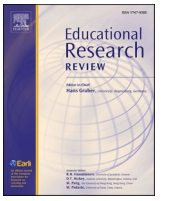
Educational Research Review 45 (2024) 100639



Contents lists available at ScienceDirect

Educational Research Review

journal homepage: www.elsevier.com/locate/edurev



Review



Learning with visualizations helps: A meta-analysis of visualization interventions in mathematics education

Johanna Schoenherr [a,*], Anselm R. Strohmaier [b], Stanislaw Schukajlow [c]

[a] Paderborn University, Faculty of Computer Science, Electrical Engineering and Mathematics, Warburger Str. 100, 33098, Paderborn, Germany
[b] University of Education Ludwigsburg, Institute of Mathematics II, Reuteallee 46, 71634, Ludwigsburg, Germany
[c] University of Münster, Institute of Mathematics Education and Computer Science Education, Henriette-Son-Straße 19, 48149, Münster, Germany

ARTICLE INFO

Keywords:
Visualization
Mathematics
Learning
Interventions
Meta-analysis

ABSTRACT

Visualization has a long tradition in mathematics education and research on this topic has become more widespread in recent decades. In a meta-analysis (41 studies, $N = 10{,}562$), we aimed to synthesize the effects of learning with external visualizations on mathematics outcomes. We analyzed intervention, learner, and outcome characteristics as moderators. Overall, results of a random-effects model indicated a medium effect ($g = 0.504$, 95% CI [0.379, 0.630]) of visualization interventions on mathematics learning, with significant heterogeneity. Moderator analyses revealed that effect sizes were higher in quasi-experimental studies and when compared with business-as-usual conditions. These results emphasize the effectiveness of external visualization as a powerful tool to support mathematics learning, with positive and lasting effects across age groups and mathematical topics.

1. Introduction

1.1. Objective and rationale

Visualization can be a powerful tool for thinking and learning in mathematics, as it can help students understand abstract concepts and make connections between mathematical ideas (Arcavi, 2003). For example, learning with two-dimensional number lines can assist students' understanding of fractions (Soni & Okamoto, 2020), and creating drawings for a mathematical modeling problems can help students solve the problems (Rellensmann et al., 2022). Thus, many curricula emphasize the use of visualizations to enhance students' mathematics learning and develop their abilities in constructing, using, and interpreting visualizations at all educational levels (e.g., NGA & CCSSO, 2010).

Research on learning with visualizations has a long tradition in mathematics education, as described in Presmeg's (2006) review, which, among other topics, was aimed at understanding the benefits and challenges of learning with visualizations in educational settings. For example, the review identified two potential reasons for why students might not learn better with visualizations: First, many students are reluctant to use visualizations (Presmeg, 2006; Rellensmann et al., 2022; Stylianou & Silver, 2004; Uesaka et al., 2010). Second, students often encounter difficulties in constructing, using, and interpreting visualizations effectively (Boels et al., 2019; Diezmann, 2000; Lem et al., 2013). The question that arises is whether visualizations effectively support learning in mathematics and if so, for whom and under what conditions they are effective. Previous reviews on visualizations specific to mathematics

* Corresponding author.
E-mail address: johanna.schoenherr@uni-paderborn.de (J. Schoenherr).

https://doi.org/10.1016/j.edurev.2024.100639
Received 23 January 2024; Received in revised form 9 September 2024; Accepted 11 September 2024
Available online 12 September 2024
1747-938X/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

learning aimed at systematically describing conceptual difficulties that arise from these specific visualizations, such as histograms (Boels et al., 2019). Meta-analyses examined visualizations for specific groups of learners, such as elementary school students (Sokolowski, 2018), included visualizations besides other tech-enhanced mathematics instruction (Ran et al., 2022; Young, 2017), or focused on specific technology (e.g., Y. Zhang et al., 2023). Thus, a comprehensive systematic review and meta-analysis of the effectiveness of visualization interventions, i.e., an environment where mathematics learners participate in activities designed to improve their mathematics learning with visualizations, and the characteristics of these interventions is needed to gain a more integrated perspective on the benefits of learning with visualizations that are specific to mathematics.

In the present meta-analysis, we synthesized empirical research findings on the effectiveness of visualization interventions to support mathematics learning. By integrating extant research, we sought to examine (1) the effectiveness of visualization interventions for learning mathematics compared with learning without visualizations and (2) investigate which characteristics of the learners, interventions, and outcomes influence the effects of visualization interventions to better understand for whom and when visualization interventions aid mathematics learning.

1.2. External visualizations in mathematics education

External visualizations refer to the product of creating, interpreting, using, and reflecting on images, pictures, and diagrams through various media to convey information, foster innovative ideas, and deepen understanding (Arcavi, 2003; Schoenherr & Schukajlow, 2023b). In contrast to internal visualizations within people's minds, external visualizations are representations of mathematical concepts or data that exist outside of people's minds and often involve the use of physical or digital aids (Goldin & Kaput, 1996). In this review, our focus centers on external visualizations in mathematics education. Throughout this review, the term *visualizations* refers specifically to external visualizations.

Examples of visualizations in mathematics are diverse, including graphs and charts, geometric shapes and figures, physical objects, such as base-ten blocks and fraction tiles, and statistical plots. To support learning, visualizations can be provided by the teacher or researcher or generated in whole or in part by the learners (Q. Zhang & Fiorella, 2019). For example, visualizing rectangular areas as the sweeping motion of one length through another can help students understand the conception of area and its measure as a multiplicative product with commutative and distributive properties (Kobiela & Lehrer, 2019). Creating visual representations of functions holds potential benefits for mathematics majors, aiding in the development of proofs in calculus (Mejia-Ramos & Weber, 2019). When confronted with Bayes problems, students find assistance in being provided with visual aids such as 2x2 tables or tree diagrams (Binder et al., 2015).

Goldin and Kaput (1996) proposed that bidirectional interactions occur between internal and external representations when learners engage with external visualizations. When visualizations are included as additional representations of abstract mathematical concepts, the visualizations can serve different functions (Ainsworth, 2008; Larkin & Simon, 1987). Visualizations can play complementary roles, help learners construct a deeper understanding, and constrain interpretation. For example, compared with a function equation, a function graph offers complementary information about characteristic points, such as extrema, and will help students take in this information through perceptual processes. Visual differentiation helps students gain an intuitive and dynamic understanding of how the slope varies across the entire function, contributing to a comprehensive understanding of the concept of slope in calculus. By visually indicating the possible input and output values of the function, the graph can constrain the interpretation, revealing where the function is not defined. Given the diverse functions of visualizations in mathematics learning, visualizations can help students understand abstract concepts better, make inferences, solve problems, connect mathematical concepts, and aid memory retention (Arcavi, 2003).

1.3. Visualization interventions in mathematics education

We refer to a school, classroom, or learning environment where students participate in activities designed to improve mathematics learning with visualizations as a *visualization intervention*. An example of such a visualization intervention is a study that uses visualizations created by dynamic geometry software GeoGebra compared to a symbolic approach to teach differential equations (Latifi et al., 2021).

Previous reviews have reported positive effects of visualization interventions in science and mathematics education (e.g., Carbonneau et al., 2013; Schroeder et al., 2018; Sokolowski, 2018; Y. Zhang et al., 2023). As one example, Schroeder et al. (2018) reported a moderate effect of $g = 0.60$ for learning with concept maps in STEM domains, with a greater benefit for constructing one's own maps than studying maps that were provided. As other examples, Sokolowski (2018) found a moderate effect size of $g = 0.53$ for learning mathematics with representations at the elementary-school level, and Y. Zhang et al. (2023) reported an effect size of $g = 0.65$ for the effect of learning with GeoGebra technology on mathematics achievement. Positive, though smaller effects were found in meta-analyses of technology-enhanced instruction on mathematics outcomes ($d = 0.15$, A. C. K. Cheung & Slavin, 2013; $d = 0.28$, Li & Ma, 2010; $g = 0.23$, Ran et al., 2022; $g = 0.38$, Young, 2017). Importantly, these meta-analyses focused technology as one type of media and partly included technology besides visualization technology, such as computation enhancement technologies (e.g., calculators, Young, 2017). The Ran et al. (2022) meta-analysis indicates differences in the effectiveness of technology use that would arise from the specificity of the mathematical outcomes investigated, such as communication, problem solving, and conceptual understanding.

One novel contribution of our meta-analysis is that we focus visualization interventions using analog and digital media specific to mathematics learning, as we differentiate between visualization types, mathematical topics (e.g., geometry and algebra), and mathematical outcomes (e.g., understanding and problem-solving) to determine when and for whom including visualizations benefits

mathematics learning.

1.4. Moderators of learning with visualizations

In addition to the overall effects of visualization interventions on mathematics learning, there has been significant variation in how the interventions have been implemented (i.e., design decisions, content of the intervention, sample selection, and the outcomes assessed in the studies). Analyzing the implementation characteristics that influence the effectiveness of visualization interventions will deepen our understanding of who may benefit from specific visualizations when learning particular mathematical topics. We categorized the potential moderating characteristics into three groups: study, intervention, learner, and outcome characteristics.

1.4.1. Study characteristics

Methodological characteristics of a visualization intervention (e.g., the study design and the type of control condition) may influence the effects of the intervention. Employing a quasi-experimental design in which pre-existing groups are assigned to visualization and control conditions can inflate effect sizes compared with experimental designs in which individual learners are randomly assigned to groups (A. C. K. Cheung & Slavin, 2016; Myers et al., 2022). As one example, the mean effect size for quasi-experimental designs was more than twice than for experimental designs in a meta-analysis by A. C. K. Cheung and Slavin (2013) on technology applications for mathematics achievement. A factor such as a teacher's interest in the research topic may favor the visualization group. Additionally, the selection of a control condition, whether it is business-as-usual or an alternative treatment, can impact the intervention's effectiveness. For example, a meta-analysis of mathematics interventions for students with learning difficulties revealed larger effects for the interventions when compared with a business-as-usual control condition (Myers et al., 2022). Explanations include inconsistent implementation and inherent limitations or inefficiencies in business-as-usual practices.

1.4.2. Intervention characteristics

Intervention characteristics include type of visualization, medium used for the visualization, and the mathematical topic taught with visualizations.

