Educational Psychology Review (2026) 38:27
https://doi.org/10.1007/s10648-025-10109-0

META-ANALYSIS

Are All Domains the Same? The Effectiveness of Multiple Representation Design in Mathematics Achievement: a Quantitative Systematic Review with Three-level Meta-analysis

Yuxin Liu[1] · Zhongtian Ji[1] · Kan Guo[1]

Received: 9 June 2025 / Accepted: 17 December 2025

Abstract
Mathematical knowledge can be represented by various external representations, such as algebraic expressions and images. While the potential benefits of multiple representations are worth exploring, the appropriate selection and design of these representations should also be addressed. We conducted a systematic review of 17 studies and a three-level meta-analysis based on 106 effect sizes derived from 15 studies to evaluate the overall effectiveness of multiple representations in mathematics achievement and how their design and effectiveness vary across mathematical domains. Additionally, we systematically coded and quantified the characteristics of the included studies, including their major themes and descriptions of their representation designs. Results showed: (1) Compared with single representations, multiple representations had a significant positive effect on students' mathematics achievement (Hedges' $g = 0.286$, $p < .05$, 95% CI [0.069, 0.502]). The additional types introduced in the experimental group significantly moderated the effects. (2) There was substantial heterogeneity across the included studies. The design and effectiveness of multiple representations differed in magnitude across various mathematical domains. More fine-grained categorization is needed in future research to draw more specific conclusions. (3) More in-depth and detailed research is needed to address the gap in current studies, which pay insufficient attention to integrating specific knowledge content when analyzing representation effects and often provide limited detail on representation design.

Keywords Multiple representations · Mathematical domain · Domain specific · Three-level meta-analysis

Extended author information available on the last page of the article

Published online: 03 March 2026

Introduction

When learning mathematics, students encounter mathematical concepts in various forms, such as graphs, tables, diagrams, and symbols. Both the Common Core State Standards for Mathematics (CCSSM, 2011) and the National Council of Teachers of Mathematics (NCTM, 2000) emphasize the importance of representation to facilitate understanding. In recent years, representation has become a popular research topic in educational research. A widely accepted view is that using multiple representations helps students establish connections between mathematical concepts and enhances their performance (Höffler & Leutner, 2007; National Council of Teachers of Mathematics, 2000; Porzio, 1999). However, some empirical studies have found no significant positive effects of multiple representations (Chandler & Sweller, 1992; Jong et al., 1998), and a few studies have even suggested that multiple representations are less effective than traditional instruction (Ainsworth et al., 2002). Such inconsistencies make it difficult to guide effective instructional practice. It is therefore essential to systematically integrate these findings and conduct deeper analysis. One potential explanation of these inconsistencies is that the effects of representation strategies vary across different types of mathematical knowledge. For example, the effects of representations in specific domains of mathematics (e.g., number and quantity) may not be applicable in other domains of mathematics (Beitzel et al., 2011; Flores et al., 2019). However, this possible heterogeneity has not been systematically investigated. The effectiveness of a specific representation design for both mathematics disciplines and their different domains requires further investigation.

In this study, multiple representations are defined as combinations of representations that include more than one type. We systematically reviewed empirical studies that compared instruction employing multiple representations with instruction solely relying on a single representation. By synthesizing these results, we examine the overall effect differences between multiple representations and single representations and explore differences in the applicability of specific multiple-representation designs to different mathematical domains. Furthermore, we synthesize the characteristics of previous studies to provide recommendations for future research, with the goal of promoting the practical applicability of representation research.

Theoretical Background

Representations in Mathematics Education

Representations can be categorized as internal or external. Representations that exist in learners' minds but cannot be observed directly are referred to as internal representations. Representations expressed through language, text, symbols, images, or physical objects are known as external representations (Hiebert & Carpenter, 1992). Meaningful interactions between internal and external representations are fundamental to effective learning (Goldin & Shteingold, 2001). Pande (2021) synthesized a two-way coupling model between external and internal representations on the basis of the perspectives of multiple scholars that suggests that the effectiveness of external representations depends not only on their design but also on learners' individual internal representations. Although internal representations cannot be observed directly, the heterogeneity in the representational effects that result from them should not be overlooked. In this paper, the term "representation" refers to external representation unless explicitly stated as being internal.

Representations can be categorized in various ways depending on the focus, and each discipline has its own classification system (Rexigel et al., 2024). Lesh et al. (1987) proposed a well-established model in the field of mathematics education, identifying five types of representations: (a) real-world contexts, (b) manipulatives, (c) static figures, (d) spoken language, and (e) written symbols. On the basis of several classical classifications, Mainali (2020) noted that graphical, numerical, algebraic, and verbal representations are the most widely used representations in mathematics teaching. Each type has its own advantages in the mathematical domain, and these representations must often be used in combination (Mainali, 2020). Some mathematical concepts can naturally be presented using multiple types of representations; for example, an ellipse can be represented by an analytical formula, a graph, or text.

Mathematical learning is inevitably dependent on multiple forms of representation. The advantages of specific representations in learning do not necessarily generalize across diverse learning contexts. Ruamba et al. (2025) pointed out that the effectiveness of representational strategies may vary across levels of education, owing to differences in cognitive development. From the perspective of cognitive load theory, Lieshout and Xenidou-Dervou (2018) suggested that suitable representation designs may differ by students' prior knowledge. Moreover, in the learning of different areas of mathematical knowledge, both the commonly used types of representations and the effects of a given type of representation may differ (He & Xin, 2025). Within the domain of algebra, representations are employed to examine, explain, estimate and verify the relationship between the variables (Demir & Zengin, 2025; Pinto & Cañadas, 2021). Within the domain of geometry, mathematical knowledge is tightly integrated with visual representations (Ni et al., 2025), which are employed to render abstract concepts more accessible and concrete. Differences inherent in mathematical content lead students to encounter various representations at unequal frequencies, which subsequently affects their ability to translate between representations (Birgin & Eryılmaz, 2025) and may further cultivate representational preferences or dependency. Further research is needed to determine which type of representation or combination of representations is most conducive to students' deep understanding of specific mathematical concepts.

Relevant Learning Theories

Cognitive load theory is a core theory in representation research and is frequently employed to guide the design of representations and to interpret their impacts on students' cognitive processing and performance (Moreno & Mayer, 1999; Ngu et al., 2014; Reed et al., 2013). When multiple representations fail to function as expected, increased cognitive load is considered the cause (Ott et al., 2018). Sweller et al. (1998) identified three types of cognitive load: intrinsic cognitive load (ICL), extraneous cognitive load (ECL), and germane cognitive load (GCL). ICL depends on both the inherent complexity of the learning material and the learner's prior knowledge, which represent the characteristics of external and internal representations, respectively. Equivalent representations that provide redundant information may increase the complexity of the learning material, leading to a higher ICL (Kalyuga et al., 2004). ECL is triggered by inappropriate instructional design (Albus et al., 2021; Hao et al., 2023). There are some benefits of using multiple representations in teaching (de Jong & van der Meij, 2012), but to fully harness their potential, students must synthesize information across multiple representations (Ainsworth, 2006). If students focus on only one type of representation, the presence of additional representations may increase the difficulty of searching for and processing information, leading to higher ECL. The GCL plays a role in reallocating cognitive resources from extraneous processing to intrinsic processing (Sweller et al., 1998). The goal of mathematics instruction is to help learners develop a schema and automate it through practice (Cooper & Sweller, 1987). The most appropriate representation for building a schema is strongly correlated with the knowledge content. For some knowledge, a single representation is sufficient (Berthold & Renkl, 2009), whereas some concepts, such as functions, cannot be fully understood through a single representation. Both the number of representations and the appropriate types of representations may differ across various mathematical domains. According to cognitive load theory, there is a fine line between the positive and negative effects of the use of representation in mathematics learning. Working memory has a limited capacity (Sweller et al., 1998). Multiple representation design is intended to maximize learners' cognitive resources and promote deeper understanding. However, inappropriate designs may induce cognitive load that exceeds working memory capacity, leading to a breakdown of cognitive processing and ultimately impaired learning performance (Mayer & Moreno, 2003). In multiple representation learning contexts, each representational element occupies working memory resources, and the interactions among these elements further increase cognitive load (Van Lieshout & Xenidou-Dervou, 2018). Therefore, it is crucial for effective representation design to understanding both the potential effects of individual representations and their interactions.

Constructivist learning theory supports the use of representations (Cuoco, 2001) and posits their use in explaining concepts (Spiro et al., 1991). Additionally, dual coding theory (Paivio, 1990), the cognitive theory of multimedia learning (CTML, Mayer & Moreno, 1998) and integrated model of text and picture comprehension (ITPC, Schnotz & Bannert, 2003) provide evidence from a cognitive processing perspective supporting multiple representation design. The CTML assumes that verbal and nonverbal information are processed within separate cognitive subsystems. When students receive information through multiple representations, such as text and pictures, the two cognitive subsystems work simultaneously, facilitating the construction of complex and coherent mental models (*The Cambridge Handbook of Multimedia Learning*, 2005). However, common multimedia learning theories cannot be applied to homogeneous combinations of representations (e.g., combinations of textual and algebraic representations). Ainsworth (2006) proposed the Design, Functions, and Tasks (DeFT) framework, which is more suitable for explaining the effects of homogeneous representations. Multiple representations affect learners' cognitive processing through three functions: complement, constrain, and construct. Representations, whether homogeneous or heterogeneous, can contribute to students' deeper understanding of the learning context through these three functions. Therefore, facilitating students' effective integration of multiple representations is a key consideration in representation design. Ainsworth additionally suggested five dimensions that should be considered in the design of multiple representations: (a) the number of representations; (b) the way that information is distributed; (c) the form of the representational system; (d) the sequence of representations; and (e) support for translation between representations. These five dimensions affect mathematics learning. The mechanisms by which representations benefit mathematical learning are intricate and require comprehensive and detailed research.

Factors that Influence Mathematics Achievement with Representation

The focus of representation strategy research should shift from whether representations work to which representations are most appropriate in specific situations. Learning through representations involves a two-way interaction between the internalization of external representations and the externalization of internal representations (Pape & Tchoshanov, 2001). With respect to internal representations, differences in the nature of content knowledge and individual differences among students should be considered. With respect to external representations, representation design (e.g., the type of representation) may also be an influential factor.

Mathematical Domain

Students' internal representations may vary across different domains of mathematics. Mathematics can be divided into closely related domains. NCTM (2000) divides the domain standards from prekindergarten through Grade 12 into five domains: (a) number and operations; (b) algebra; (c) geometry; (d) measurement; and (e) data analysis and probability. CCSSM (2011) divides high school mathematics into the following domains: (a) number and quantity; (b) algebra; (c) functions; (d) geometry; (e) statistics and probability; and (f) modeling, which is integrated across all domains. On the basis of the NCTM (2000) and CCSSM (2011), which balance the need for specificity with the risk of excessive fragmentation, we categorize the mathematics domain from prekindergarten through Grade 12 into (a) numbers and quantity; (b) algebra; (c) functions; (d) geometry; (e) statistics and probability; and (f) measurement. According to *Mathematics: Its Contents Methods and Meaning* (Alexandrov et al., 1963), the number and quantity domain emphasizes number systems and arithmetic principles, the algebra domain abstracts numbers and focuses on transformations of expressions and solving equations, the function domain includes an abstract conceptualization of how one quantity depends on another, and the geometry domain focuses on the quantitative and spatial properties of figures and geometric bodies. *The Teaching and Learning of Statistics: International Perspectives* (Ben-Zvi & Makar, 2016) suggests that the domain of statistics and probability emphasizes approaches to processing data, analyzing data variability, and evaluating event probabilities. Moreover, the measurement domain aims to develop students' understanding of the measurable attributes of objects and their ability to apply suitable techniques, tools, and formulas. It is evident that different mathematical domains highlight distinct core topics.

The internal properties of the learning material influence students' intrinsic cognitive load (Moreno & Mayer, 1999). Kokkonen and Schalk (2021) revealed notable differences between mathematics and the natural sciences with regard to the types of representations used and the learning processes triggered by multiple representations. Studies in different math domains have also reported conflicting results. For example, Beitzel et al. (2011) reported that in the domain of probability and statistics, the integration of information embedded in visual representations consumes substantial cognitive resources and increases cognitive load, thereby reducing available resources for other tasks and impairing performance on probabilistic word problems. In the functions domain, a three-core representation (algebraic, numerical, and graphical) curriculum assisted by graphing calculators concretized abstract concepts, freed up cognitive resources for other tasks, and promoted students' understanding of exponential and logarithmic functions (Ford, 2008). Thus, the differences in this aspect across different mathematical domains must be investigated. If disparities exist between mathematical domains, representation strategies must be explored in more specific mathematical domains to effectively improve teaching and avoid misunderstandings.

Educational Level

Students' educational levels may also influence the structure of students' internal representations. In a current review of multimedia learning theories, the researchers considered educational level a moderating factor (Noetel et al., 2022). According to Piaget's theory of cognitive development (Piaget, 1970), students in different age groups have varying levels of cognitive development. Students in elementary school are generally in the concrete operational stage, so physical or visual representations are easier for them to understand and symbolic representations may impose additional cognitive load on them. Students in middle school and above have entered the formal operations stage, and the simplicity of algebraic representations helps prevent attentional distraction, reduces the cognitive load of gathering relevant information, and may facilitate a deeper understanding of mathematical concepts. Additionally, the knowledge mastered by students varies considerably as their educational level increases. Kalyuga et al. (2000) noted that only learners with very little prior knowledge benefit from multirepresentational environments and that the facilitating effect diminishes as knowledge increases. However, no significant moderating effect of educational level has been found (Noetel et al., 2022). It remains unclear whether educational level plays a moderating role in the effects of multiple representations on math learning.

