Kappa's were 0.67 for the affective, 0.72 for the attitudinal, and 0.73 for the motivational category). 501 of 1578 articles were tagged as meta-analyses (Cohen's Kappa for the article type was 0.89). Of these 501 meta-analyses, 127 articles related to motivational, affective, or attitudinal aspects.

These 127 articles were then subject to stricter inclusion/exclusion criteria to create the final dataset for this review. To be included in our review, we checked the full text of all 127 articles to assess the following aspects: (a) The meta-analysis needed to aggregate effect sizes from interventional primary studies, (b) the meta-analysis needed at least one non-cognitive characteristic as an outcome of an intervention, (c) the meta-analysis needed to present an aggregated, quantitative outcome for at least one non-cognitive characteristic. (d) The outcome and intervention needed to relate to at least one of the subjects of STEMM education. (e) The outcome(s) needed to be student-related, pre-K-12, or university-level. Examples of non-cognitive outcomes are listed in Table 1.

Correlational meta-analyses that did not focus on interventions were excluded. Some meta-analyses were only partly about STEMM. If these meta-analyses did not report effect sizes separately for STEMM and non-STEMM outcomes, they were excluded. Similarly, some meta-analyses focused only partly on motivational, affective, or attitudinal outcomes; if they did not report effect sizes separately for these non-cognitive and other cognitive outcomes, they were excluded. Meta-analyses with non-student samples (e.g., teachers) were also excluded. The PRISMA flow chart in Fig. 1 specifies the number of manuscripts excluded based on these criteria. After applying these criteria to the 127 articles, 48 interventional meta-analyses remained in our final data set. The PRISMA flow chart in Fig. 1 details the reasons for the exclusions.

Table 1 Examples for inclusions in our full-text coding

Non-cognitive categories	Examples for inclusion
Motivational	e.g., motivation towards a STEMM course, interest in a STEMM subject, self-efficacy in a STEMM activity
Affective	e.g., anxiety towards a STEMM subject, satisfaction towards a STEMM course
Attitudinal	e.g., attitudes towards a STEMM subject, self-concept in a STEMM subject, attitudes towards a STEMM course

Meta-analysis quality

We aimed to include only meta-analyses that adhere to established scientific standards for quality, thereby allowing robust conclusions about the interventions analyzed. To assess the quality of the included meta-analyses, we drew on existing standards for systematic reviews and meta-analyses (Moher et al., 2009; Page et al., 2021) as well as a recent systematic review of meta-analyses (Knogler et al., 2022) to construct a list of quality indicators. We thus identified nine quality indicators related to the methods and results sections of the meta-analyses. These criteria were relatively broad, as we did not aim to verify whether new, best-practice methods were followed, but only whether a basic, good-practice procedure was followed and reported. The complete coding manual for quality is presented in the Supplemental Material S2. We considered whether the databases and search terms were described and whether the coding process was specified. We checked the inclusion/exclusion criteria and whether the study designs included in the meta-analyses were specified. Another quality criterion was that the meta-analyses should give an overview of all included studies and their characteristics. The following aspects were related to the effects: Did the authors define a statistical model for the overall effect? Did they check for heterogeneity? Did they assess the risk of bias? Did they assess publication bias? Finally, we considered whether the authors reported a confidence interval for the overall effect.

All nine quality criteria were double-coded by the paper's first and second authors in accordance with the quality coding manual for 10 of the 48 included meta-analyses (~21%). Overall, an agreement of 94.4% and a Cohen's Kappa value of 0.86 showed substantial agreement. The paper's first author then assessed the quality of the remaining articles.

To achieve a robust set of meta-analytic effects, we used a combination of our quality measure and the number of individual effects to exclude meta-analyses. Meta-analyses that adhered to fewer than five criteria were excluded from further analyses due to limited quality. Meta-analyses adhering to five or six quality criteria were also excluded if they contained fewer than 10 individual effect sizes that were aggregated. Meta-analyses adhering to more than six criteria were included. Based on these criteria, six meta-analyses were excluded due to low quality, leaving 42 for the final analyses. To strengthen the robustness of our results, we conducted all analyses using the complete data set, excluding the 'limited quality' studies.

Data extraction and coding

For the remaining dataset of 42 interventional meta-analyses, general information to describe the dataset and data to answer the research questions were extracted from each article and coded into a spreadsheet. As general information, we extracted the country of the first author's institution, the STEMM subjects in the meta-analysis, and the researched sample. The number of included studies was also extracted, as was whether the meta-analysis focused solely on non-cognitive outcomes or also on achievement-related outcomes and cognitive competencies.

Content-wise, details on the interventions and non-cognitive outcomes were extracted at two levels: the first level was the broad term the authors used to describe the intervention or non-cognitive outcome. The second, more fine-grained level was implemented when the authors provided more detailed information about the interventions and the outcomes investigated. As an example of the interventions, the meta-analysis by Higgins et al. (2019) investigated the effects of technology in mathematics instruction (first level), and in their manuscript, they specify that this included games, e-mentoring, computer-assisted instruction, etc. (second level). For the outcomes, the meta-analysis by Arztmann et al. (2023) is a fitting example. They reported effects on motivational outcomes (first level) and later specified that these contained intrinsic motivation, interest, enjoyment, and self-efficacy (second level).

To answer our research questions, we needed a first-level categorization of non-cognitive outcomes, interventions, and STEMM fields. For clarification of the terminology, primary effects are effects from the individual primary studies aggregated in a meta-analysis. These aggregated effects, which are contained in the meta-analyses of our dataset, are called first-order meta-analytic effects. Non-cognitive outcomes, intervention types, and STEMM fields were categorized by each first-order meta-analytic effect. Meta-analyses that included multiple relevant first-order meta-analytic effects had each effect separately categorized by non-cognitive outcomes, interventions, and STEMM fields. In what follows, we refer to these first-order meta-analytic effects when discussing effect sizes in the meta-analyses we cover. An exception are the primary effects of the individual interventions included in the meta-analyses, which will be made explicit in the text when they are mentioned.

For the non-cognitive outcomes, we separated motivational, affective, and attitudinal outcomes. The remaining meta-analytic effects that did not fit any of these three non-cognitive outcomes (e.g., by being too unspecific about which non-cognitive outcomes were contained in the effect or by the effect being a mixture of our three categories of non-cognitive outcomes) were labeled miscellaneous. Motivational outcomes were categorized as motivation, whether presented as broad motivational outcomes (e.g., a mixture of motivational variables such as intrinsic or extrinsic motivation) or as sub-components of motivation (e.g., self-efficacy or interest). Affective outcomes were those outcomes that focused on types of affect, e.g., anxiety, satisfaction, or anger, in relation to a STEMM subject or a STEMM lesson. Attitudinal outcomes were either general attitudes towards a STEMM subject or STEMM lesson, or the individual's self-concept or stereotypes towards a STEMM subject.

First-level categories of interventions were applications of student-centered learning, applications of educational technology, and game-based learning. We categorized the remaining interventions into a miscellaneous category because they occurred in only a single meta-analysis and did not fit with other interventions. Examples of interventions in the miscellaneous category were out-of-school programs or role-model interventions. Finally, categorizing the different STEMM subjects was straightforward. The different subjects that fall under the STEMM acronym (e.g., science, mathematics, chemistry, computer science), as well as the acronyms themselves (e.g., STEM, STEMM), were handled as separate categories. Cases in which multiple STEMM subjects were combined into one outcome (e.g., a meta-analysis focusing on science and technology education, but not STEM or STEMM education as a whole) were labeled mixed.

Meta-meta-analytical method

In handling and transforming the meta-analytic effect sizes and then calculating the effects for our second-order meta-analysis, we followed the procedures displayed in a recent second-order meta-analysis by Jansen et al. (2024). The first-order meta-analytic effects described above were the basis for the present analysis. We aim to leverage these first-order meta-analytic effects to obtain second-order effects.

Extracting and transforming effect sizes

We extracted all effect sizes that fit the scope of our study from the remaining 42 meta-analyses. Some meta-analyses focused on a single intervention for a single non-cognitive outcome, and thus, only one meta-analytic effect was extracted. If a meta-analysis reported effect sizes for multiple interventions and/or multiple non-cognitive outcomes, we recorded multiple effect sizes from these meta-analyses. We recorded multiple effect sizes from a single meta-analysis only if they were disjunct at the level of included primary effect sizes. For example, Lawner et al. (2019) aggregated the effect sizes of role-model interventions on interest separately for lab and field studies. We can record both effect sizes, as none of the primary effect sizes used to calculate the effect size for lab studies were used in the calculation for the field studies. In contrast, if a meta-analysis reported, for example, the overall effects of games on motivation and then the effect of digital games as a subset, we included only the overall effect of games to avoid inflating the influence of certain primary effects. Following this procedure, we extracted 64 effect sizes from 42 meta-analyses.

For some effect sizes, random- and fixed-effect estimates were reported in the meta-analysis. In those cases, only the random-effects estimate was recorded. Effect sizes that were reported as Hedges' g were transformed into Cohen's d by using the correction factor given in Borenstein (2009). The risk ratio in Baashar et al. (2022) was transformed using an online calculator (Lenhard & Lenhard, 2017), resulting in all of our effect sizes being presented as standardized mean differences. One final transformation concerned the effect sizes of *negative* non-cognitive characteristics, such as anxiety. Interventions targeting these characteristics aim to reduce the negative characteristics. To make the effect sizes of these interventions comparable to those that aim to increase a characteristic (e.g., interventions focusing on motivation), we reversed their signs. This results in our final effect sizes being positive when the effect points in the intended direction.