According to the integrated model of text and picture comprehension, visualizations can be broadly categorized into two groups (Schnotz, 2002): those that exhibit a physical resemblance to the original (e.g., photographs, schematic images, and augmented reality) and those that demonstrate a more abstract, structural resemblance (e.g., number lines, bar graphs, tree diagrams, and function graphs). The first is associated with the original by similarity, making it particularly suited to transfer spatial relationships. By contrast, the structure of the latter is specified by convention and is often used to represent abstract concepts. The two types of visualizations have distinct uses to serve different purposes, and their effectiveness might depend on the relationship between the specific visualization and the learning demands (Schnotz, 2002).

Both types of visualizations are central to mathematics learning, as identified in a recent scoping review of visualizations in mathematics education research (Schoenherr & Schukajlow, 2023b). Sokolowski's (2018) review of the use of representations in elementary-school mathematics suggested that the effectiveness of visualizations may depend on the type of visualization, with manipulatives showing the largest effect sizes. Given that some types of visualizations were represented by a single primary study in this review, the moderating role of different types of visualizations needs to be investigated further.

Second, visualization requires a medium through which it can be physically manifested and observed from an external perspective (Goldin & Kaput, 1996). Media include analog and digital tools for creating and using visualizations (Presmeg, 1986; Schoenherr & Schukajlow, 2023b). For example, students may use pencil and paper to create drawings when solving real-world problems (Rellensmann et al., 2021), make use of plastic blocks to enhance comprehension of the partial products algorithm (Flores et al., 2020), or utilize dynamic geometry software to understand differential equations (Latifi et al., 2021). While analog materials limit the presentation to static visual displays, technology-based visualizations enable dynamic, animated displays (Schnotz, 2002).

A recent qualitative synthesis of current visualization interventions suggested positive outcomes when technology was used for visualization, in contrast to more mixed results when other analog media were employed (Schoenherr & Schukajlow, 2023b). Explanations include more opportunities to dynamically explore, manipulate, and transform content, which can lead to a deeper understanding of the mathematical topic. A recent meta-analysis of 92 studies revealed a positive impact of the use of technology on learning in mathematics and science compared with learning without technology (Hillmayr et al., 2020). The size of the effect depended on various factors, including the specific tool that was used, with dynamic mathematical tools showing larger effects (Hillmayr et al., 2020). In the domain of science learning, a meta-analysis also showed benefits of learning with dynamic visualizations compared with static visualizations, with gender, education level, and learning domain being important moderators (Castro-Alonso et al., 2019). Thus, previous research in science learning has suggested positive effects of using technology for visualizations in mathematics learning. Importantly, the effectiveness appears to be significantly influenced by the specific method of implementation.

As the integrated model of text and picture comprehension suggests, the effectiveness of visualization interventions for learning depends on the relationships between the visualizations and the specific learning demands (Schnotz, 2002). Consequently, the impact of visualization interventions may differ across various mathematical topics. From a broader perspective, these interventions might demonstrate particular advantages in visually oriented domains, such as geometry, where spatial and visual structures play a significant role (Hershkowitz et al., 1989), in contrast to less visually oriented domains, such as algebra. Taking a more fine-grained view, the benefits of learning with visualizations may vary for specific mathematical topics. For example, meta-analytic findings by Y. Zhang et al. (2023) on GeoGebra visualizations indicated particular advantages for the topics of geometry and calculus. To the best of our knowledge, the effectiveness of learning with visualizations in both visually oriented and less visually oriented mathematical topics or

specific mathematical topics has yet to be investigated.

1.4.3. Learner characteristics

Because visualizations support learning only if they interact appropriately with the individual cognitive system (Schnotz, 2002), different learners may benefit differently from learning with visualizations. For instance, the effects of visualization interventions may depend on learners' level of education (i.e., primary or secondary and later education levels) and their prior knowledge.

Available evidence suggests age-related differences in the ability to benefit from visualizations, such as concept maps and drawing (Brod, 2021). As Brod (2021) concluded for concept mapping, we suggest that most visualizations might be similarly effective in students of different ages, if adjusted appropriately to prior knowledge, working memory, and self-regulation skills considering the complexity of the concept and the support given. This is supported by a meta-analysis on the drawing-to-learn strategy in various domains, including science and humanities, suggesting that there is no straightforward developmental trajectory but rather that a complex interplay of factors influences the effectiveness of visualizations for learners at different levels of education (Cromley et al., 2020).

Additionally, the effectiveness of visualization interventions might depend on learners' prior knowledge, including their prerequisite knowledge and abilities in relation to their peers (Booth & Thomas, 1999; van Garderen et al., 2014). Theoretically, substantial prior knowledge might be a prerequisite for the effective use of visualizations, with learners who possess lower prior knowledge having a greater likelihood of facing challenges and lagging behind their peers. This phenomenon, often termed the *Matthew effect* (Stanovich, 1986), suggests that initial advantages or disadvantages tend to compound over time. For example, previous research has shown that low-achieving students have more difficulty using diagrams than their high-achieving peers, thus impacting the performance of the low achievers (van Garderen et al., 2013, 2014). In a group of students with learning disabilities, those with lower visuo-spatial skills performed worse on arithmetic problems presented with a visualization compared with those with higher visuo-spatial skills (Booth & Thomas, 1999).

As a result, the questions that remain are whether the effectiveness of visualizations varies across learners with different educational levels and with different levels of prior knowledge.

1.4.4. Outcome characteristics

The effectiveness of visualization interventions may depend on the type of outcome measure a study employs.

In the learning of mathematics, the effectiveness of interventions might depend on the particular mathematical outcomes that are measured. For example, in their meta-analysis, Ran et al. (2022) differentiated between the use of technology for various mathematical purposes, finding the highest effect sizes for communication and collaboration, followed by problem-solving and conceptual development. Given the assumption that visualizations support information organization and inference-making (Arcavi, 2003), their advantages may manifest in various mathematical achievements, including understanding a mathematical concept, solving problems, or developing an argumentation or proof (Schoenherr & Schukajlow, 2023b). Additionally, learning with visualizations can enhance students' knowledge and abilities in constructing, using, or interpreting visualizations effectively, thereby providing them with tools to improve their mathematics performance. For example, Lowrie et al. (2019) conducted a 3-week visualization intervention on geometry with 5th and 6th grade students. Compared to the business-as-usual control group, participants in the intervention group showed significant improvements in visual abilities, including spatial visualization and spatial orientation, as well as in their understanding of geometry and word problem solving. The extent to which visualization interventions affect students' mathematics-related visual abilities remains an open question.

Additional variability might be attributed to the alignment between the intervention and the posttest (proximal vs. distal outcome measure), the use of researcher-developed or standardized outcome measures, as well as the timing of the posttest (immediate posttest vs. delayed posttest). We call a proximal outcome measure a measure that aims at assessing knowledge and skills that have been specifically trained during the intervention. On the other hand, a distal outcome measure aims at assessing whether learners can apply the knowledge or skills acquired in the intervention to non-trained contexts and skills. For example, if a student has previously learned to use the rectangle model to show fraction equivalence, a proximal measure would assess the student's skills in showing fraction equivalence and a distal measure might assess the students' ability to employ the same concept to perform fraction addition. While a second-order meta-analysis did not reveal transfer effects on distal outcome measures in cognitive training programs in various domains (Sala et al., 2019), optimism exists regarding visualization interventions specific to mathematics (Gilligan et al., 2020; Sterner et al., 2020). For example, not only did providing spatial scaling training to 250 elementary school students lead to improvements in spatial scaling (i.e., proximal outcome), but it also resulted in effects on distal outcomes such as number line estimations, likely due to enhanced proportional reasoning (Gilligan et al., 2020). In addition, whether studies use standardized or non-standardized tests (i.e., researcher-developed tests) to measure mathematical outcomes could influence the effects and their interpretation: While standardized tests might provide more reliable and comparable results, unstandardized measures might be more tailored to the specific intervention and could allow for more content valid interpretations. In line, a previous meta-analysis of the effects of computer technology on students' mathematical achievement reported larger effects in studies using non-standardized than standardized tests (Li & Ma, 2010). Additionally, the effects of learning with visualizations can be expected to be achieved primarily during the intervention, such that it is highest immediately after the intervention but weakens over time if the acquired knowledge and skills are not applied in between (Sterner et al., 2020).

1.5. The present study

This meta-analytic study aimed to synthesize the existing research on visualization interventions compared with a control group at all levels of mathematics education. The research questions were: (1) What is the overall effect of visualization interventions on mathematics learning? (2) Do intervention characteristics (type of visualization, media, and mathematical topic), learner characteristics (level of education and prior knowledge), or outcome characteristics (mathematical outcomes, proximal and distal outcomes, standardized and non-standardized tests, and timing) moderate the effects of the visualization interventions on mathematics learning?

Given the diverse functions that visualizations can serve in mathematics learning—such as deepening understanding, making inferences, constraining interpretation, and connecting concepts (Ainsworth, 2008; Arcavi, 2003; Larkin & Simon, 1987)—we expected a positive overall effect of mathematics learning with visualizations compared to learning without visualizations. We expected that the overall effect size would fall in a range that is similar to findings reported in previous meta-analyses of specific visualizations in STEM education, visual representations in elementary mathematics, and learning strategy instruction interventions; $g = 0.60$ in Schroeder et al. (2018), $g = 0.53$ in Sokolowski (2018), and $g = 0.50$ in Boer et al. (2018). However, as visualizations often serve very different purposes in mathematics than in other STEM subjects, substantial differences were also expected.