Characteristics of Representations

Representation design directly influences the effectiveness of representations and is one of the most important characteristics of representations. Different types of representations activate distinct mathematical processes (Gavilán-Izquierdo & Gallego-

Sánchez, 2025). Internal representations are formed based on the encoding format of external representations (Schnotz, 2002; Schnotz & Bannert, 2003). No single representation type or combination of representations is suitable for all learning scenarios (Ainsworth, 2006). Each type of representation has strengths and limitations; for example, numerical representations are intuitive but may lack universality (Ford, 2008). Representation design should align with students' cognitive needs and support them along appropriate learning pathways; otherwise, misalignment may lead to unnecessary exploration and may impose extra cognitive load (Reed et al., 2013). However, more types of representations do not necessarily lead to better learning outcomes. In certain situations, a single representation may be self-contained and intelligible, and any additional representation without essential complementary information may cause additional cognitive load (Ott et al., 2018). The number of representation types may also influence their effectiveness. Representational heterogeneity is a key characteristic of multiple representations. Heterogeneous representations refer to multiple representations that differ in codality, whereas homogeneous representations are the opposite (Ott et al., 2018). Multimedia learning theory emphasizes that heterogeneous representations are processed by different cognitive subsystems, which supports learning (Mayer, 2009). However, some researchers have noted that students may find it difficult to understand the relationships between heterogeneous representations (Ainsworth, 2006), which may lead to learning difficulties.

Empirical Research Trends and Prior Reviews

Ainsworth (2008) identified the early generation of research that evaluated multimedia learning environments. In first-generation experiments, participants were usually given a prior knowledge pretest before being randomly assigned to either a multimedia group or a text-only control group. Then, they engaged in short-term learning followed by a posttest, which was more difficult. If the multimedia group outperformed the control group on the posttest, the findings supported multimedia learning. She noted that these studies did not provide sufficient detail about the design of representations and that it was difficult to apply their findings to other domains, thereby creating considerable challenges for their application in instructional practice. Second-generation researchers must determine who benefits from learning with specific forms of multimedia. A growing body of researchers has recognized this issue, and research on representations is experiencing a paradigmatic shift. Jain and Mitra (2025) employed a qualitative case study design to explore how different types of representations strengthen statistical thinking in secondary education. Birgin and Eryılmaz (2025) investigated students' performance in translating among different representations of fractions. It is necessary to summarize the insights from first-generation studies and provide directions for subsequent second-generation studies.

Researchers have employed meta-analyses and systematic reviews to consolidate findings on the efficacy of representations. For example, Hu et al. (2021) found multimedia problem-solving interventions significantly improved students' response accuracy (Hedges' $g = 0.32$) and response certainty (Hedges' $g = 0.74$). Çeken and Taşkın (2022) provided a systematic overview of multimedia learning principles and reported that most studies were conducted in STEM fields. Rexigel et al. (2024) also reported the frequent presence of multiple representations in STEM education. These authors conducted a meta-analysis in the STEM field to determine whether more representations were better and found that mathematics was the most frequently studied discipline within STEM. Moreover, their results showed that participant characteristics such as educational level and the effects of the same representational design differed across disciplines. Following Rexigel et al. (2024), possible heterogeneity in the influence of representations within a single discipline due to differences in knowledge content must be investigated. However, reviews in the field of mathematics education are lacking. Sokolowski (2018) reported a weighted mean effect size of 0.53 for the impact of using representations on mathematics learning among PreK-5 students. To date, however, no review has explored possible heterogeneity.

The Present Study

We focus on the field of mathematics education and explore the effectiveness of multiple representations on students' mathematics achievement. We aim to comprehensively summarize and consolidate the dispersed results of first-generation studies, in order to guide instructional practice more effectively and offer recommendations for second-generation studies. First, we assess the overall trend of the impact of multiple representations on students' mathematics learning through a meta-analysis. Second, we investigate the heterogeneity of the effects in different mathematical domains through a systematic review. Third, inspired by Ainsworth (2008), we analyze and synthesize the features of prior studies and identify potential pitfalls that future studies should consider.

Specifically, we address the following questions:

1. What is the size of the effect of multiple representations on mathematics achievement compared with a single representation?
2. Is there domain specificity in the design and effectiveness of multiple representations across studies, and what are the similarities and differences in different mathematical domains?
3. What are the features of prior studies, and what aspects still need to be improved?

Method

The research methodology was structured in alignment with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines (Page et al., 2021).

Study Search and Selection

We conducted a literature search using the Web of Science, Education Resources Information Center (ERIC), Scopus, PsycINFO, and ProQuest databases. To ensure the quality of the included studies, only peer-reviewed journal articles and dissertations were included. We did not search for grey literature or trial registries, such as unpublished dissertations or conference papers. We did not set a restriction on the starting date of publication. The search for journal articles was conducted on January 23, 2025, and the search for dissertations was conducted on February 28, 2025. An additional filter was applied in PsycINFO to restrict the participants to human participants.

Search terms were shown in Table 1. Because the term "representation" has many other meanings, we used the search terms *multiple representation** OR *multi-representation** to avoid retrieving many irrelevant papers. Since this study focused on the effects of representations on students, the initial search used the term *student**. After reviewing many of the included articles and their references, we expanded the search terms *student** OR *child** OR *people* OR *kid** OR *learner**. Our study focused only on mathematics education, so the initial search used the term *math**. However, after reviewing many of the included articles and their references, we expanded the search terms to *math** OR *algebra** OR *numer** OR *geometr** OR *statistic** and used the Boolean operator AND to link them. Following Rexigel et al. (2024), we retrieved all records from the title, abstract, and/or keywords.

To answer our research questions, we reviewed and synthesized studies that met the following criteria: (a) the full text was available in English; (b) the study was between-subjects empirical research; (c) the participants in the studies were students; (d) the experimental group received multiple representation interventions, whereas the control group received single representation interventions; (e) the study was related to mathematics education and investigated the impact of representations on mathematics achievement; (f) representations were presented on paper or via electronic devices, and studies that used representations involving physical manipulatives were excluded; and (g) the reported data allowed for the calculation of effect sizes. Studies that met the first six criteria were included in the systematic review. Studies that met all seven criteria were included in the meta-analysis.

Figure 1 presents the PRISMA flowchart of the literature selection process. A total of 5833 studies were retrieved from the selected databases. In the first step of screening, we identified duplicate studies using the web tool Rayyan (Ouzzani et al., 2016) and manually verified these studies for removal. After 1840 duplicates were excluded, 3993 studies remained for the subsequent screening. In the second step, we retrieved the titles and abstracts of the remaining studies on the basis of the predefined inclusion criteria. To avoid incorrect exclusion, studies with insufficient information on the study design in the title or abstract were subjected to full-text screening. After this step, 3840 studies were excluded, leaving 153 studies for the

Table 1 Search terms

Terms related representation		Terms related students		Terms related mathematics
multiple representation* OR multi-representation*	**AND**	student* OR child* OR people OR kid* OR learner*	**AND**	math* OR algebra* OR numer* OR geometr* OR statistic*

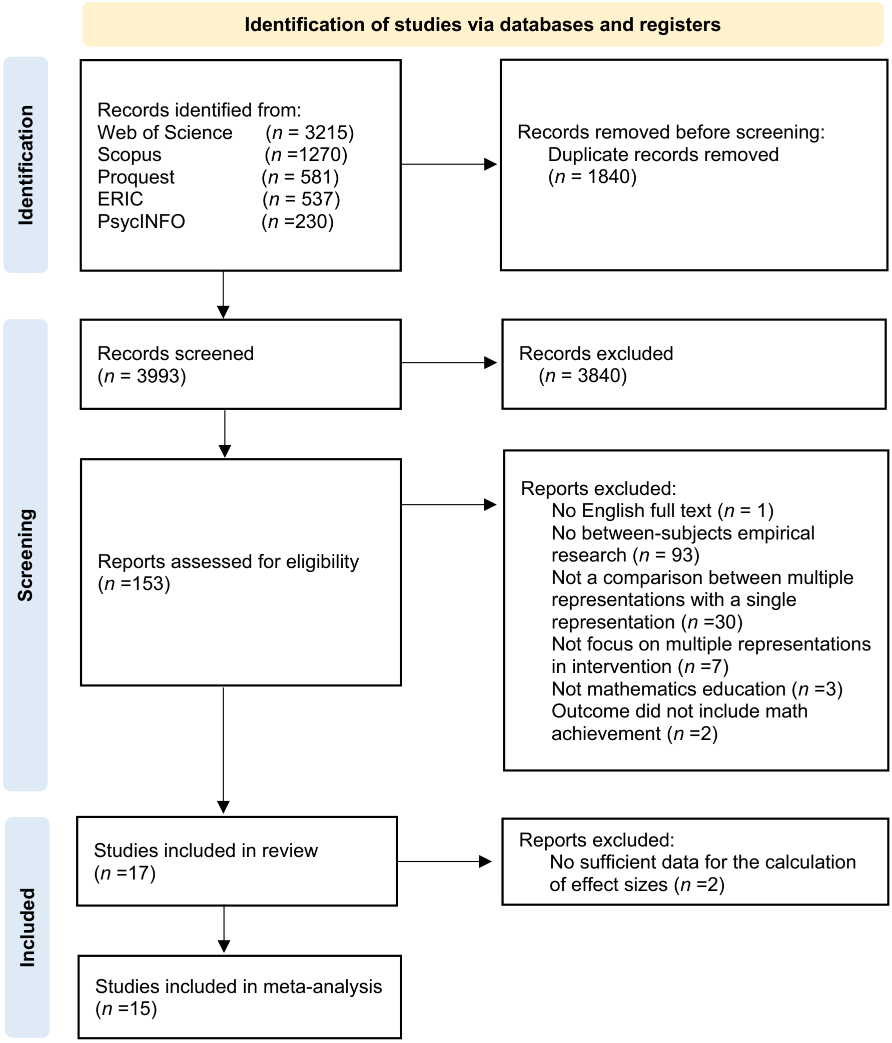

Fig. 1 Flow diagram for the search and inclusion of studies

next step. In the third step, we downloaded the full texts of the remaining studies and assessed them for eligibility. Since we defined multiple representations as involving at least two types of representation, we excluded studies that compared multiple representation instructions with traditional instructions if the latter also emphasized at least two representation types. Studies with data that were insufficient for calculating effect sizes were excluded from the meta-analysis but were included in the systematic review. Ultimately, 15 studies were included in the meta-analysis, and 17 studies were included in the systematic review.

Coding Procedure and Moderator Variables

The first and second authors independently coded all of the studies. Discrepancies were resolved through discussion and re-examination until full agreement was reached. For each eligible study, a detailed coding scheme was developed to document its descriptors and characteristics (Lipsey & Wilson, 2001). The study descriptors covered basic information, including the article title, authors, publication year, study country, sample size, and effect sizes. Inspired by Lu and Ran (2024), study characteristics that might moderate the relationship between representation and mathematics achievement were classified into two types: sample characteristics and research design characteristics. In the present study, the sample characteristics included students' educational level, whereas the research design characteristics initially covered the mathematical domain and the characteristics of representations. To better characterize prior research, we added codes for the description of the representation design of the major/minor themes. The study descriptors were directly coded, whereas the coding criteria for study characteristics were more complex. The coding scheme is described below.

Mathematical Domain We divided the mathematical domain into (a) number and quantity; (b) algebra; (c) functions; (d) geometry; (e) statistics and probability; and (f) measurement. After reviewing all eligible studies, we found that one study examined representations in the mathematical domain of propositional logic (Ott et al., 2018). Although propositional logic is not part of German school education, it is considered important for mathematics learning. Thus, we coded it as "propositional logic".

Educational Level We coded participants' educational level on the basis of the International Standard Classification of Education (ISCED) (Schneider, 2013). According to the UNESCO classification, there are nine educational levels (ISCED 0 to ISCED 8). The studies we included covered four levels: ISCED 1 (primary education), ISCED 2 (lower secondary education), ISCED 3 (upper secondary education), and ISCED 6 (bachelor's degree or equivalent).

Characteristics of Representations Inspired by Mainali (2020), we categorized the types of representations into graphic, numeric, algebraic, and verbal representations. Any graphical expression, such as images, coordinate planes, and function graphs, was coded as a graphic representation. Systematic data displays, such as data tables, were coded as numeric representations. Formulas and symbolic expressions were coded as algebraic representations. Not all written language qualifies as verbal representation; therefore, only the use of written language to describe mathematical knowledge was coded as verbal representation. The types of representations used in both the control and experimental groups were systematically coded. Furthermore, since the experimental group's multiple representations were often developed by supplementing the control group's representation with additional types, the additional types introduced in the experimental group were also coded. Measurements in which the experimental group's representations were not developed by supplementing the control group's representation with additional types were excluded from this coding. To explore whether a greater number of representation types was helpful, the number of representation types was also coded. The heterogeneity of representations is also an important characteristic. Graphic representations and numeric representations are analogically encoded representations, whereas algebraic representations and verbal representations are symbolically encoded representations. Multiple representations with different codalities (e.g., text + graphic) were considered heterogeneous representations, whereas multiple representations with the same codality (e.g., algebraic + text) were considered homogeneous representations (Ott et al., 2018).