Meta-analytic procedure

To calculate second-order effect sizes, we also followed the approach of Jansen et al. (2024), who adhered to the recommendations of Schmidt and Oh (2013). We used R (Version 4.2.3, R Core Team, 2020) and the *metafor* package (Version 4.6-0, Viechtbauer, 2010) within the RStudio (RStudio Team, 2020) environment for our analyses. The data sheet and analysis script we used are available on OSF (https://doi.org/10.17605/OSF.IO/UKESG). We used random-effects models to calculate the overall effect. The effects were modeled in a multilevel structure, as we included between 1 and 6 effect sizes per meta-analysis. Including multiple effect sizes in the calculation of second-order effects is advised (Tipton et al., 2019), as averaging effect sizes from a single meta-analysis or selecting a single effect would underestimate between-study heterogeneity (Schmidt & Hunter, 2015). Although, as mentioned above, we didn't include multiple effect sizes from a single meta-analysis if they overlapped, dependencies between effect sizes from a single meta-analysis still exist. The effect sizes from one meta-analysis were generated using the same meta-analytic method and can possibly contain effect sizes from the same primary studies (even if not the same primary effect sizes). This dependency was corrected using robust variance estimation (the robust function in the metafor package; Viechtbauer, 2010).

The weights used to calculate the second-order effects were determined as follows: for each effect, the inverse of the first-order sampling error, divided by the number of studies, was used as its weight (Schmidt & Oh, 2013). The first-order sampling error was calculated using formula 4 from Marín-Martínez and Sánchez-Meca (2010), based on the first-order effect and sample size. To address the problem of overlapping primary studies across meta-analyses and to reduce the weight of effect sizes from meta-analyses with high overlap, the number of studies used to calculate the weights was adjusted based on the number of unique primary studies and those that appeared in multiple meta-analyses. Primary studies that occurred once in our data were counted as 1, studies that occurred twice were counted as 0.5 (uniqueness value), and so on (Jansen et al., 2024; Munder et al., 2013). Meta-analyses for which the primary studies were not listed got the adjusted number of primary studies by multiplying the number of primary studies by the mean uniqueness value. Overlapping primary studies were not a major issue: Of 1344 listed primary studies,[2] 1264 appeared in only one meta-analysis, 33 appeared twice, and 2 appeared three or four times. Sometimes, the number of participants needed for calculating the first-order sampling error was missing. We then estimated the sample size by the total number of participants multiplied by the fraction of primary studies included in the first-order effect size and the number of primary studies in the meta-analysis. If the overall number of participants was missing, we used the median per effect size to estimate the total.

Moderators

To answer the research questions RQ2-RQ4, the overall effect from the second-order meta-analysis was not the critical result; rather, the moderator analyses were. We conducted three separate analyses for these three research questions: the effectiveness of different intervention types, the effectiveness of interventions targeting different outcome types, and the effectiveness of interventions across different STEMM subjects. For each moderator analysis, the different levels conferred to the first-level categorization that we made for each of the three research questions. As additional exploratory analyses, two-way interaction effects for the different outcome types, intervention types, and STEMM subjects were estimated.

Results

Descriptive information

After removing six meta-analyses from the final analysis due to limited quality, we included 42 meta-analyses and 64 effect sizes. The references for all included

---

[2] Not all meta-analyses listed citations for the primary studies they included in the analysis, but all meta-analyses report how many primary studies they analyzed. This explains the lower number of listed citations for the primary studies compared to the final overall number of primary studies reported in the Results section.

meta-analyses are listed in the Supplemental Material S3. Overall, these meta-analyses included 1903 primary studies, underscoring the breadth of the data we used for this analysis. Only 10 meta-analyses contained exclusively non-cognitive outcomes, and 32 contained both non-cognitive and achievement-related outcomes. Except for the risk of bias (26% of studies reported it), the other quality criteria we assessed were mostly adhered to (above 73% for the remaining criteria), indicating a reliable data basis for further analyses. Our data included students from pre-school through post-secondary education. 76% of meta-analyses included students in the K-12 range in their samples. Around 40% of these meta-analyses also included students in higher education. 24% of meta-analyses were only focused on higher education students.

We assessed the countries of the first authors' institutions. Overall, our data included meta-analyses from 16 different countries. First authors from institutions in the US ($n = 12$) and Turkey ($n = 8$) dominated, while those from East Asian countries contributed 12 studies overall, led by Taiwan ($n = 4$) and China ($n = 3$). Eight publications were authored by first authors from institutions in European countries (excluding Turkey), and two from the Middle East.

The publication years for the individual meta-analyses revealed an increase in the number of publications in recent years. Four meta-analyses from 2000 to 2010 were present in our data set, and an additional 20 meta-analyses were published from 2011 to 2020. The years from 2021 to 2023 added 18 more meta-analyses, signifying the steep increase in publications on this topic.

On the first, broader level, we found 27 effect sizes of interventions with attitudinal outcomes; 17 with motivational outcomes, and 15 with affective outcomes. The remaining five effect sizes consisted of miscellaneous non-cognitive outcomes or a mixture of the other outcomes, for example, in the meta-analysis by Jeong et al. (2019) in which the effect size was an overall affective measure that combined attitudes, motivation, interest, and satisfaction. Table 2 provides a detailed count of effect sizes for the different outcome (sub)types.

Table 2 Detailed listing of the number of first-order meta-analytic effect-sizes for the different outcome (sub-)types

Outcome types			
Affect ($k = 15$)	**Attitudes ($k = 27$)**	**Motivation ($k = 17$)**	**Miscellaneous ($k = 5$)**
Anxiety (7)	General/mixed attitudes (25)	General Motivation (10)	Mixed non-cognitive variables (5)
Satisfaction (6)	Self-concept (2)	(Self-)confidence (4)	
Stress (1)		interest (3)	
Depression (1)			

Except for two dissertations, meta-analyses were published as journal articles. *Nurse Education Today* ($n = 3$) and *BMC Medical Education* ($n = 2$) were the only journals with more than one meta-analysis included in our data set. The remaining publications came from various high-impact (e.g., *Psychological Bulletin*: IF = 17.3, *Educational Research Review*: IF = 9.6) and lower-impact journals (e.g., *School Science and Mathematics*: IF = 0.8, *Journal of Baltic Science Education*: IF = 1.5), including educational journals with and without a STEMM focus, as well as journals on general psychology and educational psychology.

Types of non-cognitive outcomes in the meta-analyses

The meta-analyses investigating the effects of interventions on attitudinal variables focused on general attitudes towards STEMM subjects ($n = 20$) and STEMM courses or lessons ($n = 5$). Two effect sizes were defined as students' self-concept regarding a specific STEMM subject as the outcome.

The motivational outcomes in the meta-analyses of different interventions were reported as general motivational outcomes. In these meta-analyses, they combined different types of motivational outcomes, like intrinsic or extrinsic motivation, to form a general motivational outcome. Three effect sizes focused on interest, and four on (self-)confidence towards a STEMM subject or a certain STEMM activity.

The affective variables in the meta-analyses mainly focused on the effect sizes of interventions on satisfaction and anxiety; two effect sizes examined the effects of interventions on stress and depression. The seven effect sizes regarding anxiety were of interventions to reduce mathematics anxiety in six cases, an important phenomenon in the STEMM education literature; the six effect sizes on satisfaction referred to satisfaction with a STEMM lesson or teaching method.

Types of interventions in the meta-analyses

There were 22 effect sizes for interventions that included implementations of educational technology. Fifteen effect sizes of interventions included different forms of student-centered teaching; five included game-based learning. Table 3 gives a detailed overview of the number of effect sizes for the different intervention (sub)types.

The effect sizes assessing the impact of educational technology interventions mostly stemmed from three types of interventions: using calculators in school ($n = 7$), implementing virtual or augmented reality in education ($n = 7$), and broad technological uses ($n = 6$). Educational robotics in school ($n = 1$) and synchronous distance education ($n = 1$) were also assessed. In the game-based learning category, all five meta-analyses included a mixture of analog and digital games (or simulations) in their

Table 3 Detailed listing of the number of first-order meta-analytic effect-sizes for the different intervention (sub-)types

Intervention types			
Educational technology (*k* = 22)	**Game-based learning (*k* = 5)**	**Student-centered teaching (*k* = 15)**	**Miscellaneous (*k* = 22)**
VR/AR (7)	games (5)	Mixed student-centered teaching methods (10)	Single-sex schooling (6)
Calculator use (7)		Problem-based learning (2)	Skill-/mindfulness-/therapeutic intervention (6)
Broad technological uses (6)		Cooperative learning (1)	Broad ed. Intervention (5)
Educational robotics (1)		Creative-drama (1)	Ingroup role models (2)
Synchronous distance education (1)		Inquiry-based instruction (1)	Interventions to reduce gender differences (2)
			Out-of-school time Programs (1)

Table 4 Co-occurrences of outcome types and intervention types for the meta-analytic effects

	Educational technology (*k* = 22)	**Games (*k* = 5)**	**Student-centered teaching (*k* = 15)**	**Miscellaneous (*k* = 22)**
Affect (*k* = 15)	7	1	1	6
Attitudes (*k* = 27)	9	0	9	9
Motivation (*k* = 17)	4	4	4	5
Miscellaneous (*k* = 5)	2	0	1	2

Table 5 Co-occurrences of the outcome types and STEMM subjects for the meta-analytic effects

	Mathematics (*k* = 26)	**Medical science (*k* = 13)**	**Science (*k* = 10)**	**STEM (*k* = 11)**	**Mixed (*k* = 4)**
Affect (*k* = 15)	6	9	0	0	0
Attitudes (*k* = 27)	16	0	7	2	2
Motivation (*k* = 17)	3	4	3	6	1
Miscellaneous (*k* = 5)	1	0	0	3	1

interventions. 10 of the 15 effect sizes for student-centered teaching interventions were derived from evaluations of multiple student-centered teaching methods. These meta-analyses examined student-centered teaching interventions (or applications of constructivism) as a whole and did not focus on a specific type of these interventions. Additionally, there were six effect sizes of interventions that focused on concrete student-centered teaching methods, namely problem-based learning ($n = 2$), cooperative learning ($n = 1$), creative-drama method ($n = 1$), and inquiry-based instruction ($n = 1$).