On the basis of methodological considerations regarding intervention characteristics and prior research findings, we expected larger effect sizes in studies using quasi-experimental designs than those using experimental designs (A. C. K. Cheung & Slavin, 2013, 2016; Ran et al., 2022) and in studies in which the intervention was compared with a business-as-usual control condition than with an alternative treatment condition (Myers et al., 2022). Additionally, on the basis of theoretically assumed benefits of dynamic visualizations created with technology (Schnotz, 2002) and previous research findings, we expected that visualizations involving technology would yield larger effect sizes than analog visualizations (Hillmayr et al., 2020; Schoenherr & Schukajlow, 2023a; Y. Zhang et al., 2023). Considering the theoretically assumed distinct purposes of visualizations using physical and structural resemblance (Schnotz, 2002) and the absence of findings on the influence of mathematical topics other than GeoGebra visualizations (Y. Zhang et al., 2023), we did not specify hypotheses about types of visualizations or mathematical topics, respectively. In terms of learner characteristics, we expected that individuals with lower prior knowledge would derive less benefit from visualization interventions than their peers with higher prior knowledge, as the effective use of visualizations requires specific individual prerequisites, including prior knowledge and cognitive skills (Schnotz, 2002), possibly indicating a *Matthew effect* (Stanovich, 1986). This expectation is supported by prior research, indicating that learners with lower knowledge had more trouble using learner-generated visualizations (van Garderen et al., 2013, 2014). Given that different visualizations may be more beneficial at distinct educational levels (e.g., number lines in elementary school and function graphs in higher secondary school), and prior research on learner-generated visualizations did not indicate a straightforward developmental trajectory (Brod, 2021; Cromley et al., 2020), we did not formulate explicit hypotheses about the effects of visualization interventions in primary and secondary schools.

In terms of outcome characteristics, we did not have clear expectations about the influence of the mathematical outcomes, as the theoretically assumed benefits of visualizations in terms of deepened understanding and inference generation might apply to various mathematical outcomes (Arcavi, 2003). Based on prior research findings (Csíkos et al., 2012; Gilligan et al., 2020; Rellensmann et al., 2021), we expected positive effects of visualization interventions on visual abilities. Given that many visualizations can be used effectively in different mathematical topics (e.g., the number line can be used to learn about the order relation on the natural numbers but also to perform fraction division (Sidney et al., 2019)) and learning with visualizations can support memory retention (Arcavi, 2003), we expected positive effects on proximal and distal measures (Gilligan et al., 2020) as well as immediate and delayed testing (Boer et al., 2018; Sterner et al., 2020), with larger effects on proximal measures (Sala et al., 2019) and immediate testing. In addition, we expected larger positive effects from unstandardized tests, which can be tailored more closely to the intervention (Li & Ma, 2010).

2. Method

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement to design this review and report its methods and findings (Page et al., 2021).

2.1. Inclusion criteria

To be included in this review, studies had to meet six criteria based on the Population Intervention Control Outcome (PICO) framework (Higgins et al., 2019). First, studies were eligible if they included participants at any level of mathematics education, ranging from preschool mathematics learning to the university level. Studies with visualization interventions in disciplines other than mathematics (e.g., STEM education or computer programming) or with participants outside the realm of mathematics (e.g., engineering or psychology students) were excluded because these visualization interventions likely have goals besides mathematics learning (e.g., visualizations of stress distributions aimed at understanding how forces and displacements are distributed throughout a bridge in an engineering context), rendering the generalization of effects and effective intervention characteristics difficult. Second, the studies had to focus on mathematics learning sessions that used external visualizations. Examples of interventions that did not meet this criterion included those involving multistep interventions that were not focused on visualizations and those that included other components, such as linguistic prompts. Third, the studies had to employ a (quasi-)experimental design with a control group that did not use or emphasize external visualization. We included studies with a control condition that explicitly excluded external visualizations as well as those with control conditions described as business-as-usual. Fourth, the primary outcome of the studies had to be related to student learning in mathematics. Fifth, studies had to report primary data and provide the data needed to calculate effect

sizes. Sixth, the full text of the study had to be available in English, as English is widely accessible to most researchers. Various types of publications were considered, including grey literature (e.g., dissertations), in order to reduce the risk of publication bias and to identify as much relevant evidence as possible as specified by Higgins et al. (2019). The search was not limited by years of publication.

2.2. Literature search and eligibility screening

In our primary search on June 21, 2023, we searched the interdisciplinary databases Web of Science, Scopus, and Education Resources Information Center. These databases include literature from the social sciences and education. As part of our extended search, we searched the highly ranked mathematics education journals included in the Web of Science Core Collection to ensure that we included relevant reports from the area of mathematics education.

Multiple search iterations using various terms were essential to ensure comprehensive and relevant results, following the methodology recommended by Higgins et al. (2019). Initially, we began by searching for key visualization terms based on Arcavi's (2003) definition, including terms such as "visual," "drawing," "picture," "diagram," "image," and "illustration." However, these keywords generated a substantial number of irrelevant hits, such as those unrelated to the educational context (e.g., "self-image" and "illustration"). As a result, the final search strategy involved four key visualization-related terms combined with terms associated with the educational context, mathematics, and experimental designs, as outlined in Table 1. Searches were restricted to titles, abstracts, or keywords. We did not impose any limitations on the years of publication. To validate our final search strategy, we cross-referenced it with intervention studies identified in a prior scoping review of the most recent literature on external visualizations in mathematics education conducted by Schoenherr and Schukajlow (2023b).

The final primary search yielded 4,250 reports. The flowchart in Fig. 1 outlines the entire process of eligibility screening, adhering to the PRISMA guidelines detailed by Page et al. (2021). Before screening, 57 duplicates were removed with the Citavi software. Titles and abstracts of the remaining 4,193 reports were manually screened by one trained screener, and 3,211 that did not meet the inclusion criteria were removed (e.g., those not related to mathematics education). Exclusions were reviewed by the first author, and disagreements were resolved through consensus by consulting the co-authors. Then, full texts of the 964 accessible reports were independently screened by two raters, and another 926 reports were removed after achieving consensus via discussion. Six reports were potentially eligible but did not report the necessary data. We contacted the first authors of these six reports and requested their primary data, but we did not receive any responses or the data necessary for the analyses. The extended search in highly ranked mathematics education journals included in the Web of Science Core Collection yielded 45 reports, with 23 duplicates identified by the primary search. Of the remaining 22 reports, 19 were excluded, as they did not meet our inclusion criteria. Finally, a total of 43 relevant studies published between 2010 and 2023 that met our inclusion criteria were included in this review. Two additional studies were excluded because they were identified as statistical outliers (see Section 2.6). Therefore, the final analyses included 41 studies (see Fig. 1).

2.3. Coding strategy

In order to systematically synthesize the studies and their findings, two experienced raters coded them with respect to study, intervention, learner, and outcome characteristics. Before the final coding process, both raters applied the deductively developed category system to 20 randomly selected studies. They compared and discussed their ratings, making the necessary refinements to the codes to achieve a mutual understanding of the categories and codes. Then, 20% of included studies were independently double-coded, with mostly high agreement (Cohen's $\varkappa > 0.85$). Agreement was substantial in high-inferential categories, such as visualization type ($\varkappa = 0.64$) and mathematical outcome ($\varkappa = 0.72$). The raters resolved disagreements by discussion and continued coding, maintaining frequent communication and validation with each other. Table S1 in the Supplementary Materials presents the information that was selected from all the reviewed studies.

To assess and discuss the results in light of the *study characteristics*, we coded (1) the study design (i.e., quasi-experimental or experimental) and (2) the type of control condition (i.e., business-as-usual or alternative treatment). Alternative treatments included control groups that received an intervention that explicitly excluded the use of visualizations.

To describe key *intervention characteristics*, we rated the visualizations and the mathematics content addressed in the interventions. We categorized the visualizations into (1) two types of visualizations based on their resemblance to the original: those using physical

Table 1
Search terms for database literature searches.

Set 1: Visualization terms	Set 2: Educational focus	Set 3: Mathematics focus	Set 4: Study design
Visual*	Learn*	Math*	Experiment*
Drawing*	Perform*		Intervent*
Picture*	Achieve*		Control group*
Diagram*	Understand*		
	Problem solv*		
	Proof*		
	Competenc*		
	Comprehen*		

Note. Sets were linked using AND, terms within sets were linked using OR.

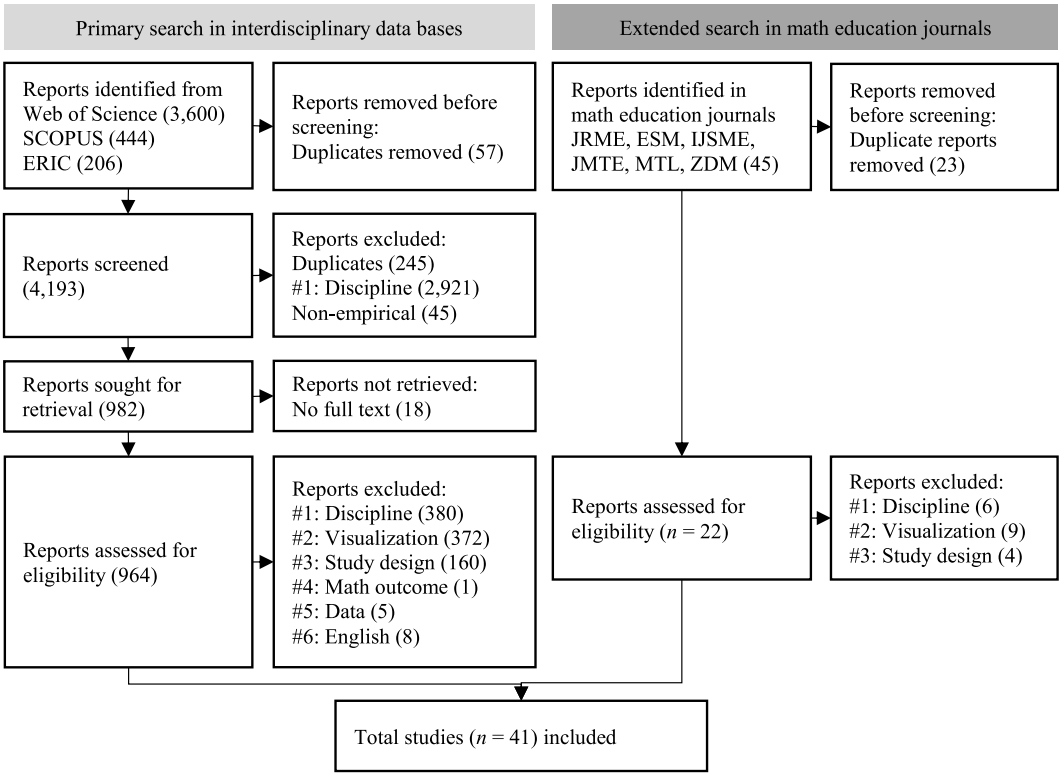


Fig. 1. Eligibility screening Process
Note. JRME = journal for research in mathematics education, ESM = educational studies in mathematics, IJSME = International journal of science and mathematics education, JMTE = journal of mathematics teacher education, MTL = mathematical thinking and learning, ZDM = ZDM – mathematics education.

resemblance (e.g., representational drawing or augmented reality application) and those using structural resemblance (e.g., tree diagram or function graph). If both types of visualizations were used in a single intervention, we assigned the code "mix." Furthermore, we categorized (2) the type of medium used into the following codes: analog or digital, which represented technology use. The mathematics content was assessed by coding (3) the mathematical topic addressed by the intervention. Using a fine-grained approach, the codes included numbers and operations, analysis, algebra, geometry, probability and data analysis, logic and set theory, or a mixture, which represented interventions covering two or more different topics. In a summarized approach, we contrasted the visually oriented topic of geometry with the other less visually oriented topics.