Description of Representation Design We divided the description of representation design into three categories: type only (showing only the representation types), examples only (showing incomplete representation designs), and full design (showing complete representation designs). Importantly, if a study did not provide the full design of representations but offered examples that were sufficiently representative to infer the overall representation design, it was classified as a full design.

Major/Minor Theme Ainsworth (2018) identified the major theme and the minor theme of the five special issue articles in "Toward a model of multisource, multimodal processing". She categorized the themes into reader, representation, task and assessment. Inspired by Ainsworth, we categorized the themes into participants, representations and content. We carefully read the full text of each article and coded it on this basis. The major theme was identified as the factor that was both experimentally classified and analyzed in depth. A factor that was only experimentally classified or analyzed in depth was identified as a minor theme.

Statistical Analysis

Calculation of Effect Sizes Hedges' *g* was used to measure effect sizes. The posttest means and standard deviations for the control and experimental groups were used to calculate the effect size, with pretest scores omitted from the calculation. We calculated Hedges' *g* value for each effect using the program Comprehensive Meta-Analysis (Borenstein, 2022). The fixed-effects model assumed that all studies shared the same true effect size, whereas the random-effects model assumed that each study had a unique true effect size, allowing for variations in true effects across studies (Borenstein et al., 2009). Three variance components distributed across the three-level meta-analysis were considered: the sampling variance of the extracted effect sizes at Level 1, the variance between effect sizes from the same study at Level 2, and the variance between studies at Level 3 (Assink & Wibbelink, 2016). This three-level meta-analysis overcomes the limitations of traditional meta-analysis, which considers only between-study variance and may misestimate overall heterogeneity. Previous studies suggest that the mathematical domain and characteristics of representations may influence the effect of representation on mathematics achievement, suggesting potential heterogeneity both within and between studies (Ainsworth, 2006; Beitzel et al., 2011). Thus, we adopted a random-effects model and three-level meta-analysis. We conducted the main meta-analysis and moderator analysis using the *metafor* package in R (Viechtbauer, 2010).

Heterogeneity of Effect Sizes Heterogeneity refers to the true heterogeneity of effect sizes among the included studies rather than the observed variation caused merely by sampling error (Borenstein et al., 2009). Given that a three-level meta-analysis model was employed, we first calculated the variance at Level 1 following the method proposed by Cheung (2014). We then used the log-likelihood ratio test to assess heterogeneity at Levels 2 and 3 and examined the significance and distribution of the total variance. We also calculated the study-level intraclass correlations and effective sample sizes. In addition, leave-one-study-out analyses were conducted to assess the influence of individual studies on the results.

Moderator Effect We examined moderator variables as possible contributors to additional variance in effect size (Ji et al., 2024). Since we use the restricted maximum likelihood (REML) method to estimate the meta-analytic model parameters, we could not compare the fit of the models using a log-likelihood ratio test (Hox et al., 2017). Thus, an omnibus test was conducted to assess whether the potential moderating variables had a significant effect (Assink & Wibbelink, 2016). To address the increased risk of type I errors from multiple moderator tests, we applied the Benjamini–Hochberg correction to control the false discovery rate separately for the moderator tests (Benjamini & Hochberg, 1995; Polanin, 2013). If the omnibus test after Benjamini–Hochberg correction was still significant, we conducted tests to examine differences between each subgroup of the moderator. If the omnibus test yielded a *p*-value below 0.05 but failed to remain significant after the Benjamini–Hochberg correction, we conducted exploratory subgroup analyses to reduce the risk of Type II errors. It should be noted that these exploratory results cannot be regarded as confirmed findings, and further studies are needed to verify their validity. The outcomes of subgroup analyses were corrected using the Benjamini–Hochberg procedure.

Publication Bias Publication bias occurs when some studies are omitted due to nonpublication or inaccessibility, leading to an incomplete representation of the research field (Gao et al., 2017). To analyze publication bias, we first qualitatively assessed the symmetry of the funnel plot through visual inspection. Next, we conducted a quantitative test of the symmetry of the funnel plot using Egger's regression test (Egger et al., 1997). Finally, we used Rosenthal's fail-safe N formula (Rosenthal, 1979) to assess the relationship between effect sizes and standard errors. If the fail-safe N exceeded $5n + 10$ (where n is the number of included studies), publication bias was unlikely to affect the results.[1]

Results

We included a total of 17 studies in the systematic review. After we excluded studies that did not have sufficient data to calculate an effect size, 15 studies were eligible for the meta-analysis. First, descriptive statistics on variables such as publication year were reported to summarize the characteristics of the studies included in the systematic review. To provide a more comprehensive overview, we present more detailed information from the included studies (Table 2). Second, we calculated the overall effect sizes of the studies using a three-level meta-analysis and examined potential heterogeneity and moderator effects. Finally, we analyzed the similarities and differ-

[1] We also assessed the risk of bias using the RoB 2 tool, with results reported in the appendix. Detailed coding procedures are available from the authors. Many of the included studies did not report a pre-specified analysis plan, indicating a potential risk of selective reporting. However, our analyses revealed no evidence of inflated effects attributable to inadequate randomisation.

Table 2 Overview of the meta-analysis results

Moderator variable		*n*	*k*	N	Hedges' *g* (95% CI)	F (df1, df2)	*p* value	Level 2 variance	Level 3 variance
a. Educational level	Primary education	2	10	96	0.384 (−0.284, 1.053)	F (3, 102) = 0.598	.618	0.049***	0.176***
	Lower secondary education	4	18	363	0.126 (−0.333, 0.585)				
	Upper secondary education	6	60	556	0.457* (0.080, 0.835)				
	Bachelor's or equivalent	3	18	486	0.118 (−0.394, 0.631)				
b. Mathematical domain	Number and Quantity	4	15	240	0.154 (−0.305, 0.613)	F (5, 100) = 0.948	.453	0.049***	0.157***
	Algebra	3	6	207	0.607* (0.045, 1.169)				
	Functions	3	13	351	0.231 (−0.264, 0.726)				
	Geometry	1	10	91	0.413 (−0.398, 1.223)				
	Statistics and Probability	3	54	466	0.031 (−0.437, 0.499)				
	Propositional logic	1	8	146	0.849* (0.020, 1.679)				
c. Representation quantity (experimental group)	2	9	79	879	0.169 (−0.101, 0.438)	F (2, 103) = 3.261	.042[a]	0.045***	0.152***
	3	6	22	398	0.609*** (0.274, 0.944)				
	4	2	5	292	0.064 (−0.540, 0.667)				
d. Types of representations (control group)	Graphic	4	30	325	0.514*** (0.220, 0.808)	F (3, 102) = 2.852	.041[a]	0.042***	0.173***
	Numeric	1	3	102	0.124 (−0.790, 1.037)				
	Algebraic	11	55	1102	0.264* (0.013, 0.515)				
	Verbal	3	18	248	0.143 (−0.171, 0.458)				

Table 2 (continued)

Moderator variable		*n*	*k*	N	Hedges' *g* (95% CI)	F (df1, df2)	*p* value	Level 2 variance	Level 3 variance
E. Additional types of representations (experimental group)	Graphic	6	23	404	−0.035 (−0.366, 0.296)	F (7, 84) = 3.560	.002	0.014***	0.193***
	Numeric	2	6	152	0.366 (−0.086, 0.819)				
	Algebraic	3	19	274	0.301 (−0.054, 0.655)				
	Verbal	4	17	242	0.425* (0.052, 0.797)				
	Graphic& Numeric	3	10	231	0.541 (−0.022, 1.103)				
	Graphic& Algebraic	2	3	83	0.681* (0.138, 1.224)				
	Graphic& Verbal	2	9	106	0.527 (−0.060, 1.115)				
	Graphic& Numeric& Verbal	2	5	292	0.061 (−0.588, 0.709)				

N Number of studies; *k* = number of effect sizes; N = number of unique participants; Level 2 variance = variance between effect sizes extracted from the same study; Level 3 variance = variance between studies

**p* < .05; ***p* < .01; ****p* < .001

[a] Variables no longer significant after Benjamini–Hochberg correction

ences in the design and effectiveness of representations across various mathematical domains.

Characteristics of the Included Studies

Publication Year

We did not limit the publication years in our literature search. Eligible studies were published between 1995 and 2022. As early as 1995, Rich (1995) validated the potential of multiple representation strategies. In the following thirty years, eligible studies continued to emerge every few years, although the number of eligible studies remained limited. There were three publications in 2018 and two publications in 2021, but either no articles or only one article was published in all other years.

Study Country

The included experiments were conducted across eight countries. The United States accounted for the majority of publications, with eight publications and 47% of all included articles. Other countries contributed fewer studies: Iran and Germany contributed two publications each, whereas Taiwan, Australia, the United Kingdom, Ethiopia, and the Netherlands contributed only one each.

Educational Level

The eligible studies covered four educational levels: primary education ($n=2$), lower secondary education ($n=5$), upper secondary education ($n=7$), and bachelor's degree or equivalent ($n=3$). The majority of the studies focused on lower secondary education and upper secondary education.

Mathematical Domain

Among the studies we included, number and quantity ($n=5$, 29%) was the most frequently studied mathematical domain. Fractions, decimals, and percentages emerged as a major focus in this domain. For example, Flores et al. (2019) and Han et al. (2016) focused on teaching fractions, decimals, and percentages. Ngu et al. (2014) addressed problems related to percentage changes, such as, "If your father wants to increase your weekly allowance of $20 by 5%, what would your new weekly allowance be?" Moreover, previous researchers have focused on arithmetic and estimation. Moreno and Mayer (1999) explored integer operations, whereas Ainsworth et al. (2002) focused on estimation in primary mathematics.

Functions were the second most studied domain ($n=4$, 23.5%). In this domain, the role of multiple representations in deepening conceptual understanding was investigated. Brenner et al. (1997) examined the understanding of various representations of functions. Amirbostaghi et al. (2021) investigated students' understanding of the limit symbol. Additionally, Arefaine et al. (2022) and Rich (1995) explored the conceptual understanding of calculus topics such as derivatives.

A portion of the included studies focused on algebra ($n=3$, 17.6%), statistics and probability ($n=3$, 17.6%). In the algebra domain, Reed et al. (2013) and Daryaee et al. (2018) focused on supporting students' equation construction and understanding of algebraic concepts such as equations. Goins (2001) specifically investigated polynomial multiplication. In the statistics and probability domain, Berthold and Renkl (2009) tested whether multiple representations benefit a conceptual understanding of probability, whereas Beitzel et al. (2011) questioned whether they truly help to solve probability word problems such as calculating the joint probability of independent events. Kolloffel (2000) integrated combinatorics and specifically emphasized the calculation of the probability of specific combinations.

Only one study focused on the geometry domain. Liang and She (2023) examined how different representational scaffolds support students' understanding of the interior and exterior angle sum theorems of triangles. Moreover, one article focused on the propositional logic domain. Ott et al. (2018) investigated the effects of multiple symbolic representations on students' learning and problem solving in propositional logic.

Design of Representations

As shown in Fig. 2, algebraic representations (k=57, 53%) were the most common representations used in the control groups, accounting for more than half of all the representations. Some measurements also used graphic representations (k=30, 28%) and verbal representations (k=18, 16%) as single representations in the control groups. In contrast, numerical representations (k=3, 3%) were seldom used independently as instructional materials.

Specifically, within measurements that employed algebraic representation in the control condition, most studies used only arithmetic operators such as addition and subtraction together with numerals or unknowns represented by letters for arithmetic or equation solving. For example, polynomial multiplication was carried out using the distributive property and simplification by combining similar terms without supportive representations (Goins, 2001). Furthermore, some measurements employ specialized symbolic expressions of knowledge. For example, propositional logic uses symbols such as ∀ (for all), ∃ (there exists), ∩ (intersection), and ∪ (union) (Ott et al., 2018). Within measurements that employ graphic representation in the control condition, representation selection is more diverse. Berthold and Renkl (2009) and Kolloffel (2000) employed tree diagrams to calculate the probability of specific events. Liang and She (2023) used triangle diagrams to determine given angles, whereas Ainsworth et al. (2002) used graphical representations such as the splatwall to support estimation. Within measurements employing verbal representation in the control condition, verbal representation was employed in problem description, analysis, and solution description. For example, Ott et al. (2018) utilized written language

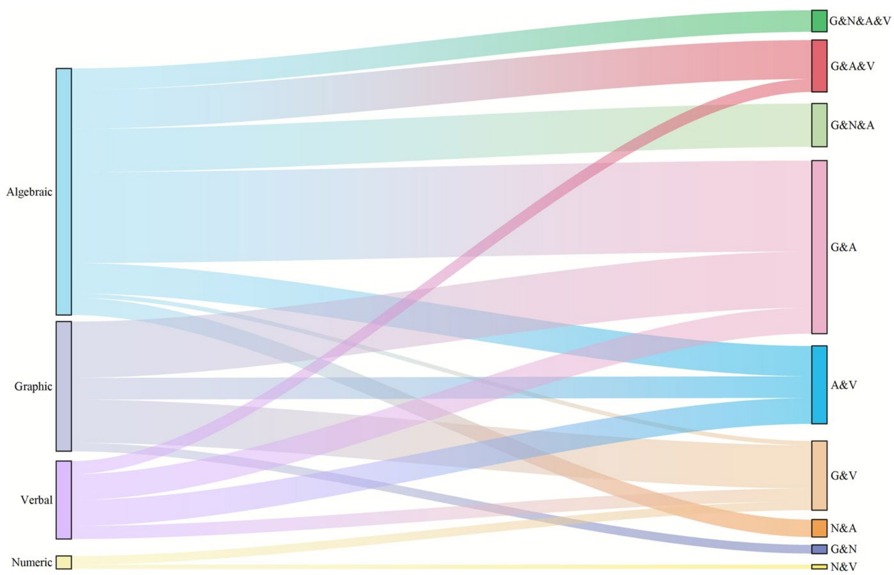

Fig. 2 Representation types. The left side indicates the representation types in the control group. The right side indicates those in the experimental group. The abbreviations are as follows: G, graphic representation; N, numeric representation; A, algebraic representation; and V, verbal representation

to express logical propositions. Ngu et al. (2014) and Kolloffel (2000) used written language to interpret problems and describe the steps to a solution. Only one study employed numerical representations in the control condition. Reed et al. (2013) used tables to analyze the relationships between variables in problem situations.