The miscellaneous category contained 22 effect sizes of interventions that could not be categorized further, either because there were too few effect sizes to form a category or because multiple effect sizes were extracted from only one or two meta-analyses. Single-sex schooling ($n = 6$), skill-/mindfulness-/therapeutic interventions ($n = 6$), broad educational interventions that were not further specified ($n = 5$), ingroup role models ($n = 2$), interventions to reduce gender differences ($n = 2$), and out-of-school time programs ($n = 1$) constituted this category.

Types of STEMM subjects in the meta-analyses

The various types of interventions differed not only in their instructional methods but also in the STEMM discipline they focused on. A majority of effect sizes focused on mathematics, with 26 interventions implemented in that domain. We found 13 effect sizes for interventions that focused on medical science, 11 on STEM, and 10 on science. Additionally, a mixed category was identified ($n = 4$), comprising studies that reported effect sizes of science and technology ($n = 3$) or science, technology, and mathematics ($n = 1$).

We listed the co-occurrences of the investigated outcome variables with the different intervention types in Table 4, and with the STEMM subjects the interventions focused on in Table 5. Affective outcomes were primarily researched in interventions that incorporated educational technology and therapeutic/mindfulness interventions. The primary focus lay on the impact of educational technology on students’ satisfaction and the use of therapeutic interventions for mathematics anxiety. Attitudinal and motivational outcomes were evenly distributed across the different intervention types. The only exception was that attitudes were not investigated in the context of game-based interventions, which were examined only for their effects on motivational aspects. An exception was one meta-analysis on reducing mathematics anxiety with game-based learning (Dondio et al., 2023).

Affective outcomes were only researched in the contexts of mathematics (anxiety) and medical science. Attitudinal outcomes have been primarily studied in the contexts of mathematics and science, with few meta-analyses in the medical sciences. Motivational outcomes were researched across all STEMM disciplines. Mathematics was the only discipline studied across all outcome variable categories.

RQ1: Are the interventions effective overall?

The overall second-order effect size was calculated with robust variance estimation in the random-effects model and resulted in a significant positive effect, $d = 0.42$ (95% CI [0.35, 0.50], $p < .001$), indicating an overall effect size of the educational interventions that resulted in small-to-medium positive effects, based on Cohen’s benchmarks

(Cohen, 1988), for the non-cognitive outcomes. The Q statistic indicated significant heterogeneity in the overall effect, Q(63) = 1,104,928 ($p < .001$). The multilevel random-effects model revealed meaningful variance between meta-analyses ($\tau^2 = 0.02$) and within meta-analyses ($\tau^2 = 0.05$). Consistent with this heterogeneity, the 95% prediction interval was wide, ranging from – 0.14 to 0.98, indicating considerable variability in true effects across contexts. All 64 effect sizes are listed in the Supplemental Table S4 with the students' grade level, country of the authors, STEMM subjects, intervention, and exact outcome type. The forest plot in Fig. 2 gives an overview of the 64 effect sizes.

Sensitivity analyses

We conducted additional sensitivity analyses to assess whether certain analytical decisions we made had a significant impact on the results. First, we tested whether the quality of meta-analyses significantly influenced the effects. Conducting the analyses with the 'limited quality' studies included did not significantly impact the results ($d = 0.43$, 95% CI [0.36, 0.50], $p < .001$); the same holds for the following moderator analyses. This revealed no significant impact of the moderator quality ($b = 0.29$, 95% CI [– 0.18, 0.76], $p = .22$), which did not change when we included the 'limited quality' studies ($b = 0.16$, 95% CI [– 0.20, 0.51], $p = .37$). Similarly, changing the quality-cutoff to 5 of 9 and 7 of 9 quality criteria without considering the number of primary effects also did not change the overall effect significantly ($d = 0.43$ and $d = 0.44$, respectively).

Second, we varied the parameter used to estimate the overlap of primary studies across the meta-analyses. Assuming a very high overlap (uniqueness of 0.25 for all primary studies) and no overlap (uniqueness of 1 for all primary studies) for the meta-analyses that did not list the primary studies they included did not change the overall results ($d = 0.42$, respectively). Conducting another analysis without the meta-analyses that did not list the primary studies they included ($k = 5$) did not impact the overall effect significantly ($d = 0.43$).

Finally, we conducted all analyses without those meta-analytic effects that we classified as 'miscellaneous non-cognitive outcomes' ($k = 5$), since they could not be fully classified as one of the other three outcome categories, and we cannot be sure which non-cognitive outcomes they contain. This change did not significantly affect the overall outcome ($d = 0.41$) or the subsequent moderator analyses.

RQ2: Does the effectiveness differ across outcome types?

For all outcome types, our moderator analyses revealed a small-to-medium positive effect of interventions, with detailed results presented in Supplemental Table S5. The test of moderators revealed no differences in the effectiveness of interventions for the different outcome types ($F(3,38) = 0.709$, $p = .55$). In other words, the different meta-analyzed interventions yielded similar outcomes, regardless of whether they targeted motivational ($d = 0.41$, 95% CI [0.31, 0.51]), affective ($d = 0.46$, 95% CI [0.27, 0.65]), attitudinal ($d = 0.39$, 95% CI [0.25, 0.52]), or the miscellaneous non-cognitive outcomes ($d = 0.53$, 95% CI [0.35, 0.71]).

RQ3: Does the effectiveness differ across intervention types?

Similar to the outcome types, the second moderator analysis also revealed significant, positive effects across all the intervention types examined in our data. The detailed results for the different interventional categories are presented in the Supplemental Table S6. The test of moderators did not indicate significant differences across the interventions ($F(3,38) = 0.674$, $p = .57$). Applications of educational technology ($d = 0.42$, 95% CI [0.24, 0.59]), games ($d = 0.39$, 95% CI [0.30, 0.49]), student-centered learning ($d = 0.48$, 95% CI [0.38, 0.58]), and the miscellaneous interventions ($d = 0.39$, 95% CI [0.25, 0.54]) all were effective in changing non-cognitive outcomes for the better with small-to-medium sized effects.

RQ4: Does the effectiveness differ across STEMM subjects?

The meta-analyzed interventions resulted in significant positive effects for all different STEMM subjects. The detailed results for the different STEMM subjects are presented in the Supplemental Table S7. The moderator analyses did not indicate overall differences between the different STEMM subjects ($F(4,37) = 2.112$, $p = .10$). Two pairwise comparisons using Tukey contrasts showed differences: Interventions in mathematics and STEM both showed significantly smaller effects ($z = -2.57$, $p < .05$ and $z = -2.69$, $p < .01$) than the interventions in mixed STEMM disciplines. No other pairwise comparison was significant; the general pattern showed small-to-medium effects for interventions in mathematics ($d = 0.32$, 95% CI [0.20, 0.45]), medical science ($d = 0.51$, 95% CI [0.29, 0.74]), science ($d = 0.45$, 95% CI [0.28, 0.62]), STEM ($d = 0.38$, 95% CI [0.29, 0.48]), and medium-to-large effects for mixed STEMM subjects ($d = 0.75$, 95% CI [0.49, 1.01]).

Exploratory analysis: interactions

As an exploratory analysis, the effect sizes of the two-way interactions between outcome type, intervention type, and STEMM subjects were estimated. Only the effect sizes for which the standard error could be estimated are included in the Supplemental Tables S8-S10. All effect size estimates were positive, indicating small-to-medium effects across all intervention and outcome type combinations. Implementations of educational technology