For *learner characteristics*, we coded (1) the learners' level of education as primary (or earlier) school level or secondary (or later) school level, and (2) their prior knowledge, categorized as low, average, or high in comparison with their peer group. Learners with learning disabilities were categorized as having low prior knowledge, whereas above-average or gifted students were categorized as having high prior knowledge compared to their peer group. Unless specified otherwise by the authors of the primary studies, all other learners were considered to have average prior knowledge.

As *outcome characteristics*, we coded (1) the type of mathematical outcome, including visual abilities, understanding or reasoning, and problem solving or problem posing. Visual abilities included visualization-related knowledge and skills in constructing and using visualizations. In addition, we coded (2) the relationship between the intervention and the posttest as either "proximal outcome measure" or "distal outcome measure." We considered the outcome to be proximal when the posttest assessed competencies and content covered in the intervention. Conversely, we categorized it as distal when the posttest referred to competencies or content beyond the scope of the intervention, such as using grades as a post-intervention measure. Moreover, we coded whether (3) standardized or non-standardized instruments were used for the posttest and (4) the timing of the posttest as either immediate or delayed.

2.4. Extraction of effect sizes

The unbiased standardized mean difference (Hedges' *g*) was used as the dependent variable (Hedges, 1981). Effect sizes were obtained by following the methods described in Borenstein et al. (2009) and Lipsey and Wilson (2000).

Effect sizes were calculated directly from the descriptive data reported in the studies. If no descriptive data were available, reported effect sizes were used directly and, if necessary, transformed into Hedges' *g* in accordance with the procedures reported in Borenstein et al. (2009) and Lipsey and Wilson (2000). The error variances of the effect sizes were estimated in proportion to the sample sizes

(Borenstein et al., 2009).

2.5. Analysis

To find the best-fitting model, we compared three multilevel meta-analytic models (see Fig. 2) using a stepwise likelihood ratio test. Model 1 was a standard random-effects model with sampling variance (extracted from the studies) at Level 1 and the random variance of effect sizes at Level 2. For Model 2, we included an additional level (Level 4) to account for the nesting of effect sizes within studies. Model 3 included an intermediate level (Level 3) that accounted for the nesting of effect sizes within subgroups, which were in turn nested within studies.

There are several approaches to tackle the issue of nested effect sizes (for an overview see Borenstein et al., 2009; M. W.-L. Cheung, 2014). M. W.-L. Cheung (2014) argues that three-level meta-analyses are the preferred way to handle dependent effect sizes compared to, for example, robust variance estimation (Hedges et al., 2010). Based on comparative analyses, M. W.-L. Cheung (2014) shows that three-level models perform best in estimating confidence intervals, variances and covariances. We used a multilevel approach instead of the structural equation modeling approach suggested by M. W.-L. Cheung (2014) because Model 3 of our analyses included a fourth level, which is not yet readily implemented in structural equation modeling tools.

To test the effects of moderators, we computed separate meta-regressions with the moderators as fixed effects. All models were fitted using the *R* package *metafor* (Version 4.4) (Viechtbauer, 2010). Dependencies in sampling errors due to the inclusion of several effect sizes from the same population were accounted for by an approximation of the variance-covariance matrix as described in Viechtbauer (2010) with an estimated correlation between concurrent measurements of $\rho = 0.7$ (sensitivity analyses for ρ indicated a minor influence on the overall effect, ranging from $g = 0.49$ for $\rho = 1$ to $g = 0.53$ for $\rho = 0$). Confidence intervals for all estimates were calculated on the basis of a *t*-distribution, similar to the Knapp and Hartung adjustment (Knapp & Hartung, 2003; Viechtbauer, 2010). I^2 (representing the proportion of variation due to heterogeneity) was calculated in accordance with the method described in Viechtbauer (2022); see also M. W.-L. Cheung (2014) and Higgins and Thompson (2002). R^2 values were calculated as described in M. W.-L. Cheung (2014). Tests for residual heterogeneity and omnibus tests of moderators were based on *F*-distributions (Viechtbauer, 2010).

2.6. Outlier identification

For outcomes, outliers were identified through sensitivity analyses for the total effect and the amount of heterogeneity, and by Cook's distance (Viechtbauer & Cheung, 2010). This was done on the study level and for individual effects. This process led to the exclusion of two full studies that had a high leverage, contributed excessively to the heterogeneity, and increased the total effect size by $\Delta g = 0.065$ and $\Delta g = 0.047$, respectively (see Figure A.1 in the Appendix). With these outliers included, the overall mean effect would increase to $g = 0.615$, $SE = 0.083$. The findings from the outlier studies (Tan, 2012; Uyen et al., 2021) should be interpreted with caution, as they may deviate from the overall effectiveness of visualization interventions due to chance or highlight particularly effective interventions, warranting further investigation in future research.

3. Results

3.1. Descriptive results

Our analysis included 41 studies, consisting of 65 subsamples with a total of $N = 10{,}562$ learners. We extracted 153 effect sizes from these studies. Studies were published in 35 journal articles, five conference proceedings, and one dissertation. An overview of all studies, coded moderators, and the corresponding effect sizes is provided as supplementary material.

3.2. Publication bias

According to Rodgers and Pustejovsky (2021), standard methods that are used to test for publication bias (including analyses of funnel plot asymmetry, trim-and-fill analyses, or Egger's regression; see Borenstein et al., 2009) do not provide reliable results for multilevel meta-analyses. They show that a multilevel (*MLMA)* Egger's regression is an appropriate approach in these situations. These analyses did not show a significant relation between effect sizes and standard errors ($\beta = 0.64$, F (1,151) $= 1.143$, $p = .287$). An alternative approach is to compare the mean effect size of non-reviewed *grey literature* against peer-reviewed publications (Borenstein

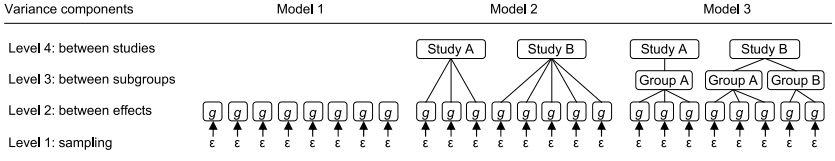


Fig. 2. Structure of the tested models.

et al., 2009). In our sample, 139 effects were extracted from peer-reviewed journal articles ($g = 0.588$), and 14 were reported in book chapters or conference proceedings ($g = 0.492$). The difference was not significant ($p = .605$) and not in the direction that would indicate publication bias.

3.3. Overall effect size

Table 2 presents the model fit comparison of the three random-effects models. We found that Model 2 with a hierarchical three-level structure fit the data better than Model 1, whereas Model 3 did not fit significantly better than Model 2. Thus, we chose the random-effect structure of Model 2 for all analyses. For this model, the mean effect of visualization interventions on outcomes in mathematics was $g = 0.504$, $SE = 0.064$, 95% CI [0.379, 0.630], $p < .001$, see Table 3. The prediction interval, [−0.45, 1.46], and Q-test, $Q\,(152) = 1522$, $p < .001$ indicated substantial heterogeneity across the studies. Most of the residual variance was explained by a moderate to high degree of heterogeneity on the between-effects level ($I^2_{(2)} = 66.3\%$), but I^2 was low on the between-studies level ($I^2_{(4)} = 23.6\%$) (Higgins et al., 2003).

3.4. Meta-regression

For each set of moderators, separate meta-regression analyses were conducted using the hierarchical three-level random-effects structure from Model 2 (as described in Section 3.2), with the moderators added as fixed effects.

3.4.1. Study characteristics

Table 4 presents the meta-regression results for the study characteristics. In line with our expectations, studies with quasi-experimental designs showed significantly higher effect sizes ($g = 0.597$) than studies with randomized designs ($g = 0.290$, $p = .016$), explaining 27% of the residual variance on the between-study level. Also as expected, effect sizes were significantly different between studies that used control groups with no visualizations ($g = 0.23$) and studies that used business-as-usual control groups ($g = 0.555$), $F\,(1, 151) = 4.523$, $p = .035$, $R^2_{(2)} = 0.27$.

3.4.2. Intervention characteristics

Table 5 presents the meta-regression results for the intervention characteristics. Contrary to our expectations, the type of medium was not a significant moderator, F (2, 146) = 0.379, $p = .539$. The type of visualization, F (2, 148) = 1.773, $p = .173$, and the mathematical topic of the intervention, $F\,(1, 151) = 0.057$, $p = .812$, were also no significant moderators.

3.4.3. Learner characteristics

Table 6 presents the meta-regression results for learner characteristics. Contrary to our expectations, prior knowledge was not a significant moderator of effect size, $F\,(1,151) = 1.409$, $p = .237$. The sample included only 17 effects from groups that reported low prior knowledge and only three effects from groups that reported high prior knowledge. Also, there were no significant differences in effect sizes between studies that differed in participants' level of education, $F\,(1,151) = 0.847$, $p = .359$.

3.4.4. Outcome characteristics

Table 7 presents the meta-regression results for outcome characteristics. As expected, effect sizes that were based on proximal outcome measures were significantly larger ($g = 0.557$) than those based on distal outcome measures ($g = 0.366$), $F\,(1, 151) = 4.352$, $p = .039$, explaining 19% of the residual variance on the between-study level. There was a trend that effects obtained with standardized instruments were smaller than with unstandardized instruments, but this difference was not significant in our data, $F\,(1, 151) = 1.120$, $p = .292$. Contrary to our expectations, timing did not moderate the effect sizes, $F\,(1, 151) = 0.244$, $p = .409$. Also, the type of outcome measure was not a significant moderator of the effect sizes, $F\,(3, 149) = 0.145$, $p = .933$, and $F\,(2, 150) = 0.060$, $p = .942$, respectively.