Most of the measurements ($k = 81$, 75%) used two different types of representations in the experimental groups. Specifically, the combinations of two representation types included graphic and algebraic ($k = 40$, 37%), algebraic and verbal ($k = 18$, 17%), graphic and verbal ($k = 16$, 15%), numeric and algebraic ($k = 4$, 4%), graphic and numeric ($k = 2$, 2%), and numeric and verbal ($k = 1$, 1%). Some measurements involved three types of representations ($k = 22$, 20%) in the experimental groups: the combination of graphic, algebraic, and verbal representations ($k = 12$, 11%) and the combination of graphic, numeric, and algebraic representations ($k = 10$, 9%). Only a small number of measurements included all four types of representations ($k = 5$, 4%). Specifically, in several studies, multiple representations in the experimental group combined the single representations used in different control groups. For example, Berthold and Renkl (2009) used tree diagrams in one control group and symbolic operations in another control group to calculate probability, whereas the experimental group combined both. In most studies, multiple representations in the experimental group were based on the single representation used in the control groups and incorporated different types of representations. For example, the control group was provided with graphic representation scaffolds only, whereas the experimental group included both graphic and algebraic scaffolds (Liang & She, 2023). Because some studies contained several single representation groups and multiple representation groups, multiple representation design in a small number of measurements was not based on a single representation.

Effects of Multiple Representations

Among all 108 measurements, 77 measurements supported multiple representation strategies, whereas 29 measurements did not.

Consistent results were observed across several measures in comparisons of multiple representations and single representations. Seven measures reported by Ott et al. (2018), Kolloffel (2000) and Amirbostaghi et al. (2021) indicated that the combination of algebraic and verbal representations was more effective than algebraic representations alone. Five measures reported by Kolloffel (2000) indicated that the combination of algebraic and verbal representations was more effective than graphic representations. Six measures reported by Ott et al. (2018) and Kolloffel (2000) indicated that the combination of algebraic and verbal representations was more effective than verbal representations. Thirteen measures reported by Berthold and Renkl (2009) and Kolloffel (2000) indicated that the combination of algebraic and graphic representations was more effective than graphic representations alone. Two measures reported by Ainsworth et al. (2002) indicated that the combination of graphic and numeric representations was more effective than graphic representations alone. One measure reported by Ott et al. (2018) indicated that the combination of graphic and verbal representations was more effective than algebraic representations. Ten measures reported by Liang and She (2023) indicated that the combination of graphic and verbal representations was more effective than graphic representations alone. One measure reported by Reed et al. (2013) indicated that the combination of numeric and verbal representations was more effective than the combination of numeric representations. Three measures reported by Ngu et al. (2014) and Ott et al. (2018) indicated that the combination of graphic, algebraic, and verbal representations was more effective than verbal representations alone. Additionally, some measures converged in favor of single representations, such as six measures reported by Ott et al. (2018). Kolloffel (2000) indicated that verbal representations were more effective than the combination of graphic and algebraic representations. Four measures reported by Beitzel et al. (2011) indicated that algebraic representations were more effective than the combination of numeric and algebraic representations.

However, discrepant results were observed across several measures in certain comparisons of multiple representations and single representations. Ten measures reported by Berthold and Renkl (2009), Ott et al. (2018), Kolloffel (2000), Goins (2001) and Han et al. (2016) suggested that the combination of graphic and algebraic representations outperformed algebraic representations, whereas eleven measures reported by Berthold and Renkl (2009), Beitzel et al. (2011) and Kolloffel (2000) supported the opposite conclusion. One measure reported by Reed et al. (2013) suggested that the combination of graphic and verbal representations outperformed numeric representations, whereas one measure reported in the same study supported the opposite conclusion. One measure reported by Ott et al. (2018) reported that the combination of graphic and verbal representations outperformed verbal representations, whereas two measures reported by Ngu et al. (2014) supported the opposite conclusion. Seven measures reported by Ott et al. (2018) and Moreno and Mayer (1999) suggested that the combination of graphic, algebraic, and verbal representations outperformed algebraic representations, whereas two measures reported by Moreno and Mayer (1999) supported the opposite conclusion. Eight measures reported by Daryaee et al. (2018) and Rich (1995) suggested that the combination of graphic, numeric, and algebraic representations outperformed algebraic representations, whereas one measure reported by Flores et al. (2019) supported the opposite conclusion. Three measures reported by Arefaine et al. (2022) and Brenner et al. (1997) suggested that the combination of graphic, numeric, algebraic, and verbal representations outperformed algebraic representations, whereas two measures reported in the same studies supported the opposite conclusion.

Description of the Representation Design

Eight studies (47.06%) did not provide a complete description of the representation design. Among them, six studies (35.30%) offered examples that were insufficient to infer the full design, whereas two studies (11.76%) offered only the types of representations. For example, Arefaine et al. (2022) mentioned that the multiple representations group focused on four types of representations (verbal, numerical, graphical, and algebraic), but these authors did not provide any details about their representation design. Similarly, Beitzel et al. (2011) stated that only the specific representation group used visual representations such as tree diagrams and Venn diagrams, while the matrix representation group employed tables and matrices without any examples.

In contrast, some studies provided clear and complete representational designs. Ngu et al. (2014) presented the complete instruction sheet in the appendix, and Goins (2001) presented the complete teaching method in the appendix. Although Moreno and Mayer (1999) did not present a full representation design, the selected frames from the single/multiple representation program were sufficiently representative to infer the remaining representation designs.

Major/Minor Themes

Representation was the major theme in all 17 studies. However, only two studies (11.76%) considered participants the major theme, and two studies (11.76%) considered content the major theme. Among the two studies that emphasized participants as the major theme, Brenner et al. (1997) examined the effects of multiple representations on students from different language backgrounds by dividing students into two groups: those with English as a first language and those with English as a second language. Moreno and Mayer (1999) investigated the possible differences among students with different academic performance and spatial skills when learning with multiple representations. Among the two studies that considered content the major theme, Ngu et al. (2014) classified tasks into simple and complex for analysis and analyzed outcomes on the basis of percentage knowledge. These authors noted that horizontal lines illustrated the underlying structure of percentage change problems and thereby reduced students' cognitive load. Like Ngu et al. (2014), Kolloffel (2000) not only categorized knowledge into different types (e.g., conceptual, procedural) but also analyzed why the combination of verbal and algebraic representations outperformed the combination of graphic and algebraic representations on the basis of probability knowledge. A possible explanation is that solving probability problems requires a specific sequence of reasoning steps to be followed. The reasoning steps in tree diagrams were more implicit. For the combination of verbal and algebraic representations, the text guided the students to think in sequence, and algebraic representations such as equations effectively reinforced this sequence, which aligns with the problem-solving requirements in the statistics and probability domains.

Only eight studies included a minor theme. Among them, six studies (35.30%) considered content the minor theme, whereas one study (5.88%) considered participants the minor theme. However, only three studies (17.65%) incorporated an analysis based on the mathematical knowledge involved, with the remaining three (17.65%) discussing only simplified classifications such as conceptual knowledge and procedural knowledge. Among the three studies that incorporated the mathematical knowledge involved, Moreno and Mayer (1999) discussed the metaphor of arithmetic as motion and suggested that representing arithmetic as a jumping rabbit moving on a number line benefited students' deep understanding of arithmetic. Rich (1995) noted that algebraic representations are commonly used for symbolic operations on equations, numerical representations are used to describe specific values of functions and derivatives, and graphical representations offer an intuitive way to identify mathematical features such as maximum and minimum points. Amirbostaghi et al. (2021) suggested that substituting the "−" and "+" symbols in limit expressions with the words "left" and "right" makes it easier for students to mentally simulate movement along the number line and obtain a better understanding of limit convergence. Arefaine et al. (2022) considered participants the minor theme and explored students' preferences in representation interpretation, such as a preference for local or global explanations.

Meta-analysis

Overall Average Effect Size

The meta-analysis included 15 studies with 106 effect sizes. The included effect sizes ranged from −0.9314 to 1.7679. Among these effect sizes, 76 were positive, 29 were negative, and one was zero. According to the three-level random effects model, the overall effect size indicated a significant positive effect of multiple representations on students' mathematics achievement (Hedges' $g = 0.286$, $p < .05$, 95% CI [0.069, 0.502]).

Heterogeneity of Effect Sizes

The Q statistic revealed significant unexplained heterogeneity ($Q\ (105) = 294.956$, $p < .001$), which supported the use of a random effects model. On the basis of log-likelihood ratio tests, both within-study variance ($p < .001$) and between-study variance ($p < .001$) were found to be statistically significant, indicating significant heterogeneity at these levels. Variance decomposition revealed that individual-level differences (Level 1) accounted for 27.60% of the total variance, within-study effect size differences (Level 2) accounted for 19.18%, and between-study variance (Level 3) contributed the most at 53.22%, indicating the main source of heterogeneity. The study-level intraclass correlation was 0.74, and the total effective sample size across all studies was 1,125.24, further supporting the use of the three-level meta-analytic model. Effect size estimates after excluding individual studies ranged from 0.226 to 0.346, compared to 0.286 in the full model. It indicated that no single study exerted an extreme influence. Variance components at Level 2 and Level 3 varied slightly across exclusions, with Beitzel et al. (2011) showing a relatively larger impact on heterogeneity. Overall, the meta-analytic results were robust to the exclusion of individual studies.

Moderating Effect Analysis

We conducted moderation analyses for several variables, including mathematical domain, educational level, and characteristics of representations (specifically, the number of representation types in the experimental group, types of representations in the control group, additional types of representations in the experimental group and heterogeneity of representations). The omnibus test revealed that the number of representation types in the experimental group, types of representations in the control group and additional types of representations in the experimental group were significant moderators. However, only the moderator effect of additional types of representations in the experimental group remained significant after the Benjamini–Hochberg correction. The effects of the number of representation types in the experimental group and the types used in the control group were challenged due to the increased risk of type I errors.

Mathematical Domain The mathematical domain did not significantly influence the overall effect (p=.453). Significant subgroup effects were detected for algebra (Hedges' g=0.607, p<.05, 95% CI [0.045, 1.169]) and propositional logic (Hedges' g=0.849, p<.05, 95% CI [0.020, 1.679]).

Educational Level Educational level did not significantly influence the overall effect (p=.618). Significant subgroup effects were detected for upper secondary education (Hedges' g=0.457, p<.05, 95% CI [0.080, 0.835]).

Number of Representation Types (Experimental Group) The moderating effect of the number of representation types in the experimental group (p=.042) did not remain significant after the Benjamini–Hochberg correction. Table 3 presents the results of exploratory subgroup comparisons. However, these results should be interpreted with caution and regarded as hypothesis-generating for future research.

Types of Representations (Control Group) The moderating effect of the type of representation in the control group (p=.041) did not remain significant after the Benjamini–Hochberg correction. Table 4 presents the results of exploratory subgroup comparisons. However, these results should be interpreted with caution and regarded as hypothesis-generating for future research.

Additional Types of Representations (Experimental Group) Additional types of representations in the experimental group remained significant in influencing the overall effect (p=.002) after the Benjamini–Hochberg correction. Because the representations used in some experimental groups were not derived by adding additional representations to the control groups, we initially excluded these measurements from the moderation analysis. To ensure result stability, we recoded the representations used in these experimental groups as additional types of representations and included them in the moderation analysis. The results indicated that the moderating effect

Table 3 Exploratory subgroup analysis of the number of representation types

	β	SE	p value	95% CI
2 vs. 3	0.440	0.181	.017	(0.081, 0.800)
2 vs. 4	−0.105	0.333	.753	(−0.766, 0.556)
3 vs. 4	−0.545	0.348	.120	(−1.236, 0.145)

SE Standard error; *CI* Confidence interval

Table 4 Exploratory subgroup analysis of the types of representations (control group)

	β	SE	p value	95% CI
Algebraic vs. Graphic	0.250	0.111	.027[a]	(0.029, 0.471)
Algebraic vs. Numeric	−0.141	0.478	.769	(−1.088, 0.807)
Algebraic vs. Verbal	−0.121	0.123	.330	(−0.366, 0.124)
Graphic vs. Numeric	−0.390	0.484	.422	(−1.350, 0.569)
Graphic vs. Verbal	−0.371	0.136	.007	(−0.639, −0.102)
Numeric vs. Verbal	0.020	0.487	.968	(−0.946, 0.986)

SE Standard error; *CI* Confidence interval

[a] Variables no longer significant after Benjamini–Hochberg correction

remained significant (p=.009). Subgroup analysis indicated that only adding verbal representations (Hedges' g=0.425, p<.05, 95% CI [0.052, 0.797]) or the combination of graphic representations and algebraic representations (Hedges' g=0.681, p<.05, 95% CI [0.138, 1.224]) yielded a significant positive effect. Given the large number of between-group comparisons, Table 5 reports only results with p-values under 0.05. As illustrated in Table 5, adding algebraic representations (β = 0.336, p<.01), verbal representations (β= 0.460, p<.01), and the combination of graphic and algebraic representations (β= 0.716, p<.01) all significantly outperformed the addition of graphic representations. Although adding numerical representations also showed an advantage over graphic representations, this effect did not remain significant after the Benjamini–Hochberg correction. It should be noted that the result for the combination of graphic and algebraic representations relies on only three effect sizes, which may reduce the stability of the result and warrants cautious interpretation.

