



# Examining the relations between spatial skills and mathematical performance: A meta-analysis

Kinnari Atit<sup>1</sup> · Jason Richard Power<sup>2</sup> · Terri Pigott<sup>3</sup> · Jihyun Lee<sup>4</sup> · Elyssa A. Geer<sup>5,6</sup> · David H. Uttal<sup>7</sup> · Colleen M. Ganley<sup>5</sup> · Sheryl A. Sorby<sup>8</sup>

Accepted: 10 September 2021  
© The Psychonomic Society, Inc. 2021

## Abstract

Much recent research has focused on the relation between spatial skills and mathematical skills, which has resulted in widely reported links between these two skill sets. However, the magnitude of this relation is unclear. Furthermore, it is of interest whether this relation differs in size based on key demographic variables, such as gender and grade-level, and the extent to which this relation can be accounted for by shared domain-general reasoning skills across the two domains. Here we present the results of two meta-analytic studies synthesizing the findings from 45 articles to identify the magnitude of the relation, as well as potential moderators and mediators. The first meta-analysis employed correlated and hierarchical effects meta-regression models to examine the magnitude of the relation between spatial and mathematical skills, and to understand the effect of gender and grade-level on the association. The second meta-analysis employed meta-analytic structural equation modeling to determine how domain-general reasoning skills, specifically fluid reasoning and verbal skills, influence the relationship. Results revealed a positive moderate association between spatial and mathematical skills ( $r = .36$ , robust standard error = 0.035,  $\tau^2 = 0.039$ ). However, no significant effect of gender or grade-level on the association was found. Additionally, we found that fluid reasoning and verbal skills mediated the relationship between spatial skills and mathematical skills, but a unique relation between the spatial and mathematical skills remained. Implications of these findings include advancing our understanding for how to leverage and bolster students' spatial skills as a mechanism for improving mathematical outcomes.

**Keywords** Meta-analysis · Spatial skills · Mathematical skills · Fluid reasoning · Verbal skills

✉ Kinnari Atit  
kinnari.atit@ucr.edu

- <sup>1</sup> School of Education, University of California, Riverside, 1207 Sproul Hall, Riverside, CA 92521, USA
- <sup>2</sup> School of Education, University of Limerick, Limerick, Ireland
- <sup>3</sup> College of Education and Human Development and School of Public Health, Georgia State University, Atlanta, GA, USA
- <sup>4</sup> College of Education, The University of Texas at Austin, Austin, TX, USA
- <sup>5</sup> Department of Psychology, Florida State University, Tallahassee, FL, USA
- <sup>6</sup> Human Development and Family Sciences, Purdue University, West Lafayette, IN, USA
- <sup>7</sup> School of Education and Social Policy, Northwestern University, Evanston, IL, USA
- <sup>8</sup> College of Engineering and Applied Science, University of Cincinnati, Cincinnati, OH, USA

## Introduction

In the past three decades, much research has focused on the relation between spatial skills and mathematical skills (e.g., Casey et al., 1995; Verdine et al., 2017). This effort has resulted in widely reported links between these two sets of skills, as well as a considerable body of evidence highlighting the benefits of well-developed spatial skills in a wide variety of science, technology, engineering, and mathematics (STEM) fields (Atit et al., 2015; Newcombe, 2010; Uttal et al., 2013). But to understand and perhaps leverage these relations within educational contexts, the magnitude of the relation as well as the moderators and mediators of the relation need to be better understood. The meta-analyses presented here aim to estimate the average magnitude of the reported relationship between spatial skills and mathematical achievement, to identify potential moderators and mediators of this relationship, and to inform the design and execution of future research in this area.

The current meta-analysis is particularly pressing in light of widespread underachievement of US students in STEM subjects (Aud et al., 2012), specifically in mathematics (Organization for Economic Development, 2016). A report from the Program of International Student Assessment (PISA) suggests that American students' weaknesses in mathematics are particularly evident when "[r]easoning in a geometric context – requiring authentic reasoning in a planar or spatial geometric context" (Organization for Economic Development, 2012, p.3). This indicates that problems engaging one's spatial skills may be particularly difficult for American students. As a concerted research and educational effort rises to address this need to improve students' mathematical outcomes, it is worth considering the existing wealth of research examining the relation between these two skill sets. A brief overview of relevant research, grouped by common themes, is presented to situate the current meta-analysis within the broader academic and societal context.

### Spatial skills and mathematics

Spatial skills enable us to mentally manipulate, organize, reason about, and make sense of spatial relationships in real and imagined spaces (e.g., Newcombe & Shipley, 2015; Uttal et al., 2013). These skills are commonly employed when completing everyday tasks such as assembling furniture or navigating from one location to another. Understanding spatial skills has been a topic of interest in psychology for much of the last 100 years (e.g., Bethell-Fox & Shepard, 1988; Carroll, 1993; Linn & Petersen, 1985; Shepard & Metzler, 1971). Historically, the interest in spatial skills has roots in the study of mechanical aptitude (Cox, 1928; Paterson et al., 1930) and in defining the factors of intelligence (Carroll, 1993; Thurstone & Thurstone, 1941). Examples of spatial skills commonly examined in studies of mathematical understanding include mental rotation (e.g., Geer et al., 2019) and spatial visualization (e.g., Burte et al., 2017).

Many studies examining the relations between spatial skills and mathematical achievement indicate that the two are significantly correlated for students at all educational levels (e.g., Casey et al., 1995; Delgado & Prieto, 2004; Verdine et al., 2017). For example, mental rotation performance relates to mathematical reasoning skills in elementary students (Geer et al., 2019) and secondary students (Delgado & Prieto, 2004), as well as mathematical aptitude in undergraduates (Casey et al., 1995). Furthermore, some studies report that spatial skills are predictive of students' future mathematics learning even after accounting for indicators of general reasoning skills, such as verbal skills and executive functioning skills (Verdine et al., 2017; Zhang et al., 2014). Yet, the details of the relations between spatial skills, mathematical skills, and factors of general intelligence are not well understood.

In addition to the performance-based evidence of the relation between spatial skills and mathematical achievement, prior research suggests that the positive relation between spatial and mathematical reasoning found in many studies may be based on shared cognitive or neural processes (Gunderson et al., 2012; Mix & Cheng, 2012; see Hawes & Ansari, 2020, for a review). Some researchers assert that number information is mentally represented in spatial formats (Mix, 2019; Mix & Cheng, 2012). For example, work by Gunderson et al. (2012) indicates that quantitative magnitudes are represented in the mind spatially as a mental number line. Further evidence comes from brain-imaging studies showing similar areas of brain activation when individuals process both spatial and numerical information (Hubbard et al., 2005). Lastly, successful interventions often focus on helping students translate mathematical symbols or problem statements into spatial representations, including number lines, diagrams, concrete models, or hand gestures (Novack et al., 2014; Valenzeno et al., 2003).