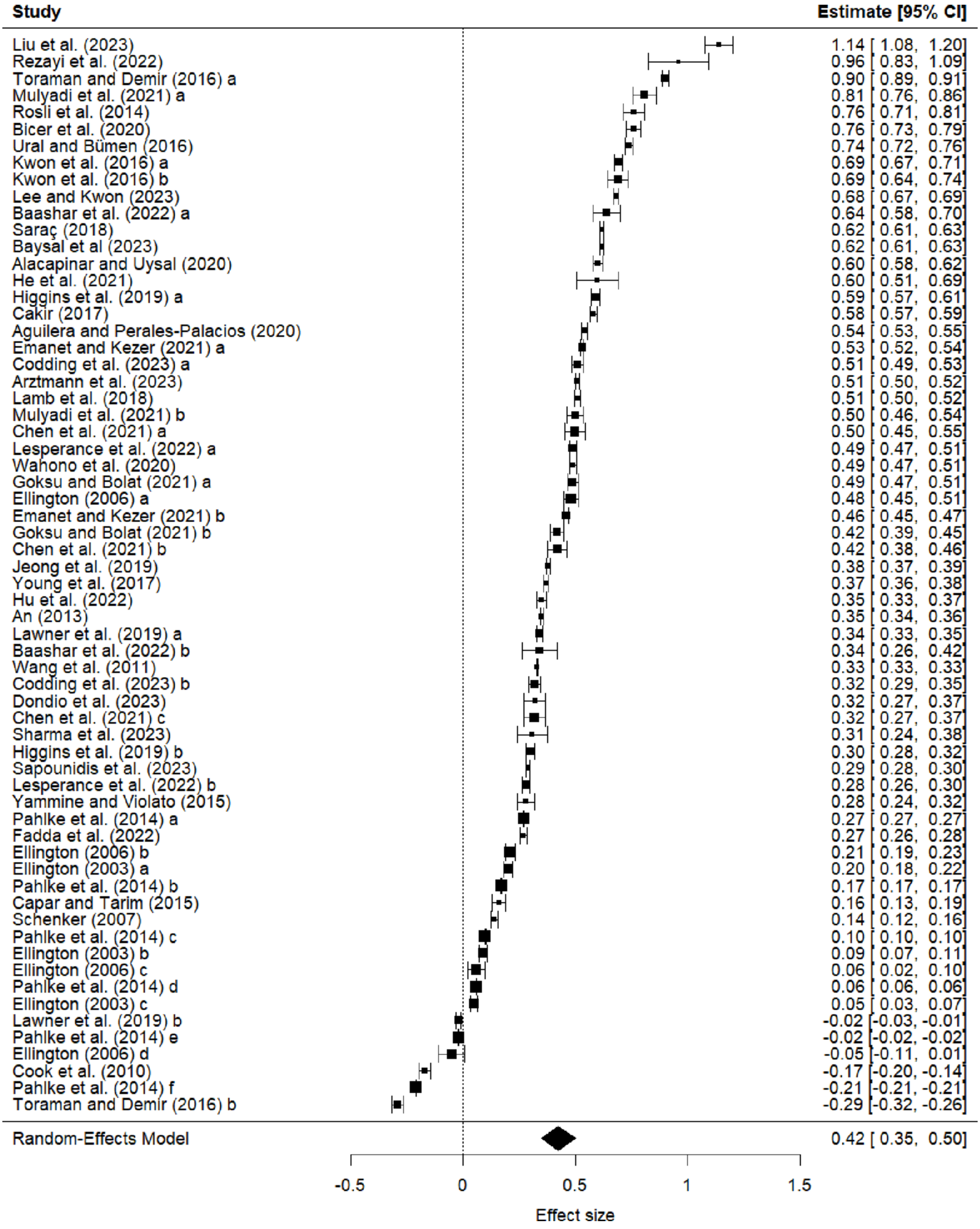

Fig. 2 Forest Plot of the 64 included effect sizes

showed positive effects on motivational ($d = 0.46$, 95% CI [0.13, 0.79]), affective ($d = 0.46$, 95% CI [0.07, 0.85]), attitudinal ($d = 0.30$, 95% CI [0.13, 0.47]), and miscellaneous outcomes ($d = 0.53$, 95% CI [0.28, 0.78]). Games showed positive effects on motivational outcomes ($d = 0.41$, 95% CI [0.29, 0.53]). Student-centered teaching methods showed positive effects on motivational ($d = 0.46$, 95% CI [0.32, 0.60]) and attitudinal outcomes ($d = 0.46$, 95% CI [0.30, 0.62]). Lastly, the miscellaneous interventions overall showed a positive effect on motivational ($d = 0.32$, 95% CI [0.16, 0.48]), affective ($d = 0.49$, 95% CI [0.32, 0.66]), and attitudinal outcomes ($d = 0.36$, 95% CI [−0.09, 0.81]). There were no significant differences across outcome types in educational technology implementation. The motivational outcome category was the only non-cognitive outcome category for which effect sizes could be estimated across all intervention types. The effect sizes of our intervention types on motivation did not differ significantly.

Interventions implemented in mathematics showed positive effects for motivational ($d = 0.27$, 95% CI [0.12, 0.52]), affective ($d = 0.33$, 95% CI [0.06, 0.60]), and attitudinal outcomes ($d = 0.31$, 95% CI [0.14, 0.48]). Interventions in medical science showed positive effects for motivational ($d = 0.45$, 95% CI [0.14, 0.86]) and affective outcomes ($d = 0.55$, 95% CI [0.22, 0.88]). For science, motivational ($d = 0.49$, 95% CI [0.34, 0.64]) and attitudinal interventions ($d = 0.41$, 95% CI [0.16, 0.66]) showed positive effects. Interventions in STEM showed positive effects for motivational ($d = 0.36$, 95% CI [0.21, 0.51]), attitudinal ($d = 0.46$, 95% CI [0.19, 0.73]), and miscellaneous outcomes ($d = 0.38$, 95% CI [0.38, 0.38]). The interventions in the mixed-subject category showed positive effects on attitudinal outcomes ($d = 1.02$, 95% CI [0.65, 1.39]).

Implementations of educational technology showed positive effects for mathematics ($d = 0.22$, 95% CI [0.07, 0.35]), medical science ($d = 0.56$, 95% CI [0.23, 0.89]), and STEM ($d = 0.34$, 95% CI [0.27, 0.41]). Games showed positive effects for mathematics ($d = 0.29$, 95% CI [0.25, 0.34]) and science ($d = 0.43$, 95% CI [0.30, 0.56]). Student-centered teaching methods showed positive effects in mathematics ($d = 0.36$, 95% CI [0.03, 0.69]), science ($d = 0.52$, 95% CI [0.39, 0.65]), and the mixed-subject category ($d = 0.80$, 95% CI [0.41, 1.19]). Finally, the miscellaneous interventions overall showed a positive effect for mathematics ($d = 0.43$, 95% CI [0.10, 0.76]), science ($d = 0.43$, 95% CI [−0.02, 0.88]), and STEM ($d = 0.38$, 95% CI [0.25, 0.51]).

Discussion

The primary aim of this second-order meta-analysis was to provide an overview of existing meta-analyses of interventions for non-cognitive outcomes in STEMM education and to examine their effectiveness. We focused on motivational, affective, and attitudinal outcomes. We included findings from 42 meta-analyses, based on 1903 primary effect sizes. In our dataset, we observed a similar number of meta-analyses focusing on motivational, affective, and attitudinal outcomes.

Our study found a significant positive effect ($d = 0.42$) of the meta-analyzed interventions on non-cognitive outcomes, independent of outcome category. We found similar effect sizes for motivational ($d = 0.41$), affective ($d = 0.46$), and attitudinal ($d = 0.39$) outcomes. In the following, we will discuss our results for each of the three outcome categories in more detail. For each of the three outcome categories, we first discuss the variables assessed, then the effectiveness of different intervention types, and finally, potential differences across STEMM subjects.

Motivational outcomes

Mostly, effect sizes on motivational outcomes were reported as an undifferentiated combination of different motivational variables. For some of these effect sizes, the authors only mentioned that the effect was on motivation, without providing additional details (e.g., Baysal et al., 2023). In other cases, the authors specified which motivational variables were assessed but did not differentiate the effects across variables (e.g., Lamb et al., 2018). Fewer effect sizes focused on facets of motivation, such as (self-)confidence (e.g., Baashar et al., 2022) and interest (e.g., Lawner et al., 2019).

The different intervention types, applications of educational technology, game-based learning, student-centered teaching methods, and miscellaneous interventions showed similar, small-to-medium effect sizes for motivational outcomes. Games can be highly motivating and engaging through satisfying needs for autonomy, competence, and relatedness (Ryan & Rigby, 2020). Student-centered teaching methods position the learner to be more actively involved in generating knowledge (Pedaste et al., 2015). This could lead to increased motivation through greater autonomy and competence. Different applications of educational technology can increase motivation by, for example, bringing novelty to STEMM education lessons (e.g., Jeno et al., 2019). Although our second-order meta-analysis showed robust, positive effects of educational technology interventions on motivation overall, there is a need for investigating long-term effects on motivation for specific types of technology.

The interventions on motivational outcomes were effective in mathematics, medical science, science, and STEM. Similar, small-to-medium effects were found for all STEMM subjects investigated. The effects did not differ significantly. These results suggest promising avenues for interventions to improve motivation across STEMM subjects. A focus of future meta-analyses could be on investigating different effects of interventions for male and female students, as there are substantial differences in motivational variables across the genders in some STEMM subjects, and only some meta-analyses included different effects for boys and girls (Lesperance et al., 2022).

Affective outcomes

Most effect sizes on affective outcomes of interventions focused on anxiety and satisfaction, with single effect sizes focusing on stress and depression. This indicates significant gaps in interventional research on affective outcomes in STEMM education. There are multiple affective sub-facets (e.g., Pekrun, 2006) that are significantly related to academic performance and other affective variables (Camacho-Morles et al., 2021; Tze et al., 2016). However, there are probably no interventions, or not enough of them, for a substantial meta-analysis focused on sub-facets such as boredom, enjoyment, or anger. This is an avenue for future research on STEMM education interventions.

The different intervention types investigated in the included meta-analyses—educational technology applications and miscellaneous interventions—showed similar small-to-medium effect sizes for affective outcomes. Various applications of educational technology were found to effectively increase students' satisfaction during concrete lessons. Anxiety was reduced following miscellaneous interventions. The miscellaneous interventions, which consisted of therapeutic, skill-based, and game-based interventions, were effective in reducing anxiety, with stronger effects following therapeutic interventions than the other types of interventions. Further investigations could consider the scalability of interventions, as some therapeutic interventions may be limited in this regard (e.g., Bicer et al., 2020).

The interventions on affective outcomes were studied only in mathematics and medical science and were effective in both domains. Studies in mathematics showed a small-to-medium effect and studies in medical science a medium effect, with no significant difference between the STEMM subjects. Previous research has shown that anxiety levels differ across STEMM fields (e.g., Boateng et al., 2025). Therefore, it should be investigated whether the interventions are equally effective across STEMM subjects or whether different approaches are needed for STEMM subjects beyond mathematics.

Attitudinal outcomes

Effect sizes on attitudinal outcomes were almost exclusively reported as broad, general attitudes towards STEMM (e.g., Ellington, 2006; Kwon et al., 2016; Pahlke et al., 2014) or as attitudes towards a specific lesson or course (e.g., Saraç, 2018; Ural & Bumen, 2016). Fewer effect sizes focused on a facet of attitudes, such as students' self-concept in a STEMM subject (e.g., Ellington, 2006). As some meta-analyses did not explicitly state how they conceptualized attitudes (e.g., Kwon et al., 2016), uncertainty about the attitudinal outcomes remains.

The different intervention types—applications of educational technology, student-centered teaching methods, and miscellaneous interventions—showed similar, small-to-medium effect sizes for attitudinal outcomes. Intervention types focusing on attitudinal outcomes were comparable to those focusing on motivational outcomes, and their effects were also comparable. Only game-based learning was not investigated for attitudinal outcomes, but in multiple meta-analyses for motivational outcomes.