Publication Bias Analysis

A visual inspection of the funnel plot (Fig. 3) indicated that the effect sizes were distributed symmetrically, suggesting a small risk of publication bias. This observation was supported by Egger's regression test, which was not statistically significant (p>.05), indicating no evidence of publication bias. Rosenthal's fail-safe N analysis suggested that at least 926 additional studies were required to invalidate the current results, far exceeding the threshold of 85 and indicating the robustness of the current findings.

Representations in Different Mathematical Domains

Design of Representations

The single representations in the control groups varied across different mathematical domains. Specifically, the function domain included only algebraic representations, and the geometry domain included only graphic representations. The algebra domain included algebraic or numerical representations, and the propositional logic domain included algebraic or verbal representations. The number and quantity domain and the statistics and probability domain included algebraic representations, graphic representations, and verbal representations. Given that only one study examined the geometry domain and the propositional logic domain, their conclusions may be

Table 5 Subgroup analysis of the additional types of representations (experimental group)

	β	SE	p value	95% CI
Graphic vs. Algebraic	0.336	0.110	.003	(0.118, 0.554)
Graphic vs. Numeric	0.402	0.189	.036[a]	(0.026, 0.777)
Graphic vs. Verbal	0.460	0.151	.003	(0.161, 0.759)
Graphic vs. Graphic& Algebraic	0.716	0.246	.005	(0.227, 1.205)

SE Standard error; *CI* Confidence interval

[a] Variables no longer significant after Benjamini–Hochberg correction

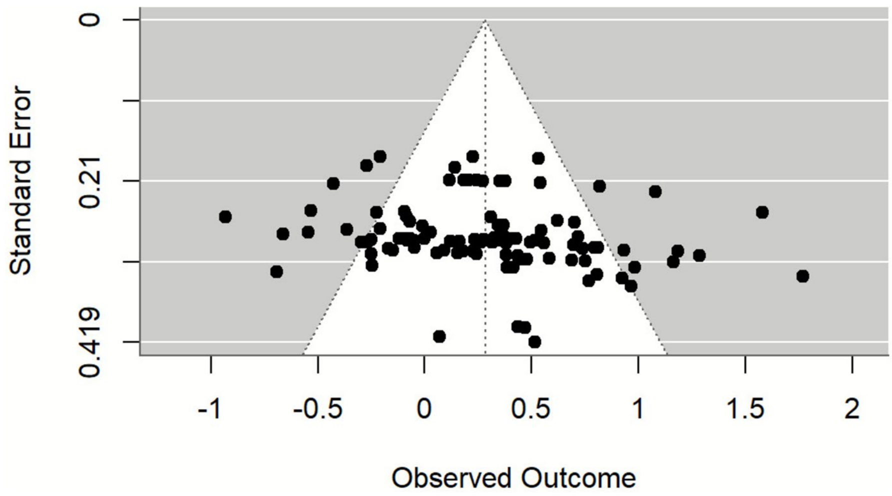

Fig. 3 Funnel plot of standard error for effect sizes

unstable and should be interpreted with caution. Although different mathematical domains sometimes adopt the same type of representation, its concrete form may differ. This phenomenon is particularly common in the use of graphic representations. For example, in the algebra domain and the number and quantity domain, number lines and percentage bar diagrams are commonly used. In the geometry domain, geometric figures such as triangles are commonly used. In the function domain, graphs of functions are commonly used. In the statistics and probability domain, tree diagrams and Venn diagrams are commonly used.

All the domains tended to adopt heterogeneous representations in the experimental group. Only two studies (11.76%) excluded heterogeneous combinations, and only four studies (23.53%) included homogeneous combinations.

Effects of Multiple Representations

All the domains had more than half of the measures, indicating that multiple representations were more effective, although specific proportions differed. Specifically, the quantity of measures supporting multiple representations was as follows: all 10 in the geometry domain, 5 out of 6 measures in the algebra domain, 7 out of 8 in the propositional logic domain, 12 out of 14 in the function domain, 11 out of 16 measures in the number and quantity domain, and 34 out of 54 in the statistics and probability domain.

Among all the comparisons, the combination of graphic and algebraic representations versus algebraic representations was included in the largest number of studies and measures. Ott et al. (2018), Goins (2001), and Han et al. (2016) provided evidence in the domains of propositional logic, algebra, and number and quantity, showing that the combination of graphic and algebraic representations was more effective for learning mathematics than algebraic representations alone. However, in the domain of statistics and probability, Berthold and Renkl (2009) and Kolloffel

(2000) reported that the combination of graphic and algebraic representations was more effective for conceptual knowledge, whereas algebraic representations alone led to better outcomes for procedural knowledge. Beitzel et al. (2011) even reported that students learned better with algebraic representations in both near- and far-transfer tasks within the domain of statistics and probability.

Discussion

The General Effect of Multiple Representations in Mathematics

Compared with single representations, multiple representations yield a positive overall effect on mathematics achievement (Hedges' $g = 0.286$, $p < .05$, 95% CI [0.069, 0.502]). This finding indicates that multiple representations are generally more effective for learning than a single representation, which aligns with the widely accepted view (Höffler & Leutner, 2007; National Council of Teachers of Mathematics, 2000; Porzio, 1999). However, the effect (Hedges' $g < 0.4$) did not reach Hattie's hinge point for visible learning (Hattie, 2008, 2023). Although the intervention showed a positive effect, it was not as substantial as hypothesized. This finding aligns with the results of our systematic review. A total of 27.36% effect sizes included in the meta-analysis were negative, indicating that inappropriate multiple representation designs are not uncommon in practice. These findings collectively suggest that confidence in the effectiveness of multiple representation strategies should be maintained, but careful attention must be paid to how representations are designed and implemented in practice. Currently, there is still a lack of research on how to design the most effective representations. It is essential for educators to first understand how to optimize their use before restructuring lessons and creating new materials. In particular, it is crucial to clarify the cognitive impacts of different representation designs on students' specific learning activities. Our results revealed several findings regarding the design of representations.

First, we may have overestimated the role of visualization strategies in multiple representation design, since they do not always free up cognitive resources for deeper learning processes. Based on dual coding theory and the cognitive theory of multimedia learning, the prevailing view was that graphic representations can serve as useful instructional aids (Levin & Lesgold, 1978) and help students connect abstract concepts with concrete images and enhance comprehension (Schnotz, 2002; Sierpinska, 2013). However, it may be insufficient for fully understanding concepts in mathematics (Duval, 1999) and may fail to convey complete information and lose their advantage when they lack additional complementary representations (Koedinger & Nathan, 2004). Additional types of representations in the experimental group significantly moderated the effects of multiple representations. The subgroup analysis indicated that adding graphic representations was less effective than the addition of algebraic ($\beta = -0.336$, $p < .01$) or verbal representations ($\beta = -0.460$, $p < .01$). The incorporation of graphic representations requires working memory resources as new elements to be processed, and the integration of representations can further increase cognitive load (Arneson & Offerdahl, 2023). For novices, working memory is more likely to become overloaded due to limited schemas for processing such information (Cook, 2006). The systematic review revealed specific examples. Ngu et al. (2014) found that the equation method led to better learning outcomes when solving percentage change problems. In contrast, the percentage bar diagram cannot clarify the solution strategy and may cause additional cognitive load because it searches for information that is not explicitly provided in the diagram. Representation design should aim to reduce extraneous cognitive load, match intrinsic cognitive load, and promote germane cognitive load to support effective learning (Hsu & Hsu, 2025). The effectiveness of representations is determined not only by the completeness of content coverage, but more importantly by how easily learners can recognize the key (Qin & Wu, 2025). Given that novices have limited capacity to simultaneously process highly interactive information, the combination of graphic and other representations with high interactivity should be implemented with caution. However, these results do not suggest that graphic representations are always less effective, but their effectiveness depends on how well they are aligned with cognitive demands and task characteristics. In many cases, the negative effects of multiple representations may result from a misalignment between the representation design and the learning tasks. For example, Berthold and Renkl (2009) and Kolloffel (2000) reported that students learn better with only algebraic representations when learning procedural knowledge. However, the combination of graphic and algebraic representations was more effective in learning conceptual knowledge than algebraic representations alone. Procedural knowledge, which focuses on actions or manipulations, requires minimal cognitive load. However, the integration of conceptual knowledge relies on germane cognitive load. Visualization strategies tend to enhance germane cognitive load, but this is unnecessary for procedural knowledge (Berthold & Renkl, 2009). The design of multiple representations should support students' specific cognitive needs. In addition, factors such as connections among representations and temporal and spatial arrangements can significantly influence their effectiveness. Thus, careful design is needed to promote learning.

Second, due to the limited number of studies involving four types of representations and the lack of significant moderating effects, we cannot conclude whether more representations are always better in mathematics education. The information provided may be insufficient when the number of representation types is not enough. Conversely, too many representation types may increase the difficulty of identifying and integrating useful information (Kalyuga et al., 2004), thereby overloading learners' working memory and raising extraneous cognitive load. Although exploratory subgroup analysis offers hypothesis-generating insights, the impact of the number of representation types remains an open question for future research. Rexigel et al. (2024) have found that more representations lead to better outcomes in STEM education. They suggested that integrating multiple representations can mitigate the complexity of single representations, lower intrinsic and extraneous cognitive load, and free up additional resources for germane cognitive processing. It is essential to determine whether such an impact remains valid within the domain of mathematics.

In addition, educational level did not significantly moderate the effects of multiple representations. While prior knowledge is widely regarded as a factor that influences the impact of representations (Kalyuga et al., 2000), increased knowledge due to educational advancement fails to significantly affect the effectiveness of multiple representations. This finding is in agreement with earlier meta-analytic findings (Noetel et al., 2022).

Finally, further refinement in the classification of mathematical domains is needed to better capture domain-specific effects. The systematic review indicated that there were variations in the design and effectiveness of representations across different mathematical domains. However, the moderation analysis showed no significant moderating effect of mathematical domain on the impact of multiple representations. This is probably because the categorization of mathematical knowledge is not sufficiently detailed. For example, procedural and conceptual knowledge in the domain of statistics and probability were not distinguished. Heterogeneity exists in the effects of multiple representations across different mathematical knowledge, but further in-depth investigations are needed to draw concrete conclusions. Particular attention should be paid to the cognitive effects of the use of specific representations across different learning tasks. Notably, although we found that multiple representations significantly benefit learning in the domain of propositional logic, this result is derived from only one study. Consequently, it should be considered preliminary and interpreted with caution.