In line with the success of spatial interventions in supporting students' mathematical reasoning, some researchers believe that the relation between spatial skills and mathematical skills is perhaps due to differences in problem-solving strategies (Delgado & Prieto, 2004). More specifically, studies show that many mathematical problems can be solved by using spatial visualization and/or analytic strategies, and there may be two different styles: one is based on algorithm memorization and automatic application, the other on the visuospatial representation of the problems. This suggests that perhaps those who have better spatial skills would employ the second style of problem solving (i.e., visuospatial strategies), which is often considered to be more efficient (Delgado & Prieto, 2004).

### Domain-general cognitive processes and mathematical achievement

Parallel to the research on spatial skills and mathematical achievement, researchers interested in mathematics learning and outcomes have also focused their efforts on examining the role of domain-general cognitive processes in students' mathematical achievement and understanding (e.g., Hawes et al., 2019; Raghobar et al., 2010; Taub et al., 2008). Domain-general cognitive processes that have been found to have a strong positive relationship with mathematical achievement include executive functions, such as working memory (Raghobar et al., 2010) and fluid reasoning (Green et al., 2017; Taub et al., 2008). Working memory refers to a mental workspace that is involved in controlling, regulating, and actively maintaining relevant information to accomplish complex cognitive tasks (Miyake & Shah, 1999). Studies have found that working memory is fundamental for mathematics learning and performance in school-aged children (Berg,

2008; McKenzie et al., 2003) and adults (Imbo et al., 2007; Seitz & Schumann-Hengsteler, 2002). For instance, Berg (2008) examined the contribution of working memory in third to sixth grade students to performance on a mathematical test of a range of skills, such as single- and multi-digit arithmetic, fractions, and algebra. Working memory contributed unique variance to mathematical performance independent of chronological age, short-term memory, reading, and processing speed. In children with and without significant mathematics difficulties, Swanson and Beebe-Frankenberger (2004) found that working memory predicted solution accuracy on word problems independent of several academic and cognitive variables, such as fluid reasoning, reading skills, and phonological processing, among others. Additionally, Imbo et al. (2007) found that working memory plays a significant role during multi-digit arithmetic problem solving in adults.

Fluid reasoning is the ability to solve novel problems flexibly and deliberately without using previous information. More specifically, it is the ability to analyze novel problems, identify patterns and relationships that underpin these problems, and apply logic (Schneider & McGrew, 2012). In a synthesis of studies investigating the concurrent relationships between cognitive abilities and achievement measures, fluid reasoning was one of the three cognitive abilities that was consistently related to mathematical performance in calculation and problem solving at all age ranges throughout development (the other two were verbal comprehension and processing speed). Additionally, fluid reasoning has been found to predict future mathematical achievement (Green et al., 2017; McGrew & Wendling, 2010). Using a longitudinal cohort sequential design, Green et al. (2017) examined how fluid reasoning measured at three assessment occasions, spaced 1.5 years apart, predicted math outcomes for a group of 69 participants between ages 6 and 21 years across all three assessment occasions. Results revealed that fluid reasoning was the only significant predictor of future mathematical achievement, while age, spatial skills, and vocabulary were not significant predictors.

Green et al.'s (2017) finding that fluid reasoning, but not spatial skills, predicts future mathematical achievement is not surprising given that spatial skills and fluid reasoning skills are reported to be very strongly related (e.g., Fry & Hale, 1996; Green et al., 2017). Moreover, many tests of fluid reasoning skills require processing spatial information, such as Matrix Reasoning and Raven's Progressive Matrices. Despite the strong correlation between fluid reasoning and spatial skills, there is evidence that suggests that the two skills rely on separable cognitive processes and brain regions (e.g., Halford et al., 1998; Holyoak, 2012; Klingberg, 2006; Krawczyk, 2012; Vendetti & Bunge, 2014). Taken together, these findings indicate that gaining an accurate understanding of the relation between spatial skills and

mathematical skills requires accounting for their shared relations with fluid reasoning.

Furthermore, though many studies simultaneously examine the contribution of executive functions and spatial skills to mathematical achievement in their analyses (e.g., Green et al., 2017; Hawes et al., 2019; Verdine et al., 2017), the findings regarding the relations between the three cognitive skill sets have not been consistent. For instance, contrary to the findings of Green et al. (2017), Hawes et al. (2019) found that spatial skills were the strongest predictor of mathematical achievement in children ages 4–11 years after controlling for age, while executive functioning skills (which included measures of working memory) was not a significant unique predictor. Moreover, the patterns of the relations between these three factors remained stable across age and grade (Hawes et al., 2019). Thus, taken together, the differences in findings between these studies indicate that the nuances of the relations between spatial skills, executive functioning skills, and mathematical skills are not well understood.

In a recent theory of intelligence called process overlap theory, Kovacs and Conway (2016) postulate that any task requires a number of domain-specific and domain-general cognitive processes. The domain-specific cognitive tests measuring different cognitive skill sets (e.g., tests of spatial skills, tests of mathematical skills) all tap into a common pool of domain-general executive functions, such as working memory and fluid reasoning (Kovacs & Conway, 2016). In particular with regard to the relations between spatial skills and mathematical achievement, the role of domain-general executive functions has yet to be ascertained. More specifically, it remains unclear whether executive functions, such as working memory and fluid reasoning, mediate the relations between spatial skills and mathematical performance.

In addition to executive functions, another domain-general cognitive factor positively related to mathematical achievement is general intelligence, or *g* (Taub et al., 2008; Wrigley, 1958). General intelligence includes the ability to think logically and systematically (Embretson, 1995) and is the best individual predictor of achievement across academic domains, including mathematics (e.g., Deary et al., 2007; Jensen, 1998; Stevenson et al., 1976; Taub et al., 2008; Walberg, 1984). For instance, in a 5-year prospective study of more than 70,000 students, Deary et al. (2007) found that general intelligence assessed at 11 years of age explained nearly 60% of the variation on national mathematics tests at 16 years of age. A study by Geary (2011) aimed at identifying cognitive factors and quantitative competencies in first-grade students that predict mathematical achievement in fifth-grade students found that general intelligence, along with processing speed and working memory, predicted fifth-grade mathematical achievement as well as growth in achievement. While these studies solidify the contribution of general intelligence on mathematical achievement and performance, what is

unclear is the magnitude of its contribution to the relation between spatial skills and mathematical achievement.

In sum, existing research suggests that spatial skills and mathematical skills may not be directly related. Theories of intelligence indicate that other domain-general cognitive factors, such as working memory, fluid reasoning, and general intelligence, may underlie their relation. Thus, a more accurate understanding of how spatial skills and mathematical skills are related requires examining if and how associated skillsets influence the relation.