The interventions on attitudinal outcomes were effective for studies in mathematics, science, STEM, and mixed STEMM subjects. Similar, small-to-medium effects were found for all subjects except for the mixed STEMM subjects, which showed large effects. Due to the small number of effect sizes, this effect did not differ significantly from those for the other STEMM subjects.

To sum up, our results show that interventions were effective, with little variation by outcome type, intervention type, and STEMM subject. Gaps for future research remain: robust meta-analytic evidence is still lacking for several outcome types, particularly sub-facets of motivation, affect, and attitudes. The strengths and limitations of our research, along with the conclusions that can be drawn from our findings, are discussed below.

Strengths and limitations

For their second-order meta-analysis on the relationship between alcohol, drugs, and violence, Duke et al. (2018) stated that this meta-meta-analytic approach "provides a 10,000 foot view" (Duke et al., 2018, p. 246) of the research area in question. This implies the benefits of a wide overview as well as the harm from losing detailed information. The main strengths of this second-order meta-analysis lie in its scope and the information on effectiveness. Because our data encompassed the full range of STEMM disciplines, the entire academic pathway, multiple interventions, and motivational, affective, and attitudinal outcomes, this overview serves as a useful reference for existing interventional meta-analyses, highlighting areas with and without significant research activity. Because the included data are meta-analyses of interventions, we can also make claims about the effectiveness of these interventions, providing important information to complement the overview. Having an overview of multiple meta-analyses, for example, multiple meta-analyses investigating the effect of one intervention type, helps contextualize the individual meta-analyses. The meta-analyses of educational games are an example of this. Four meta-analyses covered the effect of educational games on students' motivation. We find, by synthesizing multiple meta-analyses, that the positive effects of educational games span the entire educational continuum and are observed in both digital and serious games. Additionally, the meta-analyses on games for science and STEM indicate larger effects than those on mathematics and chemistry. The overview of all meta-analytic effects in Supplemental Table S4 facilitates comparison of the different effects on a single topic.

The overall value of the pooled effect size should still be taken with caution. First, we encountered considerable heterogeneity and variances of effect sizes between and within meta-analyses. Thus, while the overall effect is robust, the effects for certain intervention types or specific interventions can deviate significantly from the overall effect. The pooled estimate remains meaningful as a reference point for the included and future meta-analyses, but the heterogeneity of effects within the second-order meta-analysis should be considered to reduce overgeneralization of the overall effect. Second, the overall effect is derived from meta-analyses that aggregate all non-cognitive outcome types. By contrast, our estimates by outcome type (and, when combined with intervention type) provide a more nuanced and interpretable picture of the effects than the pooled average alone. Third, the evidence used for calculating the overall effects stemmed from enough data for the findings to be robust, but probably too large for what could be found for actual implementations of these interventions in practice. Several investigations have shown that interventions have larger effects on treatment-inherent measures than on independent measures (Slavin & Madden, 2011). This can also be found for small-scale trials compared to larger trials, quasi-experiments compared to RCTs, and unpublished documents compared to published articles (Cheung & Slavin, 2016). The effect sizes included in our analyses comprised primary effects from small- and large-scale trials, published and unpublished documents, RCTs, and quasi-experiments. It is largely unclear what the effect of RCTs (or large-scale trials, or unpublished studies) is on effect sizes, making an exact estimation of the impact of those factors impossible for our analysis specifically. Additionally, differences in the concrete meta-analytic methods (e.g., outlier adjustment or methods for identifying and correcting publication bias) can distort the results of a second-order meta-analysis and potentially inflate them for certain intervention or outcome types. Still, assuming the estimated overall effect is at the upper bound of the actual effect seems reasonable.

This leads to another limitation of our analysis: the second-order meta-analysis perspective provides a broad overview, but the effects of smaller intervention-level features such as duration, educational level, or sample size were not assessed. These effects are not only relevant for understanding how these features affect educational interventions in general (Cheung & Slavin, 2016), but also for informing practitioners about the optimal conditions for using the analyzed interventions in their own educational practice. The effect of smaller intervention-level features could have been estimated by extracting primary studies from the meta-analyses and then computing their effects. Because some meta-analyses did not provide a list of the primary studies they included, this approach would have resulted in a significant loss of data, which we found inappropriate for the goal of providing a broad, comprehensive overview. Another difficulty in aggregating effects across multiple meta-analyses is the potential for publication bias to be amplified within each meta-analysis. Still, we do not think that this aspect had a strong impact on our results. We also checked all meta-analyses ($n = 36$) that investigated publication bias. Only nine meta-analyses reported indications of publication bias, and six of these reported corrected effect sizes, which we used for our analyses. As a limitation, we did not thoroughly investigate the methods used to identify or correct publication bias, which leaves the possibility that bias-free effects are smaller. Risk of bias (e.g., selection or reporting bias) in the primary studies was assessed in only 26% of the included meta-analyses. This is a limitation of the included meta-analyses, and thus of this second-order meta-analysis, as these biases at the primary study level could distort the effect estimates at the meta-analytic and second-order meta-analytic levels. The overall effect and the effects for certain moderator levels could be overestimated by specific biased primary studies.

While second-order estimates should be interpreted with caution, the consistently small-to-medium effect sizes across more than 1900 primary studies highlight the significant potential of interventions to impact non-cognitive outcomes in STEMM education. Notably, the average effect ($d = 0.42$) is comparable in magnitude to, or even exceeds, typical effects found in learning-focused interventions (Hattie, 2008), despite significant variability in effect sizes. Thus, non-cognitive interventions deserve greater prominence in policy and practice, not as supplementary but as central levers for fostering sustainable engagement in STEMM fields.

Another limitation arising from the more distanced perspective of a second-order meta-analysis concerns the level of detail regarding outcomes and interventions. For interventions, the included meta-analyses described the interventions in sufficient detail; however, information on non-cognitive outcomes was lacking in some cases. As can be seen in major motivational theories (e.g., Eccles & Wigfield, 2020; Hidi & Renninger, 2006), there are multiple aspects of motivation in the learning process. Although some meta-analyses explicitly specified the motivational aspects they included (Lamb et al., 2018), others did not (Higgins et al., 2019), making more detailed analyses at the second-order meta-analysis level difficult.

The fact that the moderator analyses for outcome type, intervention type, and STEMM subject were not significant should not be interpreted as a definitive equality among the moderator levels, as wide confidence intervals and some categories with few effect sizes limit the power to detect differences. Thus, indications of lower effects for studies with attitudinal outcomes, studies in mathematics, or for interventions in the context of student-centered teaching methods can be considered.

The second-order meta-analysis can only depict the state of research included in meta-analyses. This means that newer or niche interventions that were not included in meta-analyses may not be represented here. For outcomes, this could also be the case. There may be informative educational interventions in STEMM whose outcomes include boredom or anger and are not included in meta-analyses and thus are not in our data.

Conclusion

Motivational, affective, and attitudinal variables play an integral role in STEMM education. This overview of the current state of meta-analytic research on interventions with non-cognitive outcomes provides practical examples of educational interventions that don't solely focus on performance or cognitive gains. For researchers in the educational domain, this second-order meta-analysis can act as a starting point for future research. Future meta-analyses on the same topic can be compared with those reviewed here; future meta-analyses focusing on a specific affective outcome, e.g., boredom, could reference their effects relative to the overall effectiveness of the included studies in this second-order meta-analysis. This analysis also suggested gaps that could be addressed through meta-analytic research: meta-analyses of interventions targeting affective variables other than anxiety or satisfaction, critical meta-analyses of the impact of the newest technologies, or meta-analyses of integrated STEM or STEMM education (Li, 2026). More fine-grained analyses of what the individual meta-analyses and their included primary studies contain when they research attitudes or motivation towards STEMM were also indicated.

This review presents a structured synthesis of the effectiveness of interventions across motivational, affective, and attitudinal outcomes in STEMM education. Building on this foundation, future meta-analytic research could extend the scope by addressing situational and contextual moderators (e.g., learning setting, cultural variation, or equity-relevant subgroup effects). Such refinements would help unpack the conditions under which interventions are most effective and for whom.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40594-026-00619-w.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Supplementary Material 4.

Supplementary Material 5.

Supplementary Material 6.

Supplementary Material 7.

Supplementary Material 8.

Supplementary Material 9.

Supplementary Material 10.

Acknowledgements
We thank our student assistants Amelie Tahedl, Nicholas Sagberger, and Sophia Horina for their valuable help with data collection. We are also very grateful to Diana Wengler, Maryam Shiani, and Mehmet Bicakci for their help in the data generation process and helpful comments in the writing process.

Author contributions

F.H. conceptualized the study with A.Z. and H.S. and curated the data with C.P., who also contributed to the paper's methodology. F.H. played the lead role in writing the manuscript as well as the formal analysis and writing of the analysis script. A.Z. and H.S. were responsible for the project administration and supervision. All authors critically reviewed and edited the manuscript, and all approved the final version.

Funding
Open Access funding enabled and organized by Projekt DEAL. This research was conducted as part of the project MesH_MINT, funded by the German Federal Ministry for Education, Family Affairs, Senior Citizens, Women and Youth (16DWMQP02A, 16DWMQP02B). The authors assume full responsibility for the contents.

Data availability

The dataset generated and analysed during the current study is available in the OSF repository (https://doi.org/10.17605/OSF.IO/UKESG).

Declarations

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Received: 29 October 2025 / Accepted: 29 April 2026

References

AAAS. (2025). *Science at work: The people and industries powering America's prosperity*. Report. American Association for the Advancement of Science. https://scienceisus.org/wp-content/uploads/2025/04/science-at-work-economic-impact-report-2025-APRIL-FINAL.pdf

Akin, A., & Kurbanoglu, I. N. (2011). The relationships between math anxiety, math attitudes, and self-efficacy: A structural equation model. *Studia Psychologica, 53*(3), 263.