3.5. Effects for strongest experimental designs

As an additional analysis, we analyzed the average effect size for studies that included the strongest experimental design with active control groups. This only applied to four studies (903 participants, 16 subgroups, 41 effects) with a mean effect size of $g = 0.16$, $SE =$

Table 2
Likelihood ratio tests for the three tested models.

	Variance (τ^2)						
	L2	L3	L4	Estimate (SE)	AIC	BIC	LRT
Model 1	0.195			0.44 (0.048)	241.00	247.05	
Model 2	0.168		0.060	0.50 (0.064)	235.98	245.05	7.02[a]
Model 3	0.168	0	0.060	0.50 (0.064)	237.98	250.08	0

Note. L2 = between effects, L3 = between subgroups, L4 = between studies. AIC = Akaike Information Criterion, BIC = Bayesian Information Criterion, LRT = Likelihood ratio test. The LRT refers to the corresponding previous model.
[a] $p < .01$.

Table 3
Results for the overall effect.

Effect	*n*	Estimate	*SE*	95% CI		*p*	I^2
				LL	*UL*		
No fixed effects							
Intercept	153	**0.504**	**0.064**	**0.379**	**0.630**	**<0.001**	
Random Effects							
Between-study variance		0.060		0.011	0.159		23.64
Between-effect variance		0.169		0.125	0.230		66.33

Table 4
Results of meta-regressions for study characteristics.

Effect	*n*	Estimate	*SE*	95% CI		*p*	I^2	R^2
				LL	*UL*			
Study design								
Fixed Effect								
RC: Experimental design	88	**0.290**	**0.103**	**0.086**	**0.493**	**0.006**		
Quasi-experimental design	65	**0.307**	**0.126**	**0.058**	**0.557**	**0.016**		
Omnibus test *F* (1, 151)		**5.920**				**0.016**		
Random Effects								
Between-study variance		0.044		0	0.134		18.22	26.90
Between-effect variance		0.168		0.124	0.229		69.52	0.58
Control Group								
Fixed Effects								
RC: Alternative treatment	46	0.230	0.138	−0.042	0.502	0.097		
Business-as-usual	107	**0.325**	**0.153**	**0.023**	**0.627**	**0.035**		
Omnibus test *F* (1, 151)		**4.523**				**0.035**		
Random Effects								
Between-study variance		0.044		0	0.139		18.26	26.91
Between-effect variance		0.170		0.126	0.233		70.60	0

Note. Number of studies = 41, number of subgroups = 65, number of effects = 153, total N = 10,562. CI = confidence interval; LL = lower limit; UL = upper limit, RC = reference category, *p*-values are two-tailed.

0.071, 95% CI [0.019, 0.307], $p = .027$. Due to the small number of studies, no moderator analyses were conducted for this subsample.

4. Discussion

4.1. Overall effect of visualization interventions

In this meta-analysis, we synthesized empirical findings on the effectiveness of visualization interventions on students' mathematics outcomes across 41 intervention studies conducted between 2010 and 2023, with 153 effect sizes.

Regarding our first research question, we found a medium overall effect of $g = 0.504$ on mathematics outcomes. This estimate suggests that, compared with students who did not receive a visualization intervention, the mathematics outcomes of those who did receive a visualization intervention improved by about 0.5 more standard deviation units. The effect size is similar in direction and magnitude to effect sizes found in previous meta-analyses of visualizations in STEM education, $g = 0.60$ in Schroeder et al. (2018); visual representations in elementary mathematics, $g = 0.53$ in Sokolowski (2018); and learning strategy instruction interventions, $g = 0.50$ in Boer et al. (2018). Due to the small number of studies employed in a prior meta-analysis on learning with representations in elementary school (i.e., 13) (Sokolowski, 2018), which limited the reliability of the estimates (Hedges & Pigott, 2004), the current meta-analysis, which included 41 studies, improved the reliability of effect size estimates and broadened the scope to include learners from various age groups. Taken together, the current and previous meta-analyses provide robust empirical support for the effectiveness of learning mathematics with visualizations across different types of visualizations, learner populations, and mathematical content, provided they are implemented appropriately according to the learning target and the learning group.

In the current meta-analysis, we were able to build on a large sample of 153 effects from various interventions that included multiple measures, such as content-related performance tests and overall grades, likely introducing sources of variation. Expectedly, we found high heterogeneity in the effect sizes. In order to explain the sources of heterogeneity, we used meta-regression models to conduct moderator analyses so that we could determine whether study, intervention, learner, and outcome characteristics were associated with the variations in the observed effects. The moderator analyses, based on theories of visualization and representation (Ainsworth, 2008; Larkin & Simon, 1987; Schnotz, 2002), while considering the specificity of mathematics learning, are a central contribution of this meta-analysis. These analyses indicate which types of visualization are effective in supporting various

Table 5
Results of meta-regressions for intervention characteristics.

Effect	*n*	Estimate	*SE*	95% CI		*p*	I^2	R^2
				LL	*UL*			
Visualization Type								
Fixed Effects								
RC: Both	8	0.279	0.211	−0.137	0.696	0.187		
Physical resemblance	49	0.163	0.213	−0.259	0.584	0.447		
Structural resemblance	45	0.394	0.242	−0.084	0.873	0.106		
Omnibus test *F* (2, 148)		1.773				0.173		
Random Effects								
Between-study variance		0.063		0.012	0.165		24.54	0
Between-effect variance		0.164		0.122	0.225		63.72	2.41
Media Type								
Fixed Effects								
RC: Analog	88	**0.463**	**0.095**	**0.276**	**0.650**	**<0.001**		
Digital	65	0.078	0.126	−0.172	0.327	0.539		
Omnibus test *F* (1, 151)		0.379				0.539		
Random Effects								
Between-study variance		0.063		0.011	0.166		24.12	0
Between-effect variance		0.168		0.125	0.230		64.83	0
Mathematical Topic (Fine Grained)								
Fixed Effects								
RC: Geometry	55	**0.419**	**0.097**	**0.229**	**0.610**	**<0.001**		
Numbers and Operations	70	0.007	0.143	−0.275	0.289	0.961		
Analysis	14	0.396	0.202	−0.004	0.796	0.052		
Algebra	8	0.139	0.213	−0.282	0.561	0.515		
Stochastic and data analysis	1	−0.581	0.097	−1.515	0.353	0.221		
Logic and set theory	2	0.171	0.466	−0.750	1.091	0.715		
Linear algebra	3	0.542	0.311	−0.072	1.157	0.083		
Omnibus test *F* (6, 146)		1.596				0.152		
Random Effects								
Between-study variance		0.050		0.004	0.145		19.70	17.53
Between-effect variance		0.167		0.124	0.228		66.42	0.90
Mathematical Topic (Summarized)								
Fixed Effects								
RC: Less visually-oriented	95	**0.519**	**0.084**	**0.352**	**0.685**	**<0.001**		
Visually oriented	58	−0.030	0.127	−0.281	0.220	0.812		
Omnibus test *F* (1, 151)		0.057				0.812		
Random Effects								
Between-study variance		0.063		0.011	0.166		24.11	0
Between-effect variance		0.169		0.125	0.231		64.84	0

Note. Number of studies = 41, number of subgroups = 65, number of effects = 153, total N = 10,562. CI = confidence interval; LL = lower limit; UL = upper limit, RC = reference category, *p*-values are two-tailed.

mathematical outcomes across different mathematical topics.

4.2. The moderating effects of study characteristics

The effect size was slightly lower in experimental studies using randomized designs ($g = 0.290$) than in quasi-experimental studies ($g = 0.597$), suggesting a tendency toward inflated effect sizes in the latter (A. C. K. Cheung & Slavin, 2013, 2016; Myers et al., 2022). Possible explanations include selective factors favoring the visualization groups, such as teachers' openness to innovation. The inflated mean effect sizes in quasi-experimental studies raise concerns, as large proportions of studies on visualizations in mathematics learning use quasi-experimental designs. For example, 42% of the effect sizes included in this meta-analysis came from studies that employed quasi-experimental designs. Thus, in order to obtain more reliable estimates of the true effect size in future visualization studies, researchers should use experimental designs in which students are randomly assigned to the visualization and control conditions.

In addition, the effect sizes were larger in studies that compared visualization interventions with a business-as-usual control condition ($g = 0.555$) than with an alternative treatment that did not include visualizations ($g = 0.230$). This finding is in line with a previous meta-analysis of mathematics interventions (Myers et al., 2022). One explanation is that additional aspects besides visualization use may be altered in business-as-usual control conditions compared to visualization conditions. Comparing visualization interventions with alternative treatments (e.g., a purposefully designed control condition that does not include visualizations) may thus provide a more focused estimate of the effect of learning with visualizations.

Table 6
Results of meta-regressions for learner characteristics.

Effect	*n*	Estimate	*SE*	95% CI		*p*	I^2	R^2
				LL	*UL*			
Level of Education								
Fixed Effects								
RC: Primary and earlier	85	**0.444**	**0.094**	**0.258**	**0.63**	**<0.001**		
Secondary and later	68	0.119	0.129	−0.136	0.374	0.359		
Omnibus test *F* (1, 151)		0.847				0.359		
Random Effects								
Between-study variance		0.065		0.014	0.167		25.05	0
Between-effect variance		0.167		0.123	0.228		64.17	1.16
Prior knowledge								
Fixed Effects								
Intercept	133	**0.690**	**0.166**	**0.363**	**1.017**	**<0.001**		
Prior Knowledge[a]	20	−0.189	0.159	−0.502	0.125	0.237		
Omnibus test *F* (1, 151)		1.409				0.237		
Random Effects								
Between-study variance		0.069		0.016	0.175		26.64	0
Between-effect variance		0.166		0.123	0.227		63.60	1.61

Note. Number of studies = 41, number of subgroups = 65, number of effects = 153, total N = 10,562. CI = confidence interval; LL = lower limit; UL = upper limit, RC = reference category, *p*-value two-tailed.

[a] Prior knowledge was coded low (−1), average (0), high (1). 17 effects of low prior knowledge groups and 3 effects for high prior knowledge groups were included.