### **The influence of participants' age or grade-level and gender on the relations between spatial and mathematical skills**

Just as the influence of domain-general cognitive processes on the relation between spatial and mathematical skills is unclear, whether the relation between the skill sets varies across development is also unknown. Much prior research suggests that the relation between spatial skills and mathematical skills could vary depending on participants' age or grade level (Battista, 1990; Stannard et al., 2001; Wolfgang et al., 2001). For instance, in a longitudinal study by Li and Geary (2013), first- to fifth-grade gains in visuospatial memory predicted the end of fifth-grade mathematical achievement. However, visuospatial memory was not related to mathematics in first grade (Li & Geary, 2013). Similarly, in a study by Hanline et al. (2010), block construction scores, an indicator of students' understanding of spatial relations and geometric knowledge, predicted preschoolers' scores on the Test of Early Reading Ability at age 8 years and the growth rate on the same test from age 5–8 years. However, there was no significant relation between block construction scores and scores on the Test of Early Math Ability at age 8 or the growth rate for the Test of Early Math Ability from ages 5–8. Similarly, in a cross-sectional study, Mix et al. (2016) examined the relations between spatial and mathematical skills in 854 students from kindergarten, and third and sixth grades and found that mental rotation skills were strongly related to mathematics performance in kindergarten and third-grade students, but was not a significant predictor of mathematics performance in sixth-grade students. Instead, working memory and visual motor integration showed the strongest relations with mathematical performance in sixth-grade students.

While some studies show that the relations between spatial and mathematical skills vary across development, there are some studies that show that the relations remain stable across ages. For example, as previously discussed, Hawes et al. (2019) found that the strong relations between spatial skills and mathematical skills were evident in children from ages 4 to 11 years. Specifically, the relation between spatial skills and mathematical skills remained consistent across the ages. Taken together, these mixed findings underline the lack

of understanding in the field about if and how the relations between spatial skills and mathematical skills vary between ages or grades. Identifying if and how the relation between spatial skills and mathematical skills varies across development is critical because it could shed light on when implementing spatial interventions may be most effective in improving students' mathematical outcomes.

Just as the relations between spatial skills and mathematics outcomes could vary by age or grade level, studies have suggested that gender could also be a moderating factor. Women are under-represented at the highest levels of STEM occupations (Ceci & Williams, 2011; Halpern et al., 2007), especially in math-intensive STEM domains (e.g., computer and information sciences, engineering, and physical and technical sciences; see Wang & Degol, 2017, for a review). Moreover, boys outperform girls on tests of mathematical aptitude, such as on the SAT mathematics subtest (SAT-M; Halpern et al., 2007), especially on the most complex problems (Hyde et al., 1990). Underlying the gender differences in mathematical outcomes are reported differences in men and women's spatial skills (e.g., Casey et al., 1995; Nuttall et al., 2005), as well as differences in the relations between spatial skills and mathematical skills (e.g., Tartre & Fennema, 1995). For instance, Casey et al. (1995) found that gender differences on the mathematics section of the SAT-M were related to gender differences on a spatial skills measure in high-achieving students. Boys/men performed better than girls/women on both mental rotation and the SAT-M (Casey et al., 1995).

Beyond these gender differences in levels of mathematical or spatial skills, there is also some evidence for gender differences in the strength of the relations between spatial skills and mathematical skills (e.g., Tartre & Fennema, 1995). Tartre and Fennema (1995) examined the relationship between spatial skills, verbal skills, and mathematical achievement in 60 students as they progressed from sixth to 12th grades. Unlike Casey et al. (1995), they found no significant difference between boys' and girls' spatial skills, verbal skills, or their mathematical achievement. However, they found that the relations between the three factors varied by gender. Spatial skills were a consistent significant predictor of mathematical achievement for girls across the years, but not for boys, whereas verbal skills were a significant predictor of mathematics across the years for boys, but not for girls (Tartre & Fennema, 1995). On the contrary, Ganley and Vasilyeva (2011) examined the relations between spatial skills and mathematical performance in middle school students and found that despite similar levels of mathematical performance for boys and girls, spatial skills predicted mathematical performance in boys, but not in girls. As the research findings on the role of gender in the association between spatial skills and mathematical skills is not consistent, a systematic

investigation synthesizing prior findings is necessary to elucidate how gender impacts the relationship between these skill sets.

### Prior meta-analytic findings on the relations between spatial skills and mathematical skills

With the aim of better understanding the relations between spatial skills and mathematical reasoning skills, a meta-analysis was recently conducted by Xie et al. (2019) summarizing the findings of 73 studies reporting correlations between these two factors. The aims of their study included: (1) determining whether there is a significant association between spatial skills and mathematical skills, and (2) determining whether the domains of spatial skills, mathematical skills, age, and developmental disability status moderated this relationship. In this meta-analysis, Xie et al. (2019) report finding a medium positive association between spatial skills and mathematical skills ( $r = .27$ ; 95% confidence interval (CI) [0.24, 0.32]), which did not differ by age, developmental disability status, or type of spatial skill. However, they did find that mathematical domain did moderate the relations between mathematical and spatial skills. Specifically, logical reasoning showed the strongest association with spatial skills in comparison with numerical skills or arithmetic skills (Xie et al., 2019). Important to the research reported here, Xie et al. (2019) did not examine if domain-general reasoning skills, such as working memory or fluid reasoning skills, mediate the relation between spatial and mathematical skills, or whether the magnitude of the relation between the two factors differ by gender. Thus, the role of domain-general reasoning skills and participants' gender on the relations between spatial skills and mathematical outcomes is still unknown.

### The current study

As discussed above, the nuances of the relationship between spatial skills and mathematical skills remains unclear. Specifically, it is unknown if the relation differs by gender and how much domain-general reasoning skills may account for the association. Furthermore, although Xie et al. (2019) found that age was not a significant moderator in their meta-analysis, we decided to include the factor in our analyses to see if we could replicate their finding.

In this study, we examined two sets of questions that each required different analytical tools. As such, we chose to conduct two separate meta-analyses that would each address one set of questions. In the first meta-analysis, we employed correlated and hierarchical effects (CHE) meta-regression models (Pustejovsky & Tipton, 2021) to answer the following research questions: (1) What is the magnitude of the relation between spatial skills and mathematical skills? and (2) What is the effect of gender and grade-level on the association