Allport, G. W. (1935). Attitudes. In *A handbook of social psychology* (pp. 798–844). Clark University Press.

Amalina, I. K., Vidákovich, T., & Karimova, K. (2025). Factors influencing student interest in STEM careers: Motivational, cognitive, and socioeconomic status. *Humanities and Social Sciences Communications, 12*(1), 1–15. https://doi.org/10.1057/s41599-025-04446-2

Anger, C., Betz, J., & Plünnecke, A. (2024). *MINT-Herbstreport 2024: MINT-Förderung - Der Schlüssel zu Innovation und Wachstum* [STEM Autumn Report 2024]. Gutachten für BDA, Gesamtmetall und MINT Zukunft schaffen.

Anger, C., & Plünnecke, A. (2022). *MINT gewinnt: Hohe Löhne in den MINT-Berufen [STEM Wins: High Wages in STEM Jobs]*.

Arztmann, M., Hornstra, L., Jeuring, J., & Kester, L. (2023). Effects of games in STEM education: A meta-analysis on the moderating role of student background characteristics. *Studies in Science Education, 59*(1), 109–145. https://doi.org/10.1080/03057267.2022.2057732

Asher, M. W., Harackiewicz, J. M., Beymer, P. N., Hecht, C. A., Lamont, L. B., Else-Quest, N. M., Priniski, S. J., Thoman, D. B., Hyde, J. S., & Smith, J. L. (2023). Utility-value intervention promotes persistence and diversity in STEM. *Proceedings of the National Academy of Sciences of the United States of America, 120*(19), Article e2300463120. https://doi.org/10.1073/pnas.2300463120

Baashar, Y., Alkawsi, G., Ahmad, W. N. W., Alhussian, H., Alwadain, A., Capretz, L. F., Babiker, A., & Alghail, A. (2022). Effectiveness of using augmented reality for training in the medical professions: Meta-analysis. *JMIR Serious Games, 10*(3), Article e32715. https://doi.org/10.2196/32715

Bai, H., Pan, W., Hirumi, A., & Kebritchi, M. (2012). Assessing the effectiveness of a 3-D instructional game on improving mathematics achievement and motivation of middle school students. *British Journal of Educational Technology, 43*(6), 993–1003. https://doi.org/10.1111/j.1467-8535.2011.01269.x

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

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall Inc.

Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin, 147*(2), 134–168. https://doi.org/10.1037/bul0000307

Baysal, Y. E., Mutlu, F., & Nacaroglu, O. (2023). The effectiveness of student-centered teaching applications used in determining motivation toward science learning: A meta-analysis study. *International Journal of Psychology and Educational Studies, 10*(1), 1–21.

Bicer, A., Perihan, C., & Lee, Y. (2020). A meta-analysis: The effects of CBT as a clinic- & school-based treatment on students' mathematics anxiety. *International Electronic Journal of Mathematics Education, 15*(2), Article em0576. https://doi.org/10.29333/iejme/7598

Boateng, S., Mudadigwa, B., & Johnston-Wilder, S. (2025). Examining gendered patterns in mathematics and science anxiety levels among physical science pre-service teachers. *Eurasia Journal of Mathematics, Science and Technology Education, 21*(1), Article em2564. https://doi.org/10.29333/ejmste/15800

Bohner, G., & Dickel, N. (2011). Attitudes and attitude change. *Annual Review of Psychology, 62*, 391–417. https://doi.org/10.1146/annurev.psych.121208.131609

Borenstein, M. (Ed.). (2009). *Effect sizes for continuous data*. Russell Sage Foundation.

Brathwaite, D. A. (2010). *A comparative analysis of single-sex schools in terms of achievement in reading and math and student* attendance*. University of Dayton.

Burnette, J. L., Hoyt, C. L., Russell, V. M., Lawson, B., Dweck, C. S., & Finkel, E. (2020). A growth mind-set intervention improves interest but not academic performance in the field of computer science. *Social Psychological and Personality Science, 11*(1), 107–116. https://doi.org/10.1177/1948550619841631

Camacho-Morles, J., Slemp, G. R., Pekrun, R., Loderer, K., Hou, H., & Oades, L. G. (2021). Activity achievement emotions and academic performance: A meta-analysis. *Educational Psychology Review, 33*(3), 1051–1095. https://doi.org/10.1007/s10648-020-09585-3

Chen, B., Chen, J., Wang, M., Tsai, C. -C., & Kirschner, P. A. (2025). The effects of integrated STEM education on K12 students' achievements: A meta-analysis. *Review of Educational Research*. https://doi.org/10.3102/00346543251318297

Cheryan, S., Drury, B. J., & Vichayapai, M. (2013). Enduring influence of stereotypical computer science role models on women's academic aspirations. *Psychology of Women Quarterly, 37*(1), 72–79. https://doi.org/10.1177/0361684312459328

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

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. *Routledge*. https://doi.org/10.4324/9780203771587

Cuder, A., Pellizzoni, S., Di Marco, M., Blason, C., Doz, E., Giofrè, D., & Passolunghi, M. C. (2024). The impact of math anxiety and self-efficacy in middle school STEM choices: A 3-year longitudinal study. *British Journal of Educational Psychology, 94*(4), 1091–1108. https://doi.org/10.1111/bjep.12707

Daker, R. J., Gattas, S. U., Sokolowski, H. M., Green, A. E., & Lyons, I. M. (2021). First-year students' math anxiety predicts STEM avoidance and underperformance throughout university, independently of math ability. *NPJ Science of Learning, 6*, Article Article 17. https://doi.org/10.1038/s41539-021-00095-7

Deeks, J. J., Higgins, J. P. T., & Altman, D. G. (Eds.). (2019). *Cochrane handbook for systematic reviews of interventions. 2nd Edition. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors)*. Wiley.

Dondio, P., Gusev, V., & Rocha, M. (2023). Do games reduce math anxiety? A meta-analysis. *Computers & Education, 194*, Article Article 104650. https://doi.org/10.1016/j.compedu.2022.104650

Duke, A. A., Smith, K. M. Z., Oberleitner, L. M. S., Westphal, A., & McKee, S. A. (2018). Alcohol, drugs, and violence: A meta-meta-analysis. *Psychology of Violence, 8*(2), 238–249. https://doi.org/10.1037/vio0000106

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology, 61*, Article 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Eidlin-Levy, H., Avraham, E., Fares, L., & Rubinsten, O. (2023). Math anxiety affects career choices during development. *International Journal of STEM Education*. https://doi.org/10.1186/s40594-023-00441-8

Ellington, A. J. (2006). The effects of non-CAS graphing calculators on student achievement and attitude levels in mathematics: A meta-analysis. *School Science and Mathematics, 106*(1), 16–26. https://doi.org/10.1111/j.1949-8594.2006.tb18067.x

Fadda, D., Pellegrini, M., Vivanet, G., & Zandonella Callegher, C. (2022). Effects of digital games on student motivation in mathematics: A meta-analysis in K-12. *Journal of Computer Assisted Learning, 38*(1), 304–325. https://doi.org/10.1111/jcal.12618

Ferdinand, R., Malanchini, M., & Rimfeld, K. (2024). Mathematics interest, self-efficacy, and anxiety predict STEM career choice in emerging adulthood. *NPJ Science of Learning, 9*, Article 66. https://doi.org/10.1038/s41539-024-00275-1

Galoyan, T., Barany, A., Donaldson, J. P., Ward, N., & Hammrich, P. (2022). Connecting science, design thinking, and computational thinking through sports. *International Journal of Instruction, 15*(1), 601–618. https://doi.org/10.29333/iji.2022.15134a

de Gioannis, E. (2022). Implicit gender-science stereotypes and college-major intentions of Italian adolescents. *Social Psychology of Education, 25*(5), 1093–1112. https://doi.org/10.1007/s11218-022-09709-3

Gurevitch, J., Koricheva, J., Nakagawa, S., & Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature, 555*(7695), 175–182. https://doi.org/10.1038/nature25753

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Hernik, J., & Jaworska, E. (2018). The effect of enjoyment on learning. In L. Gómez Chova, A. López Martínez, & I. Candel Torres (Eds.), *INTED Proceedings, INTED2018 Proceedings* (pp. 508–514). IATED. https://doi.org/10.21125/inted.2018.1087

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Higgins, K., Huscroft-D'Angelo, J., & Crawford, L. (2019). Effects of technology in mathematics on achievement, motivation, and attitude: A meta-analysis. *Journal of Educational Computing Research, 57*(2), 283–319. https://doi.org/10.1177/0735633117748416

Hong, Z.-R. (2010). Effects of a collaborative science intervention on high achieving students' learning anxiety and attitudes toward science. *International Journal of Science Education, 32*(15), 1971–1988. https://doi.org/10.1080/09500690903229304

Hunt, J. H., Taub, M., Marino, M., Holman, K., & Womack-Adams, K. (2025). Increasing student engagement, fraction knowledge, and STEM interest through game-based intervention. *Journal of Special Education Technology, 0*(0), 1–13. https://doi.org/10.1177/01626434251314014

Ioannidis, J. P. A. (2016). The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *The Milbank Quarterly, 94*(3), 485–514. https://doi.org/10.1111/1468-0009.12210

Jamieson, J. P., Peters, B. J., Greenwood, E. J., & Altose, A. J. (2016). Reappraising stress arousal improves performance and reduces evaluation anxiety in classroom exam situations. *Social Psychological and Personality Science, 7*(6), 579–587. https://doi.org/10.1177/1948550616644656

Jansen, T., Meyer, J., Hattie, J., & Möller, J. (2024). Who am I? A second-order meta-analytic review of correlates of the self in childhood and adolescence. *Psychological Bulletin, 150*(11), 1287–1317. https://doi.org/10.1037/bul0000449