4.3. The moderating effect of intervention characteristics

In the current study, different types of visualizations emerged as powerful tools to support mathematics learning in various domains. No difference in effectiveness was found between the two types of visualizations, which relied on physical and structural resemblance (Schnotz, 2002), respectively. Thus, the effectiveness of learning with visualizations does not depend on whether the visualization has physical or structural similarity to the original. By means of the visualization intervention, the students learn how the visualization encodes and presents information (Ainsworth, 2008), thus enabling them to effectively interpret, create, and use visualizations on the basis of the different representation principles. The two types of visualizations can be further differentiated into mathematics-specific visualization subtypes, such as pictorial and schematic visualizations or length, area, and relational visualizations (Schoenherr & Schukajlow, 2023b). Open questions that remain for further research to address are whether the different subtypes have different effectiveness or place different demands on an effective visualization intervention.

Contrary to our expectations, this meta-analysis did not provide support for the hypothesis that using technology to create and use visualizations is more effective than other types of media. Our findings suggest that digital and analog visualizations can be equally effective in supporting mathematics learning when their use is aligned with the learning goal. The digital visualization interventions were based on various technological tools, including programming platforms (Amevor et al., 2021), game-based learning platforms (Ke, 2019), graphing calculators (Johnson et al., 2019), dynamic geometry software (Zulnaidi et al., 2020), and augmented and virtual reality applications (Demitriadou et al., 2020). Some studies included in this meta-analysis, which contrasted different types of media, provided mixed findings. For example, Borba et al. (2015) did not find differences in elementary-school students' combinatorial problem-solving when students used paper-pencil versus technology to construct tree diagrams. As another example, Baki et al. (2011) examined virtual and physical manipulatives for preservice teachers who were learning geometry and found benefits of virtual manipulatives in some components of spatial reasoning. This finding underscores that the conditions for effective technology-supported visualizations in mathematics are complex. The Hillmayr et al. (2020) meta-analysis on mathematics and science learning with and without the use of technology additionally identified important moderators of effective technology use, including types of technological tools, such that intelligent tutoring systems and simulations were more beneficial, as they provided teachers with technology training and complemented but did not substitute for other instruction methods. Aside from the question of whether technology works (better), future research is needed to examine how different types of technological tools change the nature of how students learn mathematics and how different types of media can be combined to create effective teaching practices (Drijvers & Sinclair, 2024).

In the current meta-analysis, no differences in effectiveness in different mathematical topics emerged. Thus, visualizations offer an effective way to help students learn about geometry, a domain that relies heavily on visual reasoning (Hershkowitz et al., 1989), but they can also be effective in less visually oriented domains involving abstract concepts, such as numbers and operations, calculus, and probability (Arcavi, 2003). One explanation is that visual encoding of structural characteristics of mathematical concepts in visualizations with structural resemblance (Schnotz, 2002) facilitates comprehension through perception (Larkin & Simon, 1987). As one example, teaching fractions with two-dimensional number lines that combine the unidimensional number line with a bar diagram can help students focus on the sections of the number line rather than the tick marks, thereby enhancing their fraction knowledge (Soni & Okamoto, 2020). Consequently, the inclusion of visualizations can contribute to a deeper conceptual understanding of various mathematical topics, even those that are not inherently visual (Ainsworth, 2008).

Table 7
Results of meta-regressions for outcome characteristics.

Effect	*n*	Estimate	*SE*	95% CI		*p*	I^2	R^2
				LL	*UL*			
Type of mathematical outcome								
Fixed Effects								
RC: Visual abilities	47	**0.478**	**0.105**	**0.271**	**0.685**	**<0.001**		
Understanding and reasoning	54	0.066	0.133	−0.196	0.329	0.618		
Problem solving and problem posing	52	−0.009	0.116	−0.239	0.220	0.937		
Omnibus test F (2, 150)		0.218				0.804		
Random Effects								
Between-study variance		0.059		0.008	0.161		22.48	1.38
Between-effect variance		0.171		0.127	0.234		65.01	0
Proximal and distal outcome measures								
Fixed Effects								
RC: Proximal measure	101	**0.557**	**0.067**	**0.424**	**0.690**	**<0.001**		
Distal measure	52	**−0.191**	**0.092**	**−0.373**	**−0.010**	**0.039**		
Omnibus test *F* (1, 151)		**4.352**				**0.039**		
Random Effects								
Between-study variance		0.049		0.003	0.141		19.82	18.75
Between-effect variance		0.168		0.124	0.230		68.14	0.44
Standardized and non-standardized test								
Fixed Effects								
RC: Non-standardized	103	**0.538**	**0.071**	**0.397**	**0.679**	**<0.001**		
Standardized	50	−0.130	0.123	−0.372	0.113	0.292		
Omnibus test *F* (1, 151)		1.120				0.292		
Random Effects								
Between-study variance		0.058		0.009	0.156		22.62	4.10
Between-effect variance		0.169		0.125	0.231		65.51	0
Immediate and delayed testing								
Fixed Effects								
RC: Immediate posttest	140	**0.501**	**0.064**	**0.375**	**0.627**	**<0.001**		
Delayed posttest	13	0.076	0.155	−0.229	0.382	0.622		
Omnibus test F (1, 151)		0.244				0.622		
Random Effects								
Between-study variance		0.060		0.010	0.159		23.13	0
Between-effect variance		0.170		0.126	0.232		65.75	0

Note. Number of studies = 41, number of subgroups = 65, number of effects = 153, total N = 10,562. CI = confidence interval; LL = lower limit; UL = upper limit, RC = reference category, *p*-value two-tailed.

4.4. The moderating effect of learner characteristics

The current meta-analysis showed that visualization interventions supported students' mathematics learning at different educational levels, ranging from elementary school (and earlier) to secondary school (and later). This finding is in line with a previous meta-analysis on the drawing-to-learn strategy suggesting that there is no set developmental trajectory, but rather a complex interplay of factors that influence the effectiveness of visualizations for learners at different educational levels (Cromley et al., 2020). For example, hands-on objects might be particularly effective in primary mathematics education (Sokolowski, 2018). Additional research is essential for enhancing our understanding of how different types of visualizations and media interact with the content that students need to learn across primary, secondary, and tertiary education levels. Notably, 88% of the effect sizes identified in this meta-analysis were related to visualization interventions in elementary and secondary education. This finding underscores the need for more research on how the learning of mathematics with visualizations can be effectively implemented in other important phases of mathematics education.

In the present meta-analysis, visualization interventions were found to be beneficial for learners with low prior knowledge and those with average prior knowledge, indicating neither a Matthew effect nor an expertise reversal effect. Although no significant differences in effect size estimates were observed among learners with varying levels of prior knowledge, we did find evidence that students' prerequisites should be taken into account when designing visualization interventions. For example, elementary-school students with mathematics difficulties showed improved learning outcomes with a visual-only strategy training, whereas students without mathematics difficulties performed better in the verbal-and-visual intervention condition than in the control condition (Swanson et al., 2015). Future research should consider individual learning characteristics, including prior knowledge, to gain a better understanding of which visualization interventions are most effective for specific learners. Crucially, additional research is needed to examine the effectiveness of visualization interventions for learners with high prior knowledge. The studies included in this meta-analysis did not specifically investigate the effectiveness of visualization interventions for this particular population.

4.5. The moderating effect of outcome characteristics

Visualization interventions have been shown to support various mathematics outcomes, including understanding new topics, problem solving, and problem posing. These findings underscore the universal applicability of visualizations in mathematics learning (Arcavi, 2003; Presmeg, 1986). Importantly, visualization interventions have been shown to support visual abilities, including spatial reasoning, spatial orientation, and mental rotation (Ke, 2019; Lowrie et al., 2017, 2019), as well as strategic knowledge about visualizations (Rellensmann et al., 2021), which can contribute to long-term mathematics learning.

In addition, our findings indicate that the visualization interventions effectively enhanced targeted knowledge and abilities. Such enhancement is evident in the positive effects of visualization interventions on proximal outcomes ($g = 0.557$), which was significantly stronger than for distal outcomes ($g = 0.366$). To some extent, this might indicate a *teaching to the test* effect for studies which closely tailor their outcome measures to their intervention, or vice versa. However, in line with findings from Gilligan et al. (2020) for spatial training in elementary school, positive effects, although decreasing, were still substantial for distal outcome measures. This indicates that the benefits of visualization interventions generalize beyond the initially targeted abilities and knowledge. These findings are complemented by the finding that the effectiveness of visualization interventions did not vary significantly for standardized and non-standardized posttests, although effect sizes were larger for non-standardized tests ($g = 0.538$) than standardized tests ($g = 0.408$) on a descriptive level as in the meta-analysis by Li and Ma (2010). This indicates that even though learning with visualizations might be particularly helpful for the learning of the targeted outcome, transfer to more distal outcomes is possible.

Importantly, the positive effects of visualization interventions are not limited to improvements in comprehension and performance on immediate posttests; they also support long-term learning, as evidenced by positive effects on delayed posttests. This finding is consistent with a previous meta-analysis on learning strategy instruction, which showed that the effects of strategy instruction are maintained beyond the intervention period and might even increase slightly over time, indicating that skills continue to develop after the intervention ends (Boer et al., 2018). Taken together, the findings from the current meta-analysis suggest that visualization interventions sustainably enhance mathematics learning beyond intervention materials and the time period of the intervention.

4.6. Limitations

Despite our implementation of extensive search strategies and a standardized study selection procedure in high-ranked databases, it is possible that we may have failed to include relevant research, for example, studies that indirectly addressed visualizations in mathematics education or were rooted in alternative theoretical frameworks (e.g., manipulatives and schema-based instruction).

Another limitation concerns the coding strategy. To extract relevant study, intervention, learner, and outcome characteristics, we relied on the information the authors provided. Inconsistent or unclear use of terms as well as incomplete descriptions in the primary studies may have impacted our findings.

In this meta-analysis, we addressed central moderators identified in previous research as sources of the large heterogeneity in effect sizes. One source of heterogeneity is the inclusion of studies that varied in the test instruments that were used to measure the effects of visualization interventions in mathematics education, as only a few studies employed standardized instruments. The included studies varied in characteristics, such as instructional methods and experimental procedures, including aspects such as teacher training. Coding these variables reliably across studies proved challenging, primarily because the information the studies provided was insufficient. Moreover, there are other characteristics based on theories of learning and instruction whose analysis can provide insights into how visualization interventions can be effectively implemented. For example, a constructivist approach might be more effective than a traditional approach to mathematics learning, as found for technology-enhanced visualizations (Li & Ma, 2010). Thus, additional studies examining moderators of the effects of visualization interventions are needed to evaluate additional sources of heterogeneity, including their implementation based on theories of learning and instruction.