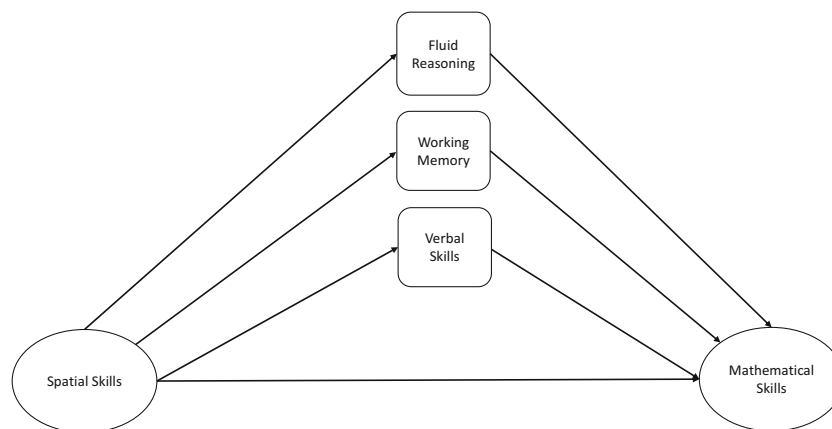
between spatial skills and mathematical skills? In the second meta-analysis, we used the meta-analytic structural equation modeling (MASEM) approach to examine how accounting for domain-general reasoning skills (working memory, fluid reasoning, and verbal skills) impacts the relation between spatial skills and mathematical skills (depicted in Fig. 1). Based on Carrol's (1993) model of cognitive abilities, Wai et al. (2009) posit that cognitive abilities center around three cognitive domains: quantitative/numerical, spatial/pictorial, and verbal/linguistic (or mathematical, spatial, and verbal domains, respectively). The shared variance across these three content domains can be attributed to the higher order construct of general intelligence ( $g$ ) (Wai et al., 2009). Thus, we included measures of verbal skills in our analyses to account for the variance contributed by  $g$  in the relationship between spatial skills and mathematical skills.

## Method

### Literature search/information sources

We began with electronic searches of PsycINFO, ProQuest, and ERIC databases. We searched all available records published since 1997 to the date of the search: 1 May 2018. We limited our search to this 20-year window for three reasons. First, the time frame was large enough to provide a wide range of studies and would encompass the large increase in studies on the topic of mathematics and spatial skills that has recently occurred. Second, the time frame is narrow enough to allow us to gather most of the relevant published and unpublished data for analysis. Third, educational structures (i.e., schooling and curriculum) have remained mostly unchanged during this time frame (Payne, 2008).

We used the following search terms: (*Intervention OR training OR practice OR class OR enhancement OR education OR quasi-experimental OR experimental*) AND ("*spatial skills*" OR "*spatial ability*" OR "*spatial cognition*" OR "*spatial thinking*" OR *visuospatial*) AND (*mathematics OR math OR "math skills" OR "math ability" OR "mathematical reasoning"*). Where possible, database search filters were employed prior to screening. For example, when searching the ERIC database, the following filters were used: *Peer-Reviewed Journal Publications*, and *Within 21 Years of the Search Date*. To obtain unpublished data, we conducted a search in ProQuest Dissertations and Abstracts. When searching ProQuest Dissertations and Abstracts, we limited our search term to within the abstracts due to the expansive nature of dissertations, which significantly reduces the effectiveness of search terms. In addition to an electronic search of databases, we acquired unpublished data through requests in social media posts and emails to various relevant research groups (e.g., *Cognitive Development Society*, *The Spatial*



**Fig. 1** Hypothesized path model between cognitive constructs. *Note.* An illustration of the hypothesized theoretical relations between spatial skills and mathematical skills, and general reasoning skills to be examined

*Intelligence and Learning Center network listserv*). Furthermore, prominent researchers in the field were contacted directly via email for any current or previous unpublished work.

Searching the databases using the Boolean phrases and filters mentioned above resulted in the acquisition of 858 articles. All resulting articles were compiled within an Endnote database and duplicates were removed with 760 unique articles remaining. An additional nine articles were acquired through social media and email requests, resulting in a total of 769 articles.

### Abstract screening

All studies found through our initial search procedure were compiled into a numbered study catalogue. As noted by Berman and Parker (2002), “Failure to blind review could lead to biases similar to those in a record review when subjects are selected by investigators who are not blinded to the outcomes of interest” (p.4). In this catalogue, each study was assigned a unique identifier and only the study title and abstract were displayed for evaluation in line with best practice (Berman & Parker, 2002). The abstracts for all acquired studies were then evaluated by the first and second authors to establish whether they included at least one measure of spatial skills and at least one measure of mathematical skills. Studies that did not mention a mathematical measure or a spatial measure in the study abstract were excluded from further analysis. Additionally, studies where manuscripts were not written in English, or an English translation was not able to be acquired, were also excluded from further analysis.

To ensure reliability of the abstract coding and selection process, 25% of the abstracts were randomly selected and screened by both reviewers for inclusion of the appropriate measures. Inter-rater reliability using Cohen’s Kappa was then calculated for this subset of studies, which indicated that

using MASEM analysis. Fluid reasoning, working memory, and verbal skills are hypothesized to be moderators between spatial skills and mathematical skills in this model

satisfactory levels of inter-rater reliability were observed ( $k = 0.77, p < .01$ ) (Altman, 1990). After inter-rater reliability was calculated for the abstract coding, any disagreements in study selection were discussed and agreed upon prior to final designation. The remaining 75% of the studies were randomly assigned to each reviewer for the completion of the abstract-coding process. 592 studies were removed from further analysis during the abstract-evaluation phase. After the study abstracts were evaluated, manuscripts coded as *included* were retained for full manuscript review.

### Full manuscript review and data extraction

One hundred and seventy-seven full manuscripts were retained for full manuscript review and data extraction, which was again conducted by the first and second authors. Each manuscript was reviewed and the following information was extracted: (1) measure(s) of spatial skills, mathematical skills, working memory, fluid reasoning, and verbal skills administered; (2) bivariate correlations between measures of spatial skills and measures of mathematical skills; (3) bivariate correlations between measures of spatial skills and measures of working memory, fluid reasoning, and verbal skills; (4) bivariate correlations between measures of mathematical skills and measures of working memory, fluid reasoning, and verbal skills; (5) the reported sample size for each correlation; (6) the age or grade-level of the participants; and (7) the proportion of males relative to female participants. If during full manuscript review, a study was found to be missing information about the spatial skills measure(s) administered, the mathematical skills measure(s) administered, or the bivariate correlation(s) between them, the manuscript authors were contacted for the information. If no response was provided, the study was excluded from further data analysis. Measures of working memory, fluid reasoning, or verbal skills were not required to retain the manuscript for further data analysis. A