Jeno, L. M., Vandvik, V., Eliassen, S., & Grytnes, J.-A. (2019). Testing the novelty effect of an m-learning tool on internalization and achievement: A Self-Determination Theory approach. *Computers & Education, 128*, 398–413. https://doi.org/10.1016/j.compedu.2018.10.008

Jeong, H., Hmelo-Silver, C. E., & Jo, K. (2019). Ten years of computer-supported collaborative learning: A meta-analysis of CSCL in STEM education during 2005-2014. *Educational Research Review, 28*, Article 100284. https://doi.org/10.1016/j.edurev.2019.100284

Jiang, S., Simpkins, S. D., & Eccles, J. S. (2020). Individuals' math and science motivation and their subsequent STEM choices and achievement in high school and college: A longitudinal study of gender and college generation status differences. *Developmental Psychology, 56*(11), 2137–2151. https://doi.org/10.1037/dev0001110

Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review, 100*(1), 109–128. https://doi.org/10.1037/0033-295X.100.1.109

Jung, A. -R., & Park, E. -A. (2022). The effectiveness of learning to use HMD-based VR technologies on nursing students: Chemoport insertion surgery. *International Journal of Environmental Research and Public Health, 19*(8), Article 4823. https://doi.org/10.3390/ijerph19084823

Kaplan, A., & Maehr, M. L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review, 19*(2), 141–184. https://doi.org/10.1007/s10648-006-9012-5

Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education, 55*(2), 427–443. https://doi.org/10.1016/j.compedu.2010.02.007

Kessels, U., & Hannover, B. (2008). When being a girl matters less: Accessibility of gender-related self-knowledge in single-sex and coeducational classes and its impact on students' physics-related self-concept of ability. *British Journal of Educational Psychology, 78*(2), 273–289. https://doi.org/10.1348/000709907X215938

Knogler, M., Hetmanek, A., & Seidel, T. (2022). Determining an evidence base for particular fields of educational practice: A systematic review of meta-analyses on effective mathematics and science teaching. *Frontiers in Psychology, 13*, Article 873995. https://doi.org/10.3389/fpsyg.2022.873995

Kollmayer, M., Schober, B., & Spiel, C. (2018). Gender stereotypes in education: Development, consequences, and interventions. *European Journal of Developmental Psychology, 15*(4), 361–377. https://doi.org/10.1080/17405629.2016.1193483

Krannich, M., Goetz, T., Lipnevich, A. A., Bieg, M., Roos, A.-L., Becker, E. S., & Morger, V. (2019). Being over- or underchallenged in class: Effects on students' career aspirations via academic self-concept and boredom. *Learning and Individual Differences, 69*, 206–218. https://doi.org/10.1016/j.lindif.2018.10.004

Kwon, H., Lee, E., & Lee, D. (2016). Meta-analysis on the effectiveness of invention education in South Korea: Creativity, attitude, and tendency for problem solving. *Journal of Baltic Science Education, 15*(1), 48–57. https://doi.org/10.33225/jbse/16.15.48

Lamb, R. L., Annetta, L., Firestone, J., & Etopio, E. (2018). A meta-analysis with examination of moderators of student cognition, affect, and learning outcomes while using serious educational games, serious games, and simulations. *Computers in Human Behavior, 80*, 158–167. https://doi.org/10.1016/j.chb.2017.10.040

Law, F., McGuire, L., Winterbottom, M., & Rutland, A. (2021). Children's gender stereotypes in STEM following a one-shot growth mindset intervention in a science museum. *Frontiers in Psychology, 12*, Article 641695. https://doi.org/10.3389/fpsyg.2021.641695

Lawner, E. K., Quinn, D. M., Camacho, G., Johnson, B. T., & Pan-Weisz, B. (2019). Ingroup role models and underrepresented students' performance and interest in STEM: A meta-analysis of lab and field studies. *Social Psychology of Education, 22*(5), 1169–1195. https://doi.org/10.1007/s11218-019-09518-1

Lazarides, R., & Lauermann, F. (2019). Gendered paths into STEM-related and language-related careers: Girls' and boys' motivational beliefs and career plans in Math and Language Arts. *Frontiers in Psychology, 10*, Article 1243. https://doi.org/10.3389/fpsyg.2019.01243

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist, 46*(8), 819–834. https://doi.org/10.1037/0003-066X.46.8.819

Leinonen, R., Asikainen, M. A., & Hirvonen, P. E. (2013). Overcoming students' misconceptions concerning thermal physics with the aid of hints and peer interaction during a lecture course. *Physical Review Special Topics - Physics Education Research, 9*(2), Article 020112. https://doi.org/10.1103/PhysRevSTPER.9.020112

Lenhard, W., & Lenhard, A. (2017). *Computation of effect sizes*. https://www.psychometrica.de/effect_size.html. https://doi.org/10.13140/RG.2.2.17823.92329

Lesperance, K., Hofer, S., Retelsdorf, J., & Holzberger, D. (2022). Reducing gender differences in student motivational-affective factors: A meta-analysis of school-based interventions. *British Journal of Educational Psychology, 92*(4), 1502–1536. https://doi.org/10.1111/bjep.12512

Li, Y. (2026). Trends of high impact educational research in scientific disciplines. *International Journal of STEM Education*. https://doi.org/10.1186/s40594-026-00607-0

Li, Q., Cho, H., Cosso, J., & Maeda, Y. (2021). Relations between students' mathematics anxiety and motivation to learn mathematics: A meta-analysis. *Educational Psychology Review, 33*(3), 1017–1049. https://doi.org/10.1007/s10648-020-09589-z

Liang, K., Ma, X., Zhang, J., Zhang, B., Xin, F., Yuan, Q., Zhang, D., Chen, R., & Tang, Y. (2023). A pilot Movement Integrity with Intelligent Play Program (MIIP): Effects on Math performance and enjoyment for preschoolers in China. *Child & Youth Care Forum, 52*(3), 703–719. https://doi.org/10.1007/s10566-022-09707-8

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

Marín-Martínez, F., & Sánchez-Meca, J. (2010). Weighting by inverse variance or by sample size in random-effects meta-analysis. *Educational and Psychological Measurement, 70*(1), 56–73. https://doi.org/10.1177/0013164409344534

Moher, D., Liberati, A., Tetzlaff, J. M., Altman, D. G., the PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *Annals of Internal Medicine, 151*(4), 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Mohr-Schroeder, M. J., Jackson, C., Miller, M., Walcott, B., Little, D. L., Speler, L., Schooler, W., & Schroeder, D. C. (2014). Developing middle school students' interests in STEM via summer learning experiences: See Blue STEM Camp. *School Science and Mathematics, 114*(6), 291–301. https://doi.org/10.1111/ssm.12079

Morris, T. T., Dorling, D., Davies, N. M., & Davey Smith, G. (2021). Associations between school enjoyment at age 6 and later educational achievement: Evidence from a UK cohort study. *NPJ Science of Learning, 6*(1), 18. https://doi.org/10.1038/s41539-021-00092-w

Munder, T., Brütsch, O., Leonhart, R., Gerger, H., & Barth, J. (2013). Researcher allegiance in psychotherapy outcome research: An overview of reviews. *Clinical Psychology Review, 33*(4), 501–511. https://doi.org/10.1016/j.cpr.2013.02.002

Namkung, J. M., Peng, P., & Lin, X. (2019). The relation between mathematics anxiety and mathematics performance among school-aged students: A meta-analysis. *Review of Educational Research, 89*(3), 459–496. https://doi.org/10.3102/0034654319843494

National Science Board, & National Science Foundation. (2024). *Science and Engineering Indicators 2024: The State of U.S. Science and Engineering*. NSB-2024-3. https://ncses.nsf.gov/pubs/nsb20243

O'Keefe, P. A., Horberg, E. J., Dweck, C. S., & Walton, G. M. (2023). A Growth-Theory-of-Interest Intervention increases interest in Math and Science coursework among Liberal Arts undergraduates. *Journal of Educational Psychology, 115*(6), 859–876. https://doi.org/10.1037/edu0000798

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education, 25*(9), 1049–1079. https://doi.org/10.1080/0950069032000032199

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 Statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

Pahlke, E., Hyde, J. S., & Allison, C. M. (2014). The effect of single-sex compared with coeducational schooling on students' performance and attitudes: A meta-analysis. *Psychological Bulletin, 140*(4), 1042–1072. https://doi.org/10.1037/a0035740

Passolunghi, M. C., de Vita, C., & Pellizzoni, S. (2020). Math anxiety and Math achievement: The effects of emotional and Math strategy training. *Developmental Science, 23*(6), Article e12964. https://doi.org/10.1111/desc.12964

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review, 14*, 47–61. https://doi.org/10.1016/j.edurev.2015.02.003

Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions. In *Advances in psychology, 131. Motivational psychology of human development: Developing motivation and motivating development* (pp. 143–163). Elsevier Science. https://doi.org/10.1016/S0166-4115(00)80010-2

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Piesch, H., Gaspard, H., Parrisius, C., Wille, E., & Nagengast, B. (2020). How can a relevance intervention in math support students' career choices? *Journal of Applied Developmental Psychology, 71*, Article 101185. https://doi.org/10.1016/j.appdev.2020.101185

R Core Team. (2020). *R: a language and environment for statistical computing*. https://www.R-project.org/

Ramos-Sandoval, R., & Ramos-Diaz, J. (2024). Does gender matter in Peruvian engineering student's career goals? *International Journal for Educational and Vocational Guidance*. https://doi.org/10.1007/s10775-024-09718-y

Respondek, L., Seufert, T., Stupnisky, R., & Nett, U. E. (2017). Perceived academic control and academic emotions predict undergraduate university student success: Examining effects on dropout intention and achievement. *Frontiers in Psychology, 8*, Article 243. https://doi.org/10.3389/fpsyg.2017.00243

Rocha, M., & Dondio, P. (2021). Effects of a videogame in Math performance and anxiety in primary school. *International Journal of Serious Games, 8*(3), 45–70. https://doi.org/10.17083/ijsg.v8i3.434

Rosenzweig, E. Q., Harackiewicz, J. M., Priniski, S. J., Hecht, C. A., Canning, E. A., Tibbetts, Y., & Hyde, J. S. (2019). Choose your own intervention: Using choice to enhance the effectiveness of a utility-value intervention. *Motivation Science, 5*(3), Article 269.