Domain-specific Effect of Multiple Representations in Mathematics

The meta-analysis reported significant heterogeneity among the included studies, with within-study and between-study variance accounting for 72.4% of the total variance. Consistent with the meta-analysis, the systematic review further revealed that divergent results were observed across six specific representation comparisons. We argue that this heterogeneity is not accidental but rather reflects the inherent complexity of learning processes when multiple representations are employed. Learning through representations involves a two-way interaction between the internalization of external representations and the externalization of internal representations (Pape & Tchoshanov, 2001). The cognitive activities occurring within this interaction may therefore be central to understanding how students engage with representations. However, the interaction between external and internal representations has rarely been explored or discussed in depth (Pande & Chandrasekharan, 2017). To explain the heterogeneity of multiple representation effects, it is essential to go beyond focusing on the power of external representations to examining how they are coupled with internal representations (Chandrasekharan & Nersessian, 2015). This aligns with the perspective of Flegr et al. (2023), highlighting the need to take into account both representation characteristics and the characteristics of learners. We observed domain-specific variations in the commonly used representations, which can be attributed to the fact that different mathematical domains emphasize distinct core topics. Algebraic expressions of functions were prevalent in the domain of functions. In the algebra domain, algebraic and numerical representations were commonly used, as it abstracts numerical values and emphasizes transforming expressions and solving equations. The number and quantity domain and the statistics and probability domain frequently employed algebraic, graphic, or verbal representations. Only one study examined the domain of geometry, and both the control and experimental groups used diagrams because geometry focuses on the quantitative and spatial properties of figures and solids. Even when different mathematical domains employ the same type of representation, the specific forms of these representations can differ significantly. For example, different mathematical domains tended to use distinct forms of graphic representations: number lines and percentage bars in the algebra domain and number and quantity domain; triangles and other geometric figures in the geometry domain; function graphs in the function domain; and tree diagrams or Venn diagrams in the statistics and probability domain. These patterns may reflect differences in the nature of the content and expose students to domain-specific external representations, thereby shaping students' internal representations in a similarly domain-specific manner. Consequently, it is essential to consider both the domain-specific nature of the knowledge itself and students' internal representations when designing representations. This is consistent with Fiorella (2023), who emphasized that multimedia design should fit the knowledge domain and focus on the conceptual core, while at the same time adapting to learner characteristics to better provide differentiated scaffolding. Although students' internal representations are not directly observable, they can be inferred from their interactions with external representations (Malinowski, 2002). We observed that the impact of specific representations may vary across different mathematical domains. The systematic review indicated that the overall advantage of multiple representations over single representations was more pronounced in the domains of geometry, algebra, and functions. This variation was especially clear in certain types of representations. For example, comparison of the combination of graphic and algebraic representations and algebraic representations alone showed that the inclusion of graphical representations facilitated learning in the domains of propositional logic, algebra, and number and quantity; however, it yielded the opposite effect in the domain of statistics and probability. This can be attributed to the nature of the concepts in the domain of statistics and probability and the cognitive demands it imposes during problem solving. The domain of statistics and probability presents considerable challenges for students with limited prior experience. Graphic representations such as tree diagrams and Venn diagrams are often regarded as effective tools for teaching statistics and probability (Greer, 2001). However, such perception may be the result of experts' mistaken belief that strategies effective for themselves are equally effective for novices. In fact, the effectiveness of representation design requires taking into consideration learners' individual representational competence (Hartmann et al., 2023). Experts benefit from graphic representations because they help trigger information stored in long-term memory, thereby reducing working memory load. Unlike other mathematical domains, problem solving in probability requires sequential reasoning. But graphic representations often lack guidance for sequential reasoning steps. For novices, the reasoning steps depicted in graphic representations are often implicit and difficult to infer. It may instead increase students' cognitive load when processing information (Kolloffel, 2000). Under these circumstances, algebraic representations that convey meaning explicitly may be more effective. Differences in internal representations between experts and students can lead to the inappropriate use of representation strategies, resulting in negative learning outcomes. However, this does not imply that multiple representation strategies or graphic representations should be avoided in the domain of statistics and probability. Once students develop internal representations similar to those of experts, graphic representations may also begin to support their learning. Overall, the effectiveness of representation strategies depends on the productive interaction between external and internal representations. The influence of intrinsic knowledge features on both the design of external representations and the formation of internal representations should not be overlooked. This finding aligns with those of numerous empirical studies that indicate that the findings obtained in certain mathematical domains may not be applicable to other domains of mathematics (Beitzel et al., 2011; Flores et al., 2019). Given the limited number of included studies, we did not analyze the heterogeneity of representation effects that have not been examined by multiple studies. The second-generation study is needed to address this issue and provide more evidence.

There are also commonalities in the design and effectiveness of representations across different mathematical domains. A consistent tendency is that all domains preferred the use of heterogeneous representations. Eye-tracking data suggest that students make more frequent gaze shifts when learning with heterogeneous representations (Ott et al., 2018), indicating a more active integration of representations (Mason et al., 2013). In contrast, students often focus on only one of them when learning with homogeneous representations, limiting the benefit of multiple representations. According to Leutner and Biele (2025), it is crucial to support students in constructing a coherent mental representation by integrating verbal and visual models. This is consistent with the integrated model of text and picture comprehension (ITPC).

Current Research Features and Suggestions

Researchers have shown sustained interest in representations. Representations emerged as a major theme in all included studies, whereas content and participants received comparatively less attention. Given the heterogeneity in the effectiveness of representations across different mathematical content, it is necessary to consider mathematical content when discussing representation strategies. It is essential to explain the cognitive reasons why specific multiple-representation designs facilitate learning of particular content; otherwise, the ability to generalize conclusions arbitrarily to other domains carries potential risks. However, only two studies focused on content as a major theme, and six regarded it as a minor theme. Among these, only five studies integrated specific mathematical knowledge content into the analysis. This is not an encouraging trend. Ainsworth (2008) summarized the progression of multimedia learning from first-generation to second-generation research and noted that the focus of research should shift from evaluating whether multimedia is an effective learning environment to exploring why specific multimedia examples benefit particular learners under certain conditions. This perspective is transferable to research on multiple representations. However, current studies still lack sufficient attention to the specific mathematical content in specific learning situations. This highlights the need for future studies to shift from asking whether particular representations are effective to understanding why they facilitate the learning of specific content (Goldman, 2003).

Ainsworth (2008) noted that some research on multimedia learning lacks sufficient detail in describing the design of multimedia environments and representations. This is consistent with our results: two studies specified only the types of representations, and six provided examples that were insufficient to convey the comprehensive design of the representations. Chen et al. (2018) highlighted that the interactivity between representational elements also influences cognitive load. Therefore, simply emphasizing the number or type of representations is insufficient. This limits the applicability of the research findings for both deepening theory and guiding instructional practice. Ainsworth (2008) showed that deepening theoretical foundations to better understand how representations influence learning is essential (Klein, 2003). Thus, future research must provide more detailed descriptions of representation design.

Limitationsations

Although our study yields valuable findings, it has limitations that warrant attention. First, only between-subjects studies were included, while within-subjects and other designs were excluded. This selection criterion may have resulted in the omission of meaningful studies and important results. In addition, we did not search for grey literature or trial registries. Some unpublished or ongoing studies may have been missed, which may introduce publication bias. Second, given that only 15 studies were included in the meta-analysis, some categories contained a small number of studies for moderator analyses. For example, only two studies involved four types of representations in the experimental group, which may have not sufficiently reflected the general trend of this category and may have limited the comprehensiveness of the results. Third, no formal study protocol was developed for this review. It constitutes a potential limitation in risk-of-bias interpretation. Additionally, the categorization of mathematical content in our study was relatively coarse, which may have led to the omission of subtle effects. Finally, although we intended to investigate potential moderators such as information redundancy, a large proportion of effect sizes did not report these variables, preventing such analyses. Future studies should explore these moderators.

Conclusion

Research on representations is undergoing a transition from initial exploration to more in-depth investigations. This study investigated the effect of multiple representations on students' mathematical achievement to provide insights for future studies. Three-level meta-analysis revealed a positive effect of multiple representations on students' mathematics achievement. However, this positive effect did not reach the anticipated level and exhibited variability across contexts. We underscore that confidence in the effectiveness of multiple representation strategies should be maintained, but careful attention must be paid to how representations are designed and implemented in practice. In particular, attention should be given to clarify how specific combinations of representation types affect students' cognitive processes. Our synthesis highlights that the design of representations must consider the specific context and avoid unquestioningly adopting popular strategies such as visualization strategies. In addition, the design and effectiveness of multiple representations differed in magnitude across various mathematical domains. It is essential to consider the intrinsic nature of mathematical content when designing representations, as it may shape students' internal representations in a similarly domain-specific manner. However, we found that the previous studies paid limited attention to knowledge content. Many studies do not incorporate specific mathematical knowledge content when analyzing the effectiveness of representations. Future studies should emphasize the integration of specific knowledge content to better investigate how representations influence learners' understanding of specific knowledge. Additionally, many studies lacked detailed descriptions of representation design and only specified representation types or provided basic examples, which limits the generalizability of their findings. Future research should offer more explicit and detailed descriptions of representation design.

Appendix 1

Author(year)	Mode(C)	Content	Hedge's g
G&N&A&V			
Arefaine et al. (2022)	Algebraic	Functions	0.1408
Arefaine et al. (2022)	Algebraic	Functions	-0.2691
Brenner et al. (1997)	Algebraic	Functions	0.5339
Brenner et al. (1997)	Algebraic	Functions	0.2268
Brenner et al. (1997)	Algebraic	Functions	-0.2076
G&A&V			
Ngu et al. (2014)	Verbal	Number and Quantity	0.7673
Ngu et al. (2014)	Verbal	Number and Quantity	0.9666
Otta et al. (2018)	Algebraic	Propositional logic	1.1642
Otta et al. (2018)	Verbal	Propositional logic	0.7388
Moreno&Mayer (1999)	Algebraic	Number and Quantity	0.3434
Moreno&Mayer (1999)	Algebraic	Number and Quantity	-0.0876
Moreno&Mayer (1999)	Algebraic	Number and Quantity	0.6221
Moreno&Mayer (1999)	Algebraic	Number and Quantity	0.7006
Moreno&Mayer (1999)	Algebraic	Number and Quantity	0.3081
Moreno&Mayer (1999)	Algebraic	Number and Quantity	0.3676
Moreno&Mayer (1999)	Algebraic	Number and Quantity	0.5429
Moreno&Mayer (1999)	Algebraic	Number and Quantity	-0.0682
G&N&A			
Floresa et al. (2018)	Algebraic	Number and Quantity	-0.4276
Daryaee et al. (2018)	Algebraic	Algebra	1.5806
Kerry (1995)	Algebraic	Functions	0.4601
Kerry (1995)	Algebraic	Functions	0.4407
Kerry (1995)	Algebraic	Functions	0.3866
Kerry (1995)	Algebraic	Functions	0.8054
Kerry (1995)	Algebraic	Functions	0.477
Kerry (1995)	Algebraic	Functions	0.4446
Kerry (1995)	Algebraic	Functions	0.4143
Kerry (1995)	Algebraic	Functions	0.9212
G&A			
Berthold&Renkl (2009)	Graphic	Statistics and Probability	0.2241
Berthold&Renkl (2009)	Graphic	Statistics and Probability	0.1785
Berthold&Renkl (2009)	Algebraic	Statistics and Probability	-0.1683
Berthold&Renkl (2009)	Algebraic	Statistics and Probability	-0.0489
Berthold&Renkl (2009)	Graphic	Statistics and Probability	0.5822
Berthold&Renkl (2009)	Graphic	Statistics and Probability	0.0912
Berthold&Renkl (2009)	Algebraic	Statistics and Probability	0.2426
Berthold&Renkl (2009)	Algebraic	Statistics and Probability	-0.1509
Berthold&Renkl (2009)	Graphic	Statistics and Probability	0.4354
Berthold&Renkl (2009)	Graphic	Statistics and Probability	0.0573
Berthold&Renkl (2009)	Algebraic	Statistics and Probability	0.1541
Berthold&Renkl (2009)	Algebraic	Statistics and Probability	-0.2515
Berthold&Renkl (2009)	Graphic	Statistics and Probability	0.9811
Berthold&Renkl (2009)	Graphic	Statistics and Probability	0.7527
Berthold&Renkl (2009)	Algebraic	Statistics and Probability	0.6859
Berthold&Renkl (2009)	Algebraic	Statistics and Probability	0.3828
Beitzel et al. (2011)	Algebraic	Statistics and Probability	-0.6624
Beitzel et al. (2011)	Algebraic	Statistics and Probability	-0.5441
Beitzel et al. (2011)	Algebraic	Statistics and Probability	-0.9314
Beitzel et al. (2011)	Algebraic	Statistics and Probability	-0.5328
Otta et al. (2018)	Algebraic	Propositional logic	0.3147
Otta et al. (2018)	Verbal	Propositional logic	-0.1205
Kolloffel et al. (2008)	Graphic	Statistics and Probability	0.3166
Kolloffel et al. (2008)	Algebraic	Statistics and Probability	0
Kolloffel et al. (2008)	Verbal	Statistics and Probability	-0.0511
Kolloffel et al. (2008)	Graphic	Statistics and Probability	0.3797
Kolloffel et al. (2008)	Algebraic	Statistics and Probability	-0.0928
Kolloffel et al. (2008)	Verbal	Statistics and Probability	-0.0683