**Table 1** Articles included in the meta-analyses

	% Male	Sample size <sup>a</sup>	Spatial measure description	Math measure description
Ackerman and Wolman (2007)	58.45%	142	Paper folding Spatial orientation Spatial analogies Verbal test of spatial ability	Math approximation Math knowledge Number series Arithmetic
Bonny and Lourenco (2015)	52.50%	40	Area discrimination task	Multiple choice geometry task
Carr et al. (2008)	nr	227	Vandenberg test of mental rotation	Criterion referenced competency test
Casey et al. (2011)	51.61%	124	Spatial visualization task and a mental rotation test	Formula Based Measurement Spatial Conceptual Measurement Numerical Skills (Items from NAEP and TIMSS 4th grade)
Casey et al. (2015)	0.00%	127	Block Design of the WISC-IV 2D Mental Transformation Task 3D Mental Rotation Task	Addition task Subtraction task
Casey et al. (1997)	45.74%	94	Mental rotation test	Geometry test SAT math
Cirino et al. (2016)	53.00%	150-162	Vandenberg and Kuse Spatial Test	Estimation task Procedural computation Fraction competency Proportional reasoning State test Fact-based retrieval
Clifford (2008)	49.00%	112	WISC Block Design Picture Completion WISC Visual Spatial Form Patterns Position & Direction Visual Spatial Factor	WJ III
Cui et al. (2017)	40.36	223	Mental rotation test	Exact computation Approximate Computation Math achievement test
Edens and Potter (2007)	50.00%	214	Case's Framework	General math test
Frick (2019)	57.14	119	Perspective taking task Mental rotation Diagrammatic representations task Cross-sectioning task Spatial scaling task Mental transformation task T2 mental rotation T3 mental rotation (figure rotation)	T1 Basic calculation T2 number line task T2 Proportional reasoning T3 mathematics (HRT composite score) T3 Geometry
Geer and Ganley (2021)	65.67	67	Spatial Orientation Task Money RoadMap Task	Math fluency
Gilles and Bailleux (2001)	57.40%	122	Spatial sub-test from the Ability Factor Battery	Numerical subtest from the Ability Factor Battery

**Table 1** (continued)

	% Male	Sample size <sup>a</sup>	Spatial measure description	Math measure description
				MAST: Mathematical achievement score on scholastic test
				General math test
Gilligan et al. (2017)	nr	12,099	Pattern Construction subscale of the BAS-II	Progress in Maths (NFER PiM)
Green et al. (2017)	48	69	Spatial span & block design	Number series Applied problems Math fluency
Haciomeroglu (2015)	54.60%	150	Cube comparison  Card Rotations Form Board Paper Folding	Calculus
Haciomeroglu (2016)	60.29%	348	KIT Reference Tests for Cognitive Factors	AP calculus test
Hegarty and Kozhevnikov (1999)	100.00%	33	Space Subtest from PMA Space	General math test
			Block Design Subtest from WISC-R(Wechsler, 1976)	
Johnson (2017)	53.00%	304	Thurstone PMA	CRCT math competency test
Kirby and Boulter (1999)	54.50%	44	Hidden Patterns, Card Rotations, Surface Development	Geometry test
Kyttälä et al. (2003)	43.50%	46	Mental rotation task	Early Numeracy Test
Kyttälä and Lehto (2008)	54.69%	128	VSWM task (mental rotation)	General math test
Lowrie et al. (2017)	nr	66	Spatial Reasoning Instrument SRI	MathT test, which we developed using released items from Australia's National Assessment Program (NAPLAN)
Mix et al. (2016)	47.81%	274	Mental rotations, Visual Spatial Working Memory, Visual Motor Integration, Block Design and Map Reading Perspective Taking	General math test
Möhring et al. (2015)	50.00%	50	Spatial localization task	Proportional reasoning task
Morosanova et al. (2016)	49.69%	318	Mental Rotation Task	Problem Verification Task
				Understanding Numbers test
Olver (2013)	47.60%	21	Mental Transformation Test Number line estimation	Number Knowledge test
Oostermeijer et al. (2014)	50.00%	128	Picture rotation task	Math word problem solving
Paunonen and Hong (2010)	30.11%	176	Spatial subscale of the Multidimensional Aptitude Battery	Arithmetic subscale of the Multidimensional Aptitude Battery
Peng et al. (2017)	58.3	94	Blocks reconstruction subtest	Approximate addition task exact addition task rapid digit naming number identification numerical reasoning
Pyke (2003)	nr	174	The Gestalt completion test	Algebra and geometry test
Reuhkala (2001)	58.06%	62	A mental rotation task based off of Shepard & Metzler	total score of a national math test
Rodán et al. (2016)	51.11%	45	Spatial Ability Subtest, EFAI-3	Numerical Ability Subtest, EFAI-3
Simms et al. (2016)	51.95%	77	Arrows subtest from NEPSY	WIAT-II Mathematical Reasoning and Numerical Operations subscales



**Table 1** (continued)

	% Male	Sample size <sup>a</sup>	Spatial measure description	Math measure description
			Mr. X mental rotation	
Skagerlund and Träff (2016)	39.80%	133	Mental rotation; paper folding	Multidigit calculation; arithmetic equations; arithmetic fact retrieval;
Tolar et al. (2009)	49.74%	195	Differential Aptitude Tests (DAT) space relations subtest	AAIMS algebra content test
van Garderen (2006)	40.90%	22	Block Design subtest from the WISC-III	WJ-III ACH Applied Problems sub-test
			MGMP-SVT	MPI - mathematical word-problem solving performance
			Block Design WISC-III	Mathematical Processing Instrument (MPI)
			Spatial Visualization Test MGMP-SVT	
Ventura et al. (2013)	39.94%	323	Spatial orientation test	SAT math test
			Mental rotation test	EMAS age 3 y
Verdine et al. (2017)	53.92%	81-102	CMTT age 5 y	EMAS age 4 y
			TOSA age 3 y	High Count age 3 y
			TOSA age 4 y	High Count age 4 y
			WPPSI Block design subtest age 4 y	WIAT Math Problem solving test age 4 y
			WJ III space relations subtest age 4 y	Number Sense Brief
			WJ III space relations subtest age 5 y	WIAT age 5 y
Verdine et al. (2014)	53.92%	102	Test of spatial assembly (TOSA) ( Dimensional)	Early Mathematics Assessment System: Number and Operations subtest
			Test of spatial assembly (TOSA) (Match)	
Voyer (1998)	43.09%	123	The Vandenberg and Kuse Mental Rotations Test	General math test
			The Water Level Task	General math test
Wei et al. (2012)	50.00%	80	3D rotation	Advanced Math Concepts
Wong (2017)	0-100%	83 - 99	Levine's Mental transformation Task	Diagnostic Tests for Metacognitions and Mathematics
Zhang and Lin (2018)	38.67	70-75	Spatial perception	Arithmetic - non symbolic
	38.67	70	Spatial visualisation	Arithmetic - word problems
				Arithmetic – written
Zhang and Lin (2017)	46.23	71-106	Spatial Perception subtest from the Test of Visual- Perceptual Skills Revised	Arithmetic word problems
			Visual-spatial subtest from the Test of Visual- Perceptual Skills Revised	
			Spatial Visualization subtest of Spatial Relations	
			Mental Rotation	

*Note.* If a commonly used acronym or name was used in reference to a measure this was included as used in the original paper. Where possible non-widely used tests were described using the language of the original paper. For the sake of conciseness full measurement titles and sources are not listed here but can be found within the relevant papers. More complete notations regarding measure composition and additional extracted variables are included in the overall dataset available upon request. Where relevant information was not provided it has been recorded as nr: not reported

list of the 45 included manuscripts as well as the study characteristics and measure information extracted from each one (except for the bivariate correlations) is provided in Table 1. A summary of the steps carried out in the abstract screening and full manuscript review components of this study, based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart template (Moher et al., 2015), are outlined in Fig. 2.