The aim of this meta-analysis was to synthesize the effects of visualization interventions in mathematics education. Given the role that mathematical visualizations play in other subjects, such as x-y function graphs for secondary school students in physics, it is a promising direction for future reviews on mathematical visualization interventions to broaden their scope. Specifically, they might investigate the effectiveness of mathematical visualization interventions in different subjects by including the moderator "mathematics vs. other subjects." As an additional line of future research, subsequent reviews might benefit from narrowing their focus to specific visualizations in order to provide deeper insights into the effective design of such interventions.

4.7. Implications for practice

Our research findings highlight the significant positive impact of using diverse visualizations, such as such as graphs, diagrams, and number lines, tailored to specific educational contexts, on learning outcomes in mathematics.

The effectiveness across different mathematical topics suggests that visualizations can be widely applied. Teachers should encourage the use of visualizations not only in geometry and algebra but also in areas such as calculus, statistics, and modeling problems to facilitate learning. By incorporating a variety of visualization exercises and strategies, educators can enhance students' visual abilities and strategic knowledge, contributing to both short-term and long-term learning in mathematics.

We recommend that educators acknowledge the potential benefits of integrating visualizations into their regular instructional practices, thereby maximizing mathematical learning outcomes. Since effects were greatest compared to business-as-usual lessons, teachers might benefit from professional development opportunities focused on best practices for creating and using visualizations effectively in their everyday teaching. Workshops and training sessions can help educators design effective visualizations and integrate

them seamlessly into their curriculum.

Educators have the opportunity to enhance students' visual abilities and strategic knowledge by incorporating a variety of visualization exercises and strategies, contributing to both short-term and long-term learning in mathematics. Importantly, after considering the unique needs of learners and the specific mathematical topic at hand, educators can choose from a diverse range of types of visualizations. This flexibility extends to the selection of the medium, allowing educators to adapt their teaching based on available options and learning goals, ensuring alignment with desired outcomes. Interactive software (e.g., Baki et al., 2011), augmented and virtual reality (e.g., Demitriadou et al., 2020), and manipulatives (e.g., Baki et al., 2011) can provide dynamic and engaging ways for students to interact with mathematical concepts through visualizations. However, paper-pencil visualizations, which require minimal resources, may be equally effective in supporting mathematical learning (e.g., Csíkos et al., 2012). It is worth noting that this flexibility in choosing the type of visualization and medium does not significantly impact the overall effectiveness of the visualization intervention.

Author statement

Johanna Schoenherr: Conceptualization; Investigation; Data curation; Writing - original draft; Writing - review & editing. Anselm R. Strohmaier: Methodology, Formal analysis, Data curation, Writing - original draft; Writing - review & editing. Stanislaw Schukajlow: Conceptualization; Writing - review and editing.

Declaration of competing interest

We have no conflicts of interest to disclose.

Data availability

We have uploaded the two data files and the syntax to https://osf.io/7ctay.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.edurev.2024.100639.

Appendix

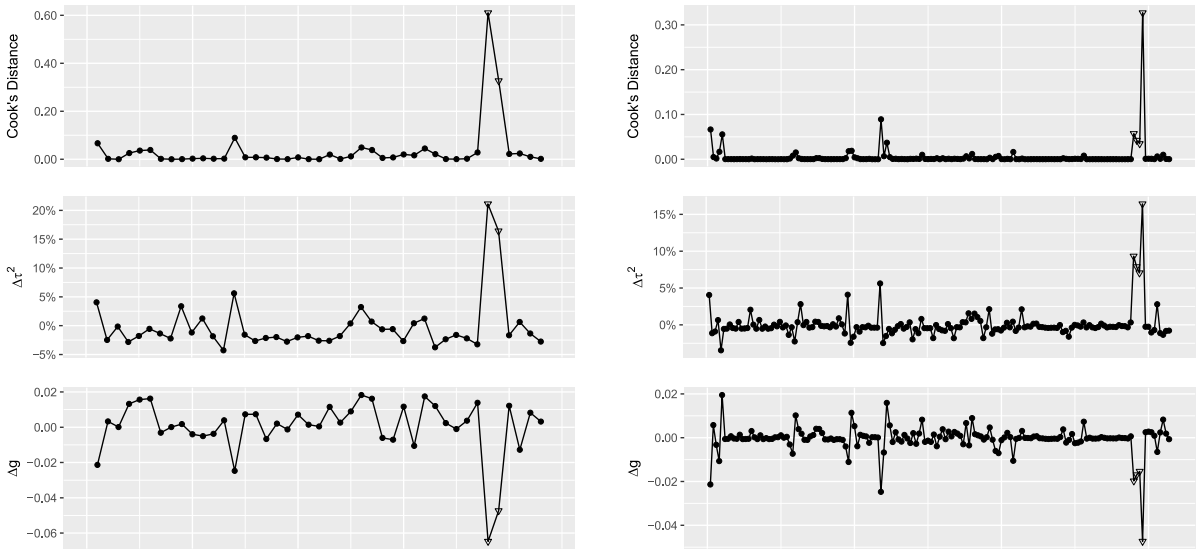


Fig. A.1. Outlier analyses on the level of studies (left) and effects (right). Excluded effects are marked as triangles.

References

Ainsworth, S. (2008). The educational value of multiple-representations when learning complex scientific concepts. In J. K. Gilbert, M. Reiner, & M. Nakhleh (Eds.), *Visualization: Theory and practice in science education* (pp. 191–208). Springer.

Amevor, G., Bayaga, A., & Bosse, M. J. (2021). Analysis of rural-based pre-service teachers spatial-visualisation skills in problem solving in vector calculus using MATLAB. *International Journal of Emerging Technologies in Learning, 16*(10), 125–149. https://doi.org/10.3991/ijet.v16i10.19269

Arcavi, A. (2003). The role of visual representations in the learning of mathematics. *Educational Studies in Mathematics, 52*(3), 215–241. https://doi.org/10.1023/A:1024312321077

Baki, A., Kosa, T., & Guven, B. (2011). A comparative study of the effects of using dynamic geometry software and physical manipulatives on the spatial visualisation skills of pre-service mathematics teachers. *British Journal of Educational Technology, 42*(2), 291–310. https://doi.org/10.1111/j.1467-8535.2009.01012.x

Binder, K., Krauss, S., & Bruckmaier, G. (2015). Effects of visualizing statistical information - an empirical study on tree diagrams and 2 × 2 tables. *Frontiers in Psychology, 6*, 1186. https://doi.org/10.3389/fpsyg.2015.01186

Boels, L., Bakker, A., van Dooren, W., & Drijvers, P. (2019). Conceptual difficulties when interpreting histograms: A review. *Educational Research Review, 28*, Article 100291. https://doi.org/10.1016/j.edurev.2019.100291

Boer, H. de, Donker, A. S., Kostons, D. D., & van der Werf, G. P. (2018). Long-term effects of metacognitive strategy instruction on student academic performance: A meta-analysis. *Educational Research Review, 24*, 98–115. https://doi.org/10.1016/j.edurev.2018.03.002

Booth, R. D. L., & Thomas, M. O. J. (1999). Visualization in mathematics learning: Arithmetic problem-solving and student difficulties. *The Journal of Mathematical Behavior, 18*(2), 169–190.

Borba, R., Azevedo, J., & Barreto, F. (2015). Using tree diagrams to develop combinatorial reasoning of children and adults in early schooling. In *Cerme 9 - Ninth congress of the European society for research in mathematics education* (pp. 2480–2486). hal-01289344.

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

Brod, G. (2021). Generative learning: Which strategies for what age? *Educational Psychology Review, 33*(4), 1295–1318. https://doi.org/10.1007/s10648-020-09571-9

Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Educational Psychology, 105*(2), 380–400. https://doi.org/10.1037/a0031084

Castro-Alonso, J. C., Wong, M., Adesope, O. O., Ayres, P., & Paas, F. (2019). Gender imbalance in instructional dynamic versus static visualizations: A meta-analysis. *Educational Psychology Review, 31*(2), 361–387. https://doi.org/10.1007/s10648-019-09469-1

Cheung, M. W.-L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods, 19*(2), 211–229. https://doi.org/10.1037/a0032968

Cheung, A. C. K., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9*, 88–113. https://doi.org/10.1016/j.edurev.2013.01.001

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

Cromley, J. G., Du, Y., & Dane, A. P. (2020). Drawing-to-learn: Does meta-analysis show differences between technology-based drawing and paper-and-pencil drawing? *Journal of Science Education and Technology, 29*(2), 216–229. https://doi.org/10.1007/s10956-019-09807-6

Csíkos, C., Szitányi, J., & Kelemen, R. (2012). The effects of using drawings in developing young children's mathematical word problem solving: A design experiment with third-grade Hungarian students. *Educational Studies in Mathematics, 81*(1), 47–65. https://doi.org/10.1007/s10649-011-9360-z

Demitriadou, E., Stavroulia, K.-E., & Lanitis, A. (2020). Comparative evaluation of virtual and augmented reality for teaching mathematics in primary education. *Education and Information Technologies, 25*(1), 381–401. https://doi.org/10.1007/s10639-019-09973-5

Diezmann, C. (2000). The difficulties students experience in generating diagrams for novel problems. In *Proceedings of the 24th conference of the international group for the psychology of mathematics education* (pp. 241–248). Nishiki Print Co.