RStudio Team. (2020). *RStudio: Integrated development environment for R*. http://www.rstudio.com/

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Ryan, R. M., & Rigby, C. S. (2020). Motivational foundations of game-based learning. In *Handbook of game-based learning* (pp. 153–176). The MIT Press.

Sakellariou, C., & Fang, Z. (2021). Self-efficacy and interest in STEM subjects as predictors of the STEM gender gap in the US: The role of unobserved heterogeneity. *International Journal of Educational Research, 109*, Article 101821. https://doi.org/10.1016/j.ijer.2021.101821

Saraç, H. (2018). The effect of science, technology, engineering and mathematics-STEM educational practices on students' learning outcomes: A meta-analysis study. *Turkish Online Journal of Educational Technology - TOJET, 17*(2), 125–142.

Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings*. SAGE Publications Ltd. https://doi.org/10.4135/9781483398105

Schmidt, F. L., & Oh, I. -S. (2013). Methods for second order meta-analysis and illustrative applications. *Organizational Behavior and Human Decision Processes, 121*(2), 204–218. https://doi.org/10.1016/j.obhdp.2013.03.002

Schunck, D. H., Meece, J. L., & Pintrich, P. R. (2014). *Motivation in education: Theory, research and applications* (4th ed.). Pearson.

Schunk, D. H., & Usher, E. L. (2012). Social cognitive theory and motivation. In R. M. Ryan (Ed.), *The oxford handbook of human motivation* (pp. 13–27). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195399820.013.0002

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*(3), 407–441. https://doi.org/10.3102/00346543046003407

Shimizu, Y. (2025). Relation between mathematics self-efficacy, mathematics anxiety, behavioural engagement, and mathematics achievement in Japan. *Psychology International, 7*(2), 36. https://doi.org/10.3390/psycholint7020036

Shin, D. D., Lee, M., Jung, S. J., & Bong, M. (2023). Relative effects of classroom utility value intervention on the science motivation of girls and boys. *Research in Science Education, 53*(3), 593–612. https://doi.org/10.1007/s11165-022-10070-w

Slavin, R., & Madden, N. A. (2011). Measures inherent to treatments in program effectiveness reviews. *Journal of Research on Educational Effectiveness, 4*(4), 370–380. https://doi.org/10.1080/19345747.2011.558986

Song, J., Zuo, B., Wen, F., & Yan, L. (2017). Math-gender stereotypes and career intentions: An application of expectancy-value theory. *British Journal of Guidance & Counselling, 45*(3), 328–340. https://doi.org/10.1080/03069885.2017.1283680

Stoeger, H., Duan, X., Schirner, S., Greindl, T., & Ziegler, A. (2013). The effectiveness of a one-year online mentoring program for girls in STEM. *Computers & Education, 69*, 408–418. https://doi.org/10.1016/j.compedu.2013.07.032

Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. A. (2011). Steming the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology, 100*(2), 255–270. https://doi.org/10.1037/a0021385

Supekar, K., Iuculano, T., Chen, L., & Menon, V. (2015). Remediation of childhood math anxiety and associated neural circuits through cognitive tutoring. *The Journal of Neuroscience, 35*(36), 12574–12583. https://doi.org/10.1523/JNEUROSCI.0786-15.2015

Tamim, R. M., Borokhovski, E., Bernard, R. M., Schmid, R. F., Abrami, P. C., & Pickup, D. I. (2021). A study of meta-analyses reporting quality in the large and expanding literature of educational technology. *Australasian Journal of Educational Technology, 37*(4), 100–115. https://doi.org/10.14742/ajet.6322

Ting, F. S. T., Shroff, R. H., Lam, W. H., Garcia, R. C. C., Chan, C. L., Tsang, W. K., & Ezeamuzie, N. O. (2023). A meta-analysis of studies on the effects of active learning on Asian students' performance in science, technology, engineering and mathematics (STEM) subjects. *The Asia-Pacific Education Researcher, 32*(3), 379–400. https://doi.org/10.1007/s40299-022-00661-6

Tipton, E., Pustejovsky, J. E., & Ahmadi, H. (2019). Current practices in meta-regression in psychology, education, and medicine. *Research Synthesis Methods, 10*(2), 180–194. https://doi.org/10.1002/jrsm.1339

Tsai, S.-P., Ting, Y.-L., & Chu, L. (2023). Development and preliminary validation of the middle school students' attitudes toward STEM learning scale. *Research in Science & Technological Education*, 1–21. https://doi.org/10.1080/02635143.2023.2235291

Tze, V. M. C., Daniels, L. M., & Klassen, R. M. (2016). Evaluating the relationship between boredom and academic outcomes: A meta-analysis. *Educational Psychology Review, 28*(1), 119–144. https://doi.org/10.1007/s10648-015-9301-y

U.S. Bureau of Labor Statistics. (2022). *Periodic table of science, technology, engineering and math occupations*. https://www.bls.gov/k12/students/careers/stem-table.htm

U.S. Bureau of Labor Statistics. (2024). *Employment in STEM occupations: U.S. Bureau of Labor Statistics*. https://www.bls.gov/emp/tables/stem-employment.htm

Ural, G., & Bumen, N. (2016). A meta-analysis on instructional applications of constructivism in science and technology teaching: A sample of Turkey. *Education and Science, 41*(185), 51–82. https://doi.org/10.15390/EB.2016.4289

Urhahne, D., & Wijnia, L. (2023). Theories of motivation in education: An integrative framework. *Educational Psychology Review*, *35*(2). https://doi.org/10.1007/s10648-023-09767-9

van den Berg, V., Singh, A. S., Komen, A., Hazelebach, C., van Hilvoorde, I., & Chinapaw, M. J. M. (2019). Integrating juggling with math lessons: A randomized controlled trial assessing effects of physically active learning on maths performance and enjoyment in primary school children. *International Journal of Environmental Research and Public Health, 16*(14), Article 2452. https://doi.org/10.3390/ijerph16142452

Vanbecelaere, S., van den Berghe, K., Cornillie, F., Sasanguie, D., Reynvoet, B., & Depaepe, F. (2020). The effects of two digital educational games on cognitive and non-cognitive math and reading outcomes. *Computers & Education, 143*, Article 103680. https://doi.org/10.1016/j.compedu.2019.103680

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

Wang, H.-H., Lin, H., Chen, Y.-C., Pan, Y.-T., & Hong, Z.-R. (2021). Modelling relationships among students' inquiry-related learning activities, enjoyment of learning, and their intended choice of a future STEM career. *International Journal of Science Education, 43*(1), 157–178. https://doi.org/10.1080/09500693.2020.1860266

Wang, Z., Lukowski, S. L., Hart, S. A., Lyons, I. M., Thompson, L. A., Kovas, Y., Mazzocco, M. M. M., Plomin, R., & Petrill, S. A. (2015). Is math anxiety always bad for math learning? The role of math motivation. *Psychological Science, 26*(12), 1863–1876. https://doi.org/10.1177/0956797615602471

Weiner, B. (1986). *An attributional theory of motivation and emotion*. Springer US. https://doi.org/10.1007/978-1-4612-4948-1

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wille, E., Stoll, G., Gfrörer, T., Cambria, J., Nagengast, B., & Trautwein, U. (2020). It takes two: Expectancy-value constructs and vocational interests jointly predict STEM major choices. *Contemporary Educational Psychology, 61*, Article 101858. https://doi.org/10.1016/j.cedpsych.2020.101858

Wu, H., Guo, Y., Yang, Y., Zhao, L., & Guo, C. (2021). A meta-analysis of the longitudinal relationship between academic self-concept and academic achievement. *Educational Psychology Review, 33*(4), 1749–1778. https://doi.org/10.1007/s10648-021-09600-1

Xie, F., Yang, Y., & Xiao, C. (2023). Gender-math stereotypes and mathematical performance: The role of attitude toward mathematics and math self-concept. *European Journal of Psychology of Education, 38*(2), 695–708. https://doi.org/10.1007/s10212-022-00631-y

Yildirim, H. I. (2018). The impact of out-of-school learning environments on 6th grade secondary school students attitude towards science course. *Journal of Education and Training Studies, 6*(12), 26–41. https://doi.org/10.11114/jets.v6i12.3624

Young, J. R., Ortiz, N., & Young, J. L. (2017). Stemulating interest: A meta-analysis of the effects of out-of-school time on student STEM interest. *International Journal of Education in Mathematics, Science and Technology, 5*(1), 62–74. https://doi.org/10.18404/ijemst.61149

Zhao, F., Zhang, Y., Alterman, V., Zhang, B., & Yu, G. (2018). Can math-gender stereotypes be reduced? A theory-based intervention program with adolescent girls. *Current Psychology, 37*(3), 612–624. https://doi.org/10.1007/s12144-016-9543-y

Zheng, P., You, Y., & Luo, G. (2025). Math interest on primary school students' mathematics achievement: Self-control/resilience chain mediation. *Journal of Psychoeducational Assessment*. https://doi.org/10.1177/07342829251326435

Živković, M., Pellizzoni, S., Doz, E., Cuder, A., Mammarella, I., & Passolunghi, M. C. (2023). Math self-efficacy and anxiety? The role of emotional and motivational contribution in math performance. *Social Psychology of Education, 26*(3), 579–601. https://doi.org/10.1007/s11218-023-09760-8

Publisher's Note