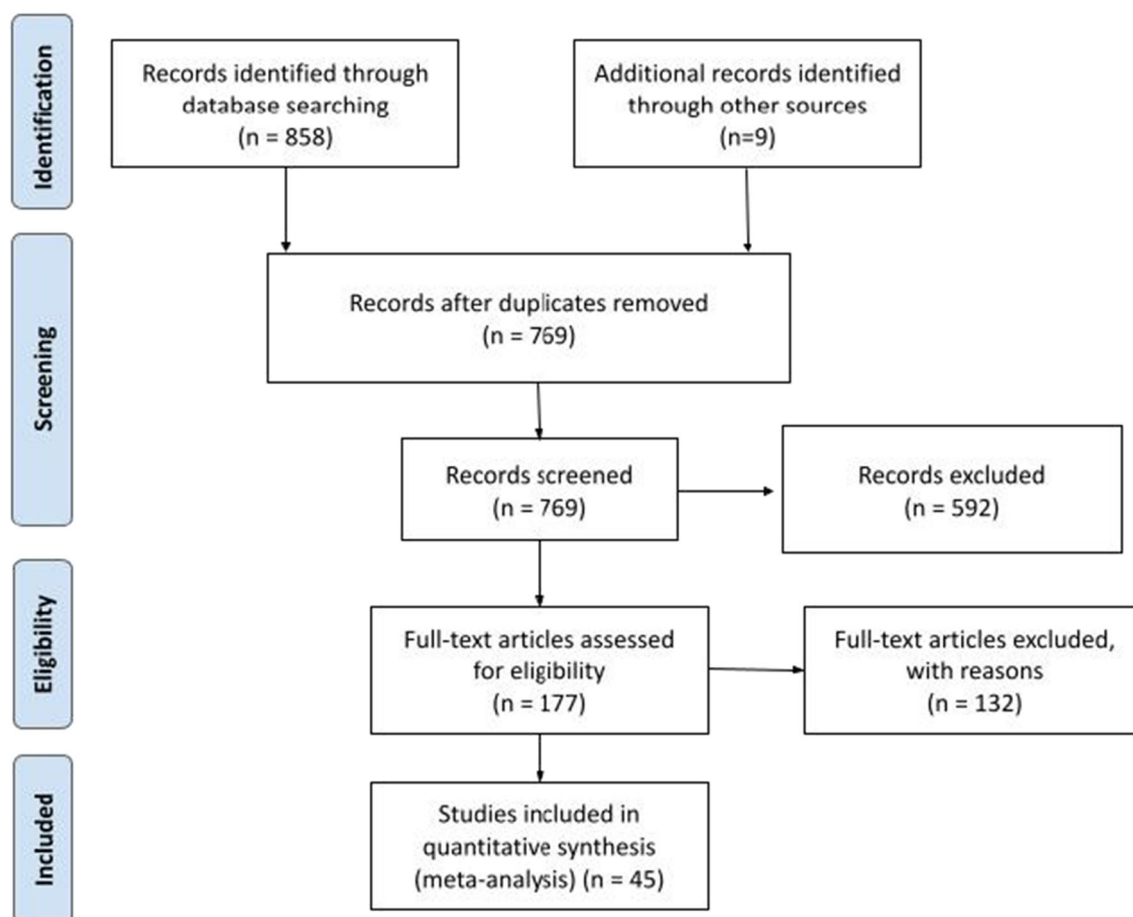
## Study characteristics and measures of interest

**Participant grade** Both the age of the participants and their grade-level reported within each study was extracted. If studies collected data on participants at multiple time points, age information at each administration of the relevant measures was extracted. As many studies examining mathematical understanding occur in a classroom setting (e.g., Verdine et al., 2017), if participants' grade-level was reported without age-related data, then age was approximated based on typical age per grade defined by the National Center for Education Statistics (2018). More studies reported on participants'

grade-level than age, so grade-level was used as a moderator of effect size in the analysis.

**Percent of male participants** To examine whether gender moderates the relationship between spatial skills and mathematical skills, the percent of male participants was extracted from each study. Few studies reported the correlation between spatial skills and mathematical skills by gender. Thus, we use percent of male participants as an indirect measure of whether the magnitude of the correlation between spatial and mathematical skills differs as a function of the percent of male participants.

**Spatial skills** Tasks that were considered as measures of spatial skills assessed participants' skills in visualizing and/or mentally manipulating objects or figures or navigating spaces (Atit, Uttal, & Stieff, 2020b; Newcombe & Shipley, 2015). Common examples of measures of spatial skills include the Mental Rotations Test (Peters et al., 1995; Vandenberg & Kuse, 1978), the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-IV) Block Design subtest (Wechsler, 2012), and the Paper Folding test (Ekstrom et al., 1976).



**Fig. 2** The process of article selection for inclusion in analysis. *Note.* After carrying out this selection process, data from 45 articles were included for analysis

**Mathematical skills** Tasks that were considered measures of mathematical skills assessed participants' skills in reasoning about "numbers and their operations, interrelations, combinations, generalizations, and abstractions and of space configurations and their structure, measurement, transformations, and generalizations" (Merriam Webster, n.d.). Examples of measures assessing mathematical skills include assessments of arithmetic (e.g., Ackerman & Wolman, 2007), geometry tests (e.g., Bonny & Lourenco, 2015), and the mathematics portion of the Criterion Referenced Competency Test (e.g., Carr et al., 2008).

**Working memory capacity** Tasks that were considered measures of working memory capacity assessed participants' ability to both retain and process information simultaneously (Conway et al., 2003). Examples of working memory measures include Backwards Digit Span (GrÉgoire & Van Der Linden, 1997), Automated Symmetry Span (Oswald et al., 2015), and the Corsi Block Task (Corsi, 1972; Miyake et al., 2001).

**Fluid reasoning** Fluid reasoning tasks measured participants' ability to flexibly and deliberately solve novel problems without using previous information. More specifically, they measured participants' ability to analyze novel problems, identify patterns and relationships that underpin these problems, and apply logic (Schneider & McGrew, 2012). Common measures of fluid reasoning include the Matrix Reasoning from the Perceptual Reasoning Index in the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV; Flanagan & Kaufman, 2004), the Culture Fair test (Cattell, 1971), and the Diagramming Relationships Test (Ekstrom et al., 1976).

**Verbal skills** Measures of verbal skills assess participants' vocabulary, verbal comprehension skills, and/or verbal reasoning capacity, and are strong predictors of achievement in school and the ability to learn in non-school settings (Gottfredson & Deary, 2004). Examples of verbal measures include the WAIS-R Information (Wechsler, 1981), the Woodcock-Johnson III Picture Vocabulary (Woodcock et al., 2001), and the verbal portion of the SAT (SAT-V; College Board, 2020).