Drijvers, P., & Sinclair, N. (2024). The role of digital technologies in mathematics education: purposes and perspectives. *ZDM Mathematics Education, 56*, 239–248. https://doi.org/10.1007/s11858-023-01535-x

Flores, M. M., Moore, A. J., & Meyer, J. M. (2020). Teaching the partial products algorithm with the concrete representational abstract sequence and the strategic instruction model. *Psychology in the Schools, 57*(6), 946–958. https://doi.org/10.1002/pits.22335

Gilligan, K. A., Thomas, M. S. C., & Farran, E. K. (2020). First demonstration of effective spatial training for near transfer to spatial performance and far transfer to a range of mathematics skills at 8 years. *Developmental Science, 23*(4), Article e12909. https://doi.org/10.1111/desc.12909

Goldin, G. A., & Kaput, J. J. (1996). A joint perspective on the idea of representation in learning and doing mathematics. In L. Steffe, P. Nesher, P. Cobb, G. A. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 397–430). Erlbaum.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*(2), 107. https://doi.org/10.2307/1164588

Hedges, L. V., & Pigott, T. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9*(4), 426–445. https://doi.org/10.1037/1082-989X.9.4.426

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*(1), 39–65. https://doi.org/10.1002/jrsm.5

Hershkowitz, R., Ben-Chaim, D., Hoyles, C., Lappan, G., Mitchelmore, M., & Vinner, S. (1989). Psychological aspects of learning geometry. In P. Nesher, & J. Kilpatrick (Eds.), *Mathematics and cognition (ICMI study series)* (pp. 70–95). University press.

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education, 153*, Article 103897. https://doi.org/10.1016/j.compedu.2020.103897

Johnson, H. L., Gardner, A., Smith, A., Olson, G., & Wang, X. (2019). Interacting with dynamic computer activities impacts college algebra students' math attitudes and performance. *North American Chapter of the International Group for the Psychology of Mathematics Education*.

Ke, F. (2019). Mathematical problem solving and learning in an architecture-themed epistemic game. *Educational Technology Research & Development, 67*(5), 1085–1104.

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine, 22*(17), 2693–2710. https://doi.org/10.1002/sim.1482

Kobiela, M., & Lehrer, R. (2019). Supporting dynamic conceptions of area and its measure. *Mathematical Thinking and Learning, 21*(3), 178–206. https://doi.org/10.1080/10986065.2019.1576000

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65–99. https://doi.org/10.1111/j.1551-6708.1987.tb00863.x

Latifi, M., Hattaf, K., & Achtaich, N. (2021). The effect of dynamic mathematics software Geogebra on student' achievement: The case of differential equations. *Journal of Educational and Social Research, 11*(6), 211. https://doi.org/10.36941/jesr-2021-0141

Lem, S., Onghena, P., Verschaffel, L., & van Dooren, W. (2013). The heuristic interpretation of box plots. *Learning and Instruction, 26*, 22–35. https://doi.org/10.1016/j.learninstruc.2013.01.001

Li, Q., & Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review, 22*(3), 215–243. https://doi.org/10.1007/s10648-010-9125-8#Sec18
Lipsey, M. W., & Wilson, D. B. (2000). *Practical meta-analysis*. Sage.
Lowrie, T., Logan, T., & Hegarty, M. (2019). The influence of spatial visualization training on students' spatial reasoning and mathematics performance. *Journal of Cognition and Development, 20*(5), 729–751. https://doi.org/10.1080/15248372.2019.1653298
Lowrie, T., Logan, T., & Ramful, A. (2017). Visuospatial training improves elementary students' mathematics performance. *British Journal of Educational Psychology, 87*(2), 170–186. https://doi.org/10.1111/bjep.12142
Mejia-Ramos, J. P., & Weber, K. (2019). Mathematics majors' diagram usage when writing proofs in calculus. *Journal for Research in Mathematics Education, 50*(5), 478–488. https://doi.org/10.5951/jresematheduc.50.5.0478
Myers, J. A., Witzel, B. S., Powell, S. R., Li, H., Pigott, T. D., Xin, Y. P., & Hughes, E. M. (2022). A meta-analysis of mathematics word-problem solving interventions for elementary students who evidence mathematics difficulties. *Review of Educational Research, 92*(5), 695–742. https://doi.org/10.3102/00346543211070049
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, n71*. https://doi.org/10.1136/bmj.n71
Presmeg, N. C. (1986). Visualisation in high school mathematics. *For the Learning of Mathematics, 6*(3), 42–46. http://www.jstor.org/stable/40247826.
Presmeg, N. C. (2006). Research on visualization in learning and teaching mathematics. In A. Gutiérrez, & P. Boero (Eds.), *Handbook of research on the psychology of mathematics education* (pp. 205–235). Sense Publishers.
Ran, H., Kim, N. J., & Secada, W. G. (2022). A meta-analysis on the effects of technology's functions and roles on students' mathematics achievement in K-12 classrooms. *Journal of Computer Assisted Learning, 38*(1), 258–284. https://doi.org/10.1111/jcal.12611
Rellensmann, J., Schukajlow, S., Blomberg, J., & Leopold, C. (2021). Does strategic knowledge matter? Effects of strategic knowledge about drawing on students' modeling competencies in the domain of geometry. *Mathematical Thinking and Learning, 25*(3), 296–316. https://doi.org/10.1080/10986065.2021.2012741
Rellensmann, J., Schukajlow, S., Blomberg, J., & Leopold, C. (2022). Effects of drawing instructions and strategic knowledge on mathematical modeling performance: Mediated by the use of the drawing strategy. *Applied Cognitive Psychology, ac3930*. https://doi.org/10.1002/acp.3930
Rodgers, M. A., & Pustejovsky, J. E. (2021). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods, 26*(2), 141. https://doi.org/10.1037/met0000300
Sala, G., Aksayli, N. D., Tatlidil, K. S., Tatsumi, T., Gondo, Y., & Gobet, F. (2019). Near and far transfer in cognitive training: A second-order meta-analysis. *Collabra: Psychology, 5*(1). https://doi.org/10.1525/collabra.203. Article 18.
Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review, 14*(1), 101–120.
Schoenherr, J., & Schukajlow, S. (2023a). Characterizing external visualisization interventions: A systematic literature review. In M. Ayalon, B. Koichu, R. Leikin, L. Rubel, & M. Tabach (Eds.), *Proceedings of the 46th conference of the international group for the psychology of mathematics education* (Vol. 4, pp. 163–170).
Schoenherr, J., & Schukajlow, S. (2023b). Characterizing external visualization in mathematics education research: a scoping review. *ZDM Mathematics Education, 56*, 73–85. https://doi.org/10.1007/s11858-023-01494-3 (2024).
Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2018). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review, 30*(2), 431–455.
Sidney, P. G., Thompson, C. A., & Rivera, F. D. (2019). Number lines, but not area models, support children's accuracy and conceptual models of fraction division. *Contemporary Educational Psychology, 58*, 288–298. https://doi.org/10.1016/j.cedpsych.2019.03.011
Sokolowski, A. (2018). The effects of using representations in elementary mathematics: Meta-analysis of research. *IAFOR Journal of Education, 6*(3), 129–152.
Soni, M., & Okamoto, Y. (2020). Improving children's fraction understanding through the use of number lines. *Mathematical Thinking and Learning, 22*(3), 233–243. https://doi.org/10.1080/10986065.2020.1709254
Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360–407. https://doi.org/10.1598/RRQ.21.4.1
Sterner, G., Wolff, U., & Helenius, O. (2020). Reasoning about representations: Effects of an early math intervention. *Scandinavian Journal of Educational Research, 64*(5), 782–800. https://doi.org/10.1080/00313831.2019.1600579
Stylianou, D. A., & Silver, E. A. (2004). The role of visual representations in advanced mathematical problem solving: An examination of expert-novice similarities and differences. *Mathematical Thinking and Learning, 6*(4), 353–387. https://doi.org/10.1207/s15327833mtl0604_1
Swanson, H. L., Lussier, C. M., & Orosco, M. J. (2015). Cognitive strategies, working memory, and growth in word problem solving in children with math difficulties. *Journal of Learning Disabilities, 48*(4), 339–358. https://doi.org/10.1177/0022219413498771
Tan, C.-K. (2012). Effects of the application of graphing calculator on students' probability achievement. *Computers & Education, 58*(4), 1117–1126. https://doi.org/10.1016/j.compedu.2011.11.023
Uesaka, Y., Manalo, E., & Ichikawa, S. (2010). The effects of perception of efficacy and diagram construction skills on students' spontaneous use of diagrams when solving math word problems. In A. K. Goel, M. Jamnik, & N. H. Narayanan (Eds.), *Diagrammatic representation and inference* (Vol. 6170, pp. 197–211). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-14600-8_19.
Uyen, B. P., Ngan, L. K., Thao, N. P., & Tong, D. H. (2021). Impulsing the development of students' competency related to mathematical thinking and reasoning through teaching straight-line equations. *International Journal of Learning, Teaching and Educational Research, 20*(6), 38–65. https://doi.org/10.26803/ijlter.20.6.3
van Garderen, D., Scheuermann, A., & Jackson, C. (2013). Examining how students with diverse abilities use diagrams to solve mathematics word problems. *Learning Disability Quarterly, 36*(3), 145–160. https://doi.org/10.1177/0731948712438558
van Garderen, D., Scheuermann, A., & Poch, A. (2014). Challenges students identified with a learning disability and as high-achieving experience when using diagrams as a visualization tool to solve mathematics word problems. *ZDM – Mathematics Education, 46*(1), 135–149.
Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3). https://doi.org/10.18637/jss.v036.i03
Viechtbauer, W. (2022). I^2 for multilevel and multivariate models. https://www.metafor-project.org/doku.php/tips:i2_multilevel_multivariate.
Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods, 1*(2), 112–125. https://doi.org/10.1002/jrsm.11
Young, J. (2017). Technology-enhanced mathematics instruction: A second-order meta-analysis of 30 years of research. *Educational Research Review, 22*, 19–33. https://doi.org/10.1016/j.edurev.2017.07.001
Zhang, Q., & Fiorella, L. (2019). Role of generated and provided visuals in supporting learning from scientific text. *Contemporary Educational Psychology, 59*, Article 101808. https://doi.org/10.1016/j.cedpsych.2019.101808
Zhang, Y., Wang, P., Jia, W., Zhang, A., & Chen, G. (2023). Dynamic visualization by GeoGebra for mathematics learning: A meta-analysis of 20 years of research. *Journal of Research on Technology in Education*, 1–22. https://doi.org/10.1080/15391523.2023.2250886
Zulnaidi, H., Oktavika, E., & Hidayat, R. (2020). Effect of use of GeoGebra on achievement of high school mathematics students. *Education and Information Technologies, 25*(1), 51–72. https://doi.org/10.1007/s10639-019-09899-y