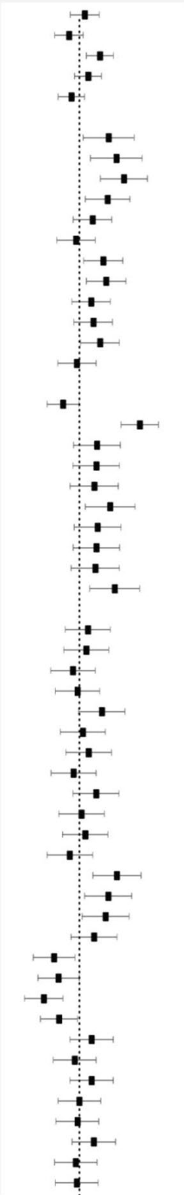

Fig. 4 Forest plot

Kolloffel et al. (2008)	Graphic	Statistics and Probability	0.1204
Kolloffel et al. (2008)	Algebraic	Statistics and Probability	-0.2485
Kolloffel et al. (2008)	Verbal	Statistics and Probability	-0.2657
Kolloffel et al. (2008)	Graphic	Statistics and Probability	0.1613
Kolloffel et al. (2008)	Algebraic	Statistics and Probability	0.2749
Kolloffel et al. (2008)	Verbal	Statistics and Probability	-0.2964
Kolloffel et al. (2008)	Graphic	Statistics and Probability	0.2647
Kolloffel et al. (2008)	Algebraic	Statistics and Probability	-0.0675
Kolloffel et al. (2008)	Verbal	Statistics and Probability	-0.0831
Kathleen (2001)	Algebraic	Algebra	0.0687
Kathleen (2001)	Algebraic	Algebra	0.5142
A&V			
Otta et al. (2018)	Algebraic	Propositional logic	1.1837
Otta et al. (2018)	Verbal	Propositional logic	0.7169
Kolloffel et al. (2008)	Graphic	Statistics and Probability	0.6979
Kolloffel et al. (2008)	Algebraic	Statistics and Probability	0.3949
Kolloffel et al. (2008)	Verbal	Statistics and Probability	0.3588
Kolloffel et al. (2008)	Graphic	Statistics and Probability	0.8098
Kolloffel et al. (2008)	Algebraic	Statistics and Probability	0.3282
Kolloffel et al. (2008)	Verbal	Statistics and Probability	0.3374
Kolloffel et al. (2008)	Graphic	Statistics and Probability	0.7945
Kolloffel et al. (2008)	Algebraic	Statistics and Probability	0.4263
Kolloffel et al. (2008)	Verbal	Statistics and Probability	0.4933
Kolloffel et al. (2008)	Graphic	Statistics and Probability	0.2319
Kolloffel et al. (2008)	Algebraic	Statistics and Probability	0.3454
Kolloffel et al. (2008)	Verbal	Statistics and Probability	0.3702
Kolloffel et al. (2008)	Graphic	Statistics and Probability	0.9316
Kolloffel et al. (2008)	Algebraic	Statistics and Probability	0.517
Kolloffel et al. (2008)	Verbal	Statistics and Probability	0.5556
G&V			
Ngu et al. (2014)	Verbal	Number and Quantity	-0.6913
Ngu et al. (2014)	Verbal	Number and Quantity	-0.2461
K. Reed et al. (2012)	Numeric	Algebra	0.0262
K. Reed et al. (2012)	Numeric	Algebra	-0.0106
Liang&She (2021)	Graphic	Geometry	0.2724
Liang&She (2021)	Graphic	Geometry	0.1156
Liang&She (2021)	Graphic	Geometry	0.2406
Liang&She (2021)	Graphic	Geometry	0.2043
Liang&She (2021)	Graphic	Geometry	0.3769
Liang&She (2021)	Graphic	Geometry	0.351
Liang&She (2021)	Graphic	Geometry	0.8182
Liang&She (2021)	Graphic	Geometry	0.1835
Liang&She (2021)	Graphic	Geometry	0.5418
Liang&She (2021)	Graphic	Geometry	1.0787
Otta et al. (2018)	Algebraic	Propositional logic	1.7679
Otta et al. (2018)	Verbal	Propositional logic	1.2869
N&A			
Beitzel et al. (2011)	Algebraic	Statistics and Probability	-0.3615
Beitzel et al. (2011)	Algebraic	Statistics and Probability	-0.2066
Beitzel et al. (2011)	Algebraic	Statistics and Probability	-0.0955
Beitzel et al. (2011)	Algebraic	Statistics and Probability	-0.2266
G&N			
Ainsworth et al. (2002)	Graphic	Number and Quantity	0.4698
Ainsworth et al. (2002)	Graphic	Number and Quantity	0.4352
N&V			
K. Reed et al. (2012)	Numeric	Algebra	0.3634

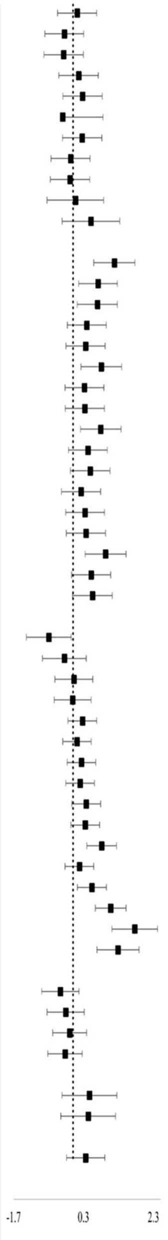

Fig. 4 (continued)

Appendix 2

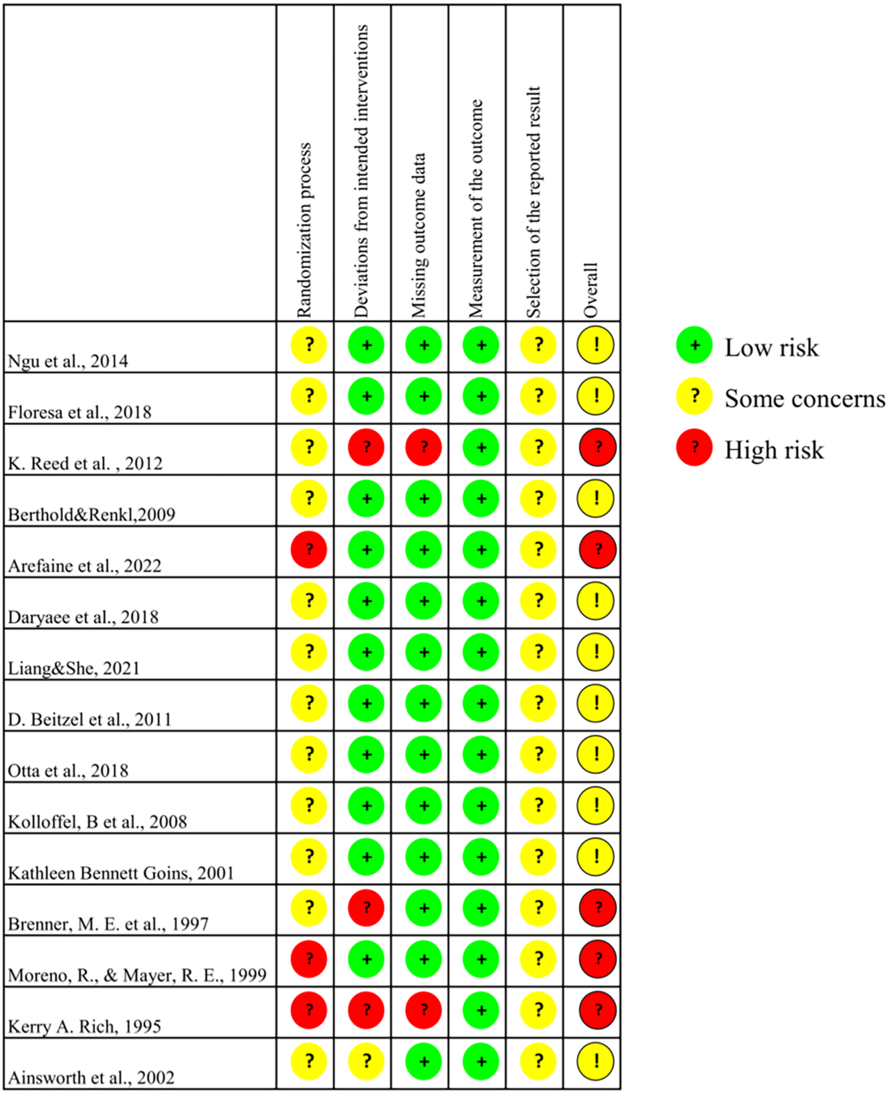

Fig. 5 Result of risk of bias

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10648-025-10109-0.

Author Contribution Guo Kan, Liu Yuxin, and Ji Zhongtian contributed to the conceptualization of the study. Liu Yuxin took the lead in writing the original draft and, together with Ji Zhongtian, was responsible for data screening and coding. All authors reviewed and edited the final manuscript for publication.

Funding This research was supported by Beijing Education Sciences Planning Foundation [CEDA24014].

Data Availability Data will be made available on request.

Declarations

This manuscript is original, has not been published elsewhere, and is not under consideration by any other journal. This research complies with ethical publication standards.

Conflicts of interest The authors declare no conflicts of interest.

References

Studies preceded by an asterisk were included in the systematic review

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*(3), 183–198. https://doi.org/10.1016/j.learninstruc.2006.03.001

Ainsworth, S. (2008). How should we evaluate multimedia learning environments?. In *Understanding multimedia documents* (pp. 249–265). Boston, MA: Springer US. https://doi.org/10.1007/978-0-387-73337-1_13

Ainsworth, S. (2018). Multi-modal, multi-source reading: A multi-representational reader's perspective. *Learning and Instruction, 57*, 71–75. https://doi.org/10.1016/j.learninstruc.2018.01.014

Ainsworth, S., Bibby, P., & Wood, D. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *The Journal of the Learning Sciences, 11*(1), 25–61.

Albus, P., Vogt, A., & Seufert, T. (2021). Signaling in virtual reality influences learning outcome and cognitive load. *Computers & Education, 166*, 104154. https://doi.org/10.1016/j.compedu.2021.104154

Aleksandrov, A. D., Kolmogorov, A. N., & Lavrent'ev, M. A. (Eds.) (1963). *Mathematics: Its content, methods, and meaning* (S. H. Gould & T. Bartha, Trans.; Vol. 1). MIT Press.

*Amirbostaghi, G., Asadi, M., Mardanbeigi, M. R., Azhini, M., & Shahvarani, A. (2021). *The impact of words in mathematics education – Case of symbols*.

Arefaine, N., Michael, K., & Assefa, S. (2022). Effect of multiple representations on students' performance on interpretations and techniques of representation in calculus. *Jurnal Pendidikan Matematika, 16*(3), 351–372. https://doi.org/10.22342/jpm.16.3.18291.351-372

Arneson, J. B., & Offerdahl, E. G. (2023). Assessing the load: Effects of visual representation and task features on exam performance in undergraduate molecular life sciences. *Research in Science Education, 53*(2), 319–335. https://doi.org/10.1007/s11165-022-10057-7

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *Tutorials in Quantitative Methods for Psychology, 12*(3), 154–174. https://doi.org/10.20982/tqmp.12.3.p154

Beitzel, B. D., Staley, R. K., & DuBois, N. F. (2011). The (in)effectiveness of visual representations as an aid to solving probability word problems. *Effective Education, 3*(1), 11–22. https://doi.org/10.1080/19415532.2011.604256

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series b: Statistical Methodology, 57*(1), 289–300.

Ben-Zvi, D., & Makar, K. (Eds.). (2016). *The teaching and learning of statistics: International perspectives*. Springer International Publishing. https://doi.org/10.1007/978-3-319-23470-0

*Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology*, *101*(1), 70–87. https://doi.org/10.1037/a0013247

Birgin, O., & Eryılmaz, E. (2025). Investigation of seventh-grade students' performance in translating among multiple representations of fractions. *Thinking Skills and Creativity, 57*, 101809. https://doi.org/10.1016/j.tsc.2025.101809

Borenstein, M. (2022). Comprehensive meta-analysis software. In *Systematic Reviews in Health Research* (pp. 535–548). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119099369.ch27

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. wiley. https://doi.org/10.1002/9780470743386

Brenner, M. E., Mayer, R. E., Moseley, B., Brar, T., Durán, R., Reed, B. S., & Webb, D. (1997). Learning by understanding: The role of multiple representations in learning algebra. *American Educational Research Journal, 34*(4), 663–689. https://doi.org/10.2307/1163353

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology, 62*(2), 233–246. https://doi.org/10.1111/j.2044-8279.1992.tb01017.x

Chandrasekharan, S., & Nersessian, N. J. (2015). Building cognition: The construction of computational representations for scientific discovery. *Cognitive Science, 39*(8), 1727–1763. https://doi.org/10.1111/cogs.12203

Chen, O., Castro-Alonso, J. C., Paas, F., & Sweller, J. (2018). Extending cognitive load theory to incorporate working memory resource depletion: Evidence from the spacing effect. *Educational Psychology Review, 30*(2), 483–501. https://doi.org/10.1007/s10648-017-9426-2

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

Common core state standards for mathematics. (2011). Common Core State Standards Initiative.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education, 90*(6), 1073–1091. https://doi.org/10.1002/sce.20164

Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79*(4), 347–362. https://doi.org/10.1037/0022-0663.79.4.347

Cuoco, A. (2001). Mathematics for teaching. *Bollettino Della Unione Matematica Italiana. Serie VIII. Sezione A. La Matematica Nella Società e Nella Cultura*, *3*.

Daryaee, A., Shahvarani, A., Tehranian, A., Hosseinzadeh Lotfi, F., & Rostamy-Malkhalifeh, M. (2018). Executable functions of the representations in learning the algebraic concepts. *PNA. Revista De Investigación En Didáctica De La Matemática, 13*(1), 1–18. https://doi.org/10.30827/pna.v13i1.6903

de Jong, T., & van der Meij, J. (2012). Learning with multiple representations. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 2026–2029). Springer US. https://doi.org/10.1007/978-1-4419-1428-6_485

Demir, M., & Zengin, Y. (2025). Investigation of generalisation processes of secondary school students using multiple representations in a pattern task. *International Journal of Mathematical Education in Science and Technology, 56*(3), 417–444. https://doi.org/10.1080/0020739X.2023.2240795

Duval, R. (1999). *Representation, vision and visualization: Cognitive functions in mathematical thinking. Basic Issues for Learning.*

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

Çeken, B., & Taşkın, N. (2022). Multimedia learning principles in different learning environments: A systematic review. *Smart Learning Environments, 9*(1), 19. https://doi.org/10.1186/s40561-022-00200-2

Fiorella, L. (2023). Making sense of generative learning. *Educational Psychology Review, 35*(2), 50. https://doi.org/10.1007/s10648-023-09769-7

Flegr, S., Kuhn, J., & Scheiter, K. (2023). When the whole is greater than the sum of its parts: Combining real and virtual experiments in science education. *Computers & Education, 197*, 104745. https://doi.org/10.1016/j.compedu.2023.104745

Flores, R., Inan, F. A., Han, S., & Koontz, E. (2019). Comparison of algorithmic and multiple-representation integrated instruction for teaching fractions, decimals, and percent. *Investigations in Mathematics Learning, 11*(4), 231–244. https://doi.org/10.1080/19477503.2018.1461050

Ford, S. J. (2008). *The effect of graphing calculators and a three -core representation curriculum on college students' learning of exponential and logarithmic functions* [Ph.D.].