### Effect size measures

Correlation coefficients representing the bivariate correlation between spatial skills and mathematical skills, spatial skills and a measure of domain-general reasoning skills (i.e., fluid reasoning, working memory capacity, and verbal skills), and mathematical skills and a measure of domain-general reasoning skills were used as measures of effect size for this study. Prior to conducting any analyses, for our first meta-analysis (examining the magnitude of the relation between spatial

skills and mathematical skills), we converted all correlation coefficients to the Fisher's  $z$  scale (Borenstein et al., 2005). After the analysis, all the results in Fisher's  $z$  scale were then converted back to correlation coefficients for the ease of interpretation. For our second meta-analysis, a MASEM approach was used to examine the contribution of domain-general reasoning skills to the relations between spatial skills and mathematical skills, and the original bivariate correlation coefficients were utilized.

### Analysis

**Study 1: What is the magnitude of the relation between spatial skills and mathematical skills?** To answer the above research question, we first fit an intercept-only random-effects model and estimated the pooled correlation coefficient between spatial and mathematical skills. Next, we examined whether the grade and gender (i.e., proportion of males in the primary study sample) explained further variability in effect sizes across different studies using a meta-regression model. Because the unit of analysis is the study, the grade and gender are the study and sample characteristics, not an individual participant characteristic.

Given that there is more than one kind of spatial skill and, thus, more than one kind of spatial test (Linn & Petersen, 1985; Newcombe & Shipley, 2015), and given that there is more than one kind of mathematical domain, many of the studies included in our meta-analyses administered more than one measure of spatial skills or mathematical skills. Therefore, multiple effect sizes between different types of spatial tests and measures of mathematical skills were reported based on the same sample. To account for the dependent structure of effect sizes within studies, robust variance estimation (RVE) with the small-sample correction technique (Hedges et al., 2010; Tipton, 2015) was used. Specifically, we incorporated an extended working model for RVE that models the correlated and hierarchical structure of effect size estimates, called the *Correlated and Hierarchical Effects* (CHE) working model (Pustejovsky & Tipton, 2021). Many studies included in the synthesis report multiple correlations from the same participants (e.g., the correlation between algebra performance and mental rotation skills and algebra performance and perspective-taking skills) leading to dependencies among correlations nested within individuals and studies. The CHE working model represents our data structure and is described by Pustejovsky and Tipton (2021) as the most broadly applicable model in social science systematic reviews. In applying CHE to the data, we assumed a constant correlation of .8 for the effect sizes within studies.

The random-effects meta-regression models were fit using the CHE working model to estimate the mean correlation between spatial skills and mathematical skills, and to examine whether the correlation between spatial skills and

mathematical skills was moderated by grade or gender. The R packages *metafor* (Viechtbauer, 2010) and *clubSandwich* (Pustejovsky, 2020) were used for the analyses, and the variance components were estimated using restricted maximum likelihood (REML).

### Study 2: How do domain-general reasoning skills influence the relation between spatial skills and mathematical skills?

To examine whether domain-general reasoning skills account for some of the relation between spatial skills and mathematical skills, we hypothesized a path model with verbal skills, and fluid reasoning as mediators of the relationship from spatial skills to mathematical skills as seen in Fig. 3. We had originally planned to use working memory in the path model, but we decided to remove it due to the insufficient numbers of studies that reported correlations with working memory. For example, the nine correlation coefficients including working memory and verbal skills are derived from only three studies (see Table 2). This adversely affects the degrees of freedom of the pooled correlation coefficients. The degrees of freedom of the pooled correlation were less than 4, which indicates unreliable estimates when using RVE (Tipton, 2015). Thus, we removed working memory from the path model.

We used MASEM with a two-stage structural equation modeling (TSSEM) process to estimate the path model (Cheung & Chan, 2005). In the first stage of MASEM, the correlation matrices are pooled across studies. In the second stage, the pooled correlation matrix from the first stage is used to fit the structural equation model. In the first stage, we implemented RVE with the small-sample correction using the CHE working model (Pustejovsky & Tipton, 2021) to pool the correlations in order to incorporate dependent effect-size estimates within study. Our approach differs from a standard TSSEM application, which assumes one correlation matrix is reported per study, a limitation of current MASEM applications (Wilson et al., 2016). Thus, we implemented RVE so that the variance-covariance matrix of the pooled correlation matrix takes into account the dependent structure of correlations within studies. As in the meta-regression analysis, the R packages *metafor* (Viechtbauer, 2010) and *clubSandwich* (Pustejovsky, 2020) were used. In the second stage, we fit the path model using the  $4 \times 4$  pooled correlation and variance-covariance matrices. We used the *metaSEM* R package (Cheung, 2015) to estimate the path model provided in Fig. 3.

### Publication bias and selective reporting

To examine possible publication bias in the data, we implemented a modified version of Egger's test of funnel plot asymmetry (Egger et al., 1997), or Egger's sandwich (Rodgers & Pustejovsky, 2020), that can be used when dependent effect sizes exist in the meta-analysis data. Egger's sandwich uses

the same approaches as we used in our meta-analysis (i.e., RVE with CHE working model using Fisher's z scale effect sizes was used to reduce any artifactual correlation between the effect size and its variance estimates (i.e., measure of precision).<sup>1</sup> The results suggested that funnel plot asymmetry was not present ( $p = .642$ ), indicating no evidence of publication bias.

## Results

### Sampling characteristics

Of the 45 studies examined for the two meta-analyses, 18 of the studies were conducted with participants outside of the USA. A summary of the different countries from which participants were recruited is provided in Table 3. Regarding the distribution of studies across educational levels, 18 of the studies were conducted with participants in preschool/primary grades (i.e., preschool to fifth grades), 18 of the studies were conducted with participants in secondary grades (sixth–12th grades), and one study was conducted with participants in both preschool/primary and secondary grades. The remaining eight studies were conducted in participants at the post-secondary level.

### Descriptive

In total, the data include 568 correlation coefficients across ten types of correlations from 45 studies. Table 2 shows the number of reported correlation coefficient effect sizes and the number of studies by the type of correlation. Again, because the unit of analysis is the study, not the individual participants, the number of studies is our focused sample size in meta-analysis. Note that for Study 1, we used the subset of the dataset only using the correlations between spatial and mathematical skills.

### Study 1

In study 1, we estimated the pooled correlations between spatial and mathematical skills and examined whether grade and gender moderated the relationship between spatial and mathematical skills. To estimate the pooled correlation between spatial and mathematical skills, we fit an intercept-only meta-regression under the CHE working model. The pooled effect size in the correlation coefficient metric was .358 (95% CI [.295, .419]) and it is statistically significantly different from zero ( $\beta = .375$ ,  $SE = .035$ ,  $t(43.277) = 10.635$ ,  $p < .001$ ). The

<sup>1</sup> We adopted this idea from Pustejovsky and Rodgers (2019), who focused on the standardized mean difference. In the current study, we applied this idea to correlation coefficients and Fisher's z transformation.