Gao, S., Assink, M., Cipriani, A., & Lin, K. (2017). Associations between rejection sensitivity and mental health outcomes: A meta-analytic review. *Clinical Psychology Review, 57*, 59–74. https://doi.org/10.1016/j.cpr.2017.08.007

Gavilán-Izquierdo, J. M., & Gallego-Sánchez, I. (2025). Developing TPACK through task design: Exploring the use of multiple modes of representation and the promotion of mathematical processes. *Journal of Education for Teaching, 51*(1), 28–45. https://doi.org/10.1080/02607476.2024.2422507

*Goins, K. B. (2001). *Comparing the effects of visual and algebra tile manipulative methods on student skill and understanding of polynomial multiplication* [Ph.D.].

Goldin, G., & Shteingold, N. (2001). Systems of representations and the development of mathematical concepts. In *The roles of representation in school mathematics* (pp. 1–23).

Goldman, S. R. (2003). Learning in complex domains: When and why do multiple representations help? *Learning and Instruction, 13*(2), 239–244. https://doi.org/10.1016/S0959-4752(02)00023-3

Greer, B. (2001). Understanding probabilistic thinking: The legacy of Efraim Fischbein. *Educational Studies in Mathematics, 45*(1), 15–33. https://doi.org/10.1023/A:1013801623755

*Han, S., Flores, R., Inan, F. A., & Koontz, E. (2016). The use of traditional algorithmic versus instruction with multiple representations: Impact on pre-algebra students' achievement with fractions, decimals, and percent. *School Mathematics*, 18(2), 257-275.

Hao, X., Xu, Z., Guo, M., Hu, Y., & Geng, F. (2023). The effect of embedded structures on cognitive load for novice learners during block-based code comprehension. *International Journal of STEM Education, 10*(1), 42. https://doi.org/10.1186/s40594-023-00432-9

Hartmann, C., Orli-Idrissi, Y., Pflieger, L. C. J., & Bannert, M. (2023). Imagine & immerse yourself: Does visuospatial imagery moderate learning in virtual reality? *Computers & Education, 207*, 104909. https://doi.org/10.1016/j.compedu.2023.104909

Hattie, J. (2008). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. *Routledge*. https://doi.org/10.4324/9780203887332

Hattie, J. (2023). Visible learning: The sequel: A synthesis of over 2,100 meta-analyses relating to achievement. *Routledge*. https://doi.org/10.4324/9781003380542

He, S., & Xin, Y. P. (2025). Design knowledge scaffolds to facilitate students' probabilistic thinking skills for solving classical probability problems: An exploratory study. *Thinking Skills and Creativity*. https://doi.org/10.1016/j.tsc.2025.101928

Höffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17*(6), 722–738. https://doi.org/10.1016/j.learninstruc.2007.09.013

Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 65–97). Macmillan Publishing Co, Inc.

Hox, J., Moerbeek, M., & Schoot, R. van de. (2017). *Multilevel analysis: Techniques and applications third edition* (3rd ed.). Routledge https://doi.org/10.4324/9781315650982

Hsu, S.-K., & Hsu, Y. (2025). Supporting young learners in learning geometric area concepts through static versus dynamic representation and imagination strategies. *International Journal of Science and Mathematics Education, 23*(2), 441–459. https://doi.org/10.1007/s10763-024-10481-3

Hu, L., Chen, G., Li, P., & Huang, J. (2021). Multimedia effect in problem solving: A meta-analysis. *Educational Psychology Review, 33*(4), 1717–1747. https://doi.org/10.1007/s10648-021-09610-z

Jain, V., & Mitra, A. (2025). The role of multiple representations in enhancing statistical thinking in secondary education: A case study approach. In S. D. Anastasiadou & L. Seremeti (Eds.), *Modes of Representation in Developing Statistical Thinking in Education* (pp. 137–172). IGI Global. https://doi.org/10.4018/979-8-3693-9934-7.ch008

Ji, Z., Guo, K., & Song, S. (2024). Effects of dynamic mathematical software on students' performance: A three-level meta-analysis. *Journal of Educational Computing Research, 62*(4), 1035–1060. https://doi.org/10.1177/07356331241226594

Jong, T. de, Ainsworth, S., Dobson, M., Hulst, A. van der, Levonen, J., Reimann, P., Sime, J. A., Someren, M. van, Spada, H., & Swaak, J. (1998). Acquiring knowledge in science and mathematics: The use of multiple representations in technology based learning environments. In *Learning with multiple representations* (pp. 9–40). Pergamon Press.

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology, 92*(1), 126–136. https://doi.org/10.1037/0022-0663.92.1.126

Kalyuga, S., Chandler, P., & Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors, 46*(3), 567–581. https://doi.org/10.1518/hfes.46.3.567.50405

Klein, P. D. (2003). Rethinking the multiplicity of cognitive resources and curricular representations: Alternatives to "learning styles" and "multiple intelligences. *Journal of Curriculum Studies*. https://doi.org/10.1080/00220270210141891

Koedinger, K., & Nathan, M. (2004). The Real Story Behind Story Problems: Effects of Representations on Quantitative Reasoning. *Journal of The Learning Sciences, 13,* 129–164. https://doi.org/10.1207/s15327809jls1302_1

Kokkonen, T., & Schalk, L. (2021). One instructional sequence fits all? A conceptual analysis of the applicability of concreteness fading in mathematics, physics, chemistry, and biology education. *Educational Psychology Review, 33*(3), 797–821. https://doi.org/10.1007/s10648-020-09581-7

*Kolloffel, B. (2000). *The role of external representations in simulation-based inquiry learning* [PhD].

Lesh, R., Post, T. R., & Behr, M. (1987). Representations and translations among representations in mathematics learning and problem solving. In *Problems of representations in the teaching and learning of mathematics* (pp. 33–40). Lawrence Erlbaum.

Leutner, D., & Biele, J. (2025). Without integration, everything is nothing: A meta-analysis of the effectiveness of instructional support for drawing-to-learn. *Educational Psychology Review, 37*(4), 93. https://doi.org/10.1007/s10648-025-10067-7

Levin, J. R., & Lesgold, A. M. (1978). On pictures in prose. E*ducational communication and technology, 26*(3), 233–243.

*Liang, C.-P., & She, H.-C. (2023). Investigate the effectiveness of single and multiple representational scaffolds on mathematics problem solving: Evidence from eye movements. *Interactive Learning Environments*, *31*(6), 3882–3897. https://doi.org/10.1080/10494820.2021.1943692

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (pp. ix, 247). Sage Publications, Inc.

Lu, L., & Ran, G. (2024). The association between trait mindfulness and sleep problems: A three-level meta-analysis. *Journal of Health Psychology*. https://doi.org/10.1177/13591053241253483

Mainali, B. (2020). Representation in teaching and learning mathematics. *International Journal of Education in Mathematics, Science and Technology, 9*(1), 1–21. https://doi.org/10.46328/ijemst.1111

Malinowski, L. T. (2002). The roles of representation in school mathematics. *Research and Teaching in Developmental Education, 19*(1), 70–72.

Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education, 60*(1), 95–109. https://doi.org/10.1016/j.compedu.2012.07.011

Mayer, R. E. (2009). *Multimedia learning, 2nd ed* (pp. xiii, 304). Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

Mayer, R. E., & Moreno, R. (1998). A cognitive theory of multimedia learning: Implications for design principles. *Journal of Educational Psychology, 91*(2), 358–368.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

Moreno, R., & Mayer, R. E. (1999). Multimedia-supported metaphors for meaning making in mathematics. *Cognition and Instruction, 17*(3), 215–248.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* National Council of Teachers of Mathematics.

Ngu, B. H., Yeung, A. S., & Tobias, S. (2014). Cognitive load in percentage change problems: Unitary, pictorial, and equation approaches to instruction. *Instructional Science, 42*(5), 685–713. https://doi.org/10.1007/s11251-014-9309-6

Ni, S., Jiang, Z., & Chiang, F. (2025). Visual attention to different types of graphical representations in elementary school mathematics textbooks: An eye-movement-based study. *Stem Education, 5*(3), 448–472. https://doi.org/10.3934/steme.2025022

Noetel, M., Griffith, S., Delaney, O., Harris, N. R., Sanders, T., Parker, P., del Pozo Cruz, B., & Lonsdale, C. (2022). Multimedia design for learning: An overview of reviews with meta-meta-analysis. *Review of Educational Research, 92*(3), 413–454. https://doi.org/10.3102/00346543211052329

*Ott, N., Brünken, R., Vogel, M., & Malone, S. (2018). Multiple symbolic representations: The combination of formula and text supports problem solving in the mathematical field of propositional logic. *Learning and Instruction*, *58*, 88–105. https://doi.org/10.1016/j.learninstruc.2018.04.010

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews, 5*(1), Article 210. https://doi.org/10.1186/s13643-016-0384-4

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 (Clinical Research Ed.)*, *372*, n71. https://doi.org/10.1136/bmj.n71

Paivio, A. (1990). *Mental representations: A dual coding approach* (1st ed., Vol. 9). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195066661.001.0001

Pande, P. (2021). Learning and expertise with scientific external representations: An embodied and extended cognition model. *Phenomenology and the Cognitive Sciences, 20*(3), 463–482. https://doi.org/10.1007/s11097-020-09686-y

Pande, P., & Chandrasekharan, S. (2017). Representational competence: Towards a distributed and embodied cognition account. *Studies in Science Education, 53*(1), 1–43. https://doi.org/10.1080/03057267.2017.1248627

Pape, S. J., & Tchoshanov, M. A. (2001). The role of representation(s) in developing mathematical understanding. *Theory into Practice, 40*(2), 118–127. https://doi.org/10.1207/s15430421tip4002_6

Piaget, J. (1970). *Science of education and the psychology of the child* (D. Coltman, Trans.). Orion Press.

Pinto, E., & Cañadas, M. C. (2021). Generalizations of third and fifth graders within a functional approach to early algebra. *Mathematics Education Research Journal, 33*(1), 113–134. https://doi.org/10.1007/s13394-019-00300-2

Polanin, J. (2013). Addressing the issue of meta-analysis multiplicity in education and psychology. *Dissertations*. https://ecommons.luc.edu/luc_diss/539

Porzio, D. (1999). Effects of differing emphases in the use of multiple representations and technology on students' understanding of calculus concepts. *Focus on Learning Problems in Mathematics, 21*(3), 1–29.

Qin, T., & Wu, C.-J. (2025). Graphic organizers' optimized design through segmenting and signaling principle: Based on generative learning theory. *Computers & Education, 239*, 105416. https://doi.org/10.1016/j.compedu.2025.105416

*Reed, S. K., Corbett, A., Hoffman, B., Wagner, A., & MacLaren, B. (2013). Effect of worked examples and cognitive tutor training on constructing equations. *Instructional Science*, *41*(1), 1–24. https://doi.org/10.1007/s11251-012-9205-x

Rexigel, E., Kuhn, J., Becker, S., & Malone, S. (2024). The more the better? A systematic review and meta-analysis of the benefits of more than two external representations in STEM education. *Educational Psychology Review, 36*(4), 124. https://doi.org/10.1007/s10648-024-09958-y

*Rich, K. A. (1995). *The effect of dynamic linked multiple representations on students' conceptions of and communication of functions and derivatives*[Ph.D.].

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Ruamba, M. Y., Sukestiyarno, Y. L., Rochmad, R., & Asih, T. S. N. (2025). The impact of visual and multimodal representations in mathematics on cognitive load and problem-solving skills. *International Journal of Advanced and Applied Sciences, 12*(4), 164–172. https://doi.org/10.21833/ijaas.2025.04.018

Schneider, S. (2013). The international standard classification of education 2011. *Comparative Social Research, 30*, 365–379. https://doi.org/10.1108/S0195-6310(2013)0000030017

Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review, 14*, 101–120. https://doi.org/10.1023/A:1013136727916

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*(2), 141–156. https://doi.org/10.1016/S0959-4752(02)00017-8

Sierpinska, A. (2013). Understanding in mathematics. *Routledge*. https://doi.org/10.4324/9780203454183

Sokolowski, A. (2018). The effects of using representations in elementary mathematics: Meta-analysis of research. *IAFOR Journal of Education, 6*(3), 129–152. https://doi.org/10.22492/ije.6.3.08

Spiro, R. J., Feltovich, P. J., Feltovich, P. L., Jacobson, M. J., & Coulson, R. L. (1991). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. *Educational Technology, 31*(5), 24–33.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296. https://doi.org/10.1023/A:1022193728205

Van Lieshout, E. C. D. M., & Xenidou-Dervou, I. (2018). Pictorial representations of simple arithmetic problems are not always helpful: A cognitive load perspective. *Educational Studies in Mathematics, 98*(1), 39–55. https://doi.org/10.1007/s10649-017-9802-3

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

Authors and Affiliations

Yuxin Liu[1] · Zhongtian Ji[1] · Kan Guo[1]

✉ Kan Guo
guokan@bnu.edu.cn

[1] School of Mathematical Sciences, Beijing Normal University, No.19, Xinjiekouwai St, Haidian District, Beijing 100875, China