**Table 2** The number of correlations, the number of studies, and the total average number of participants for each pairing

Pairing	Number of correlations	Number of studies	Total averaged <i>n</i>
Math and Fluid Reasoning	48	14	1,711
Math and Verbal	81	18	14,286
Math and Working Memory	67	11	1,716.86
Spatial and Fluid Reasoning	32	12	1364
Spatial and Math	181	45	17833.93
Spatial and Verbal	96	19	14326
Spatial and Working Memory	25	10	1401
Verbal and Fluid Reasoning	20	6	596
Working Memory and Fluid Reasoning	9	6	976
Working Memory and Verbal	9	3	377

*Note.* The “Total Averaged *n*” is the sum of the averaged sample size across the studies reporting the relevant correlation. Averaged sample size is the mean of the sample size for each relevant correlation within each study

*Spatial* denotes spatial skills, *Math* denotes mathematical skills, and *Verbal* denotes verbal skills

estimated between-study heterogeneity ( $\tau^2$ ) was .039 and the estimated within-study heterogeneity ( $\omega^2$ ) was .017.

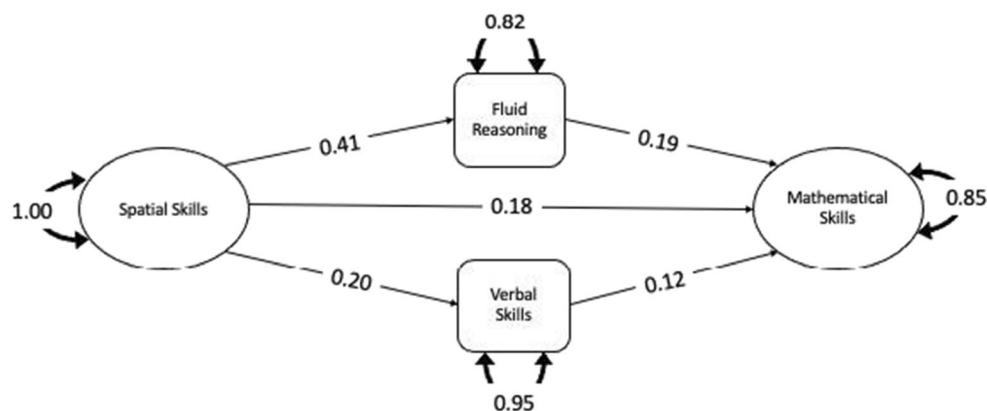
Next, we conducted a meta-regression using grade and proportion of males as moderator variables. There were three missing values in grade from two studies and five missing values in proportion of males from four studies among 181 effect size estimates. The effect sizes with missing moderators were deleted to fit the meta-regression model. Thus, 173 correlation coefficients from 41 studies were used to examine the effects of grade and proportion of males. The results of the meta-regression model are presented in Table 4. Neither moderator was significantly related to effect size. The degrees of freedom for the test of proportion of males in the sample is less than 4, indicating little information in the data about the relationship between the proportion of males in the sample and the correlation between spatial skills and mathematical skills. The proportion of males in the studies has a mean of 0.496 with a standard deviation of 0.13, indicating little variation among

studies in this variable. As discussed by Hedges et al. (2010) as well as Tipton (2015), results from RVE with small degrees of freedom should be interpreted cautiously. In sum, results indicated a positive moderate association between spatial skills and mathematical achievement. However, neither grade nor gender significantly moderated the relationship.

## Study 2

In study 2, a path model was examined to understand the relationships between spatial and mathematical skills as well as fluid reasoning and verbal skills as mediators. In this set of analyses, the  $4 \times 4$  correlation matrix of these four variables is the unit of analysis.

For the first stage of TSSEM, we pooled correlation matrices across studies and estimated the variance-covariance matrix using the CHE model given the presence of dependent correlations within studies (Pustejovsky & Tipton, 2021).



**Fig. 3** Path model examining the relations between spatial skills and mathematical skills. *Note.* This figure shows the paths between spatial skills, mathematical skills, and general reasoning skills (i.e., fluid reasoning and verbal skills) examined using a modified TSSEM approach for MASEM.

**Table 3** Summary of the study locations for the studies included in the meta-analyses

Country	No. of studies
China	6
Switzerland	1
Finland	3
Australia	1
Russia	1
Canada	2
Netherlands	1
Spain	1
United Kingdom	1
Sweden	1
United States	27

*Note.* Study location is the country from which participants were recruited in the study

Table 5 provides the pooled correlation matrix and Table 6 provides its corresponding standard error estimates. The pooled correlation coefficients among four variables ranged between .277 (verbal skill and fluid reasoning) and .418 (spatial skill and fluid reasoning). All six correlation coefficients were statistically significant at  $\alpha = .05$  and generally moderate in magnitude.

We also examined whether the pooled correlation coefficients differed due to the moderators of grade and gender (i.e., proportion of males in a sample). Neither of the moderators were significantly related to the magnitude of the correlation coefficients. Thus, we used a single pooled correlation matrix to fit a path model in the second stage.

In the second stage of TSSEM, we fit a path model using the pooled correlation matrix and variance-covariance matrix in the first stage. The total sample size used in this stage was 17,824.80 obtained by averaging sample size within studies and summing across studies. The total averaged sample size within study per pooled correlation coefficient is presented in Table 2. The second stage ensures that the parameter estimates take into account the precision of each pooled correlation

**Table 4** Results of a meta-regression examining the effect of age and gender on the relations between mathematics and spatial skills

Coefficient	<i>r</i>	<i>SE</i>	<i>t</i>	<i>df</i>	<i>p</i>
Intercept	0.27	0.06	4.32	3.12	0.02
Grade	-0.0002	0.01	-0.03	22.72	0.97
Proportion Male	0.20	0.07	2.76	1.65	0.14

*Note.* The reported estimated coefficients (*r*) are Inverse-Fisher's *z* transformed to correlation metric and test results are in Fisher's *z* metric. The degrees of freedom (*df*) and *p* values are based on a Satterthwaite approximation for small-sample correction

**Table 5** The pooled correlation matrix in the lower triangle and its standard error estimates in the upper triangle

Variables	Spatial	Verbal	Fluid Reasoning	Math
Spatial		.036	.031	.031
Verbal	.286*		.038	.045
Fluid Reasoning	.418*	.277*		.032
Math	.346*	.313*	.334*	

*Spatial* denotes spatial skills, *Math* denotes mathematical skills, and *Verbal* denotes verbal skills. \*  $p < .05$

coefficient from different sample sizes. The hypothesized path model included a direct path from spatial skills to mathematical skills and indirect paths through verbal skills and fluid reasoning (see Fig. 3). We implemented likelihood-based confidence intervals (Neale & Miller, 1997) to test indirect and direct effects. The approximate goodness-of-fit indices for the model were mixed with root mean square error of approximation (RMSEA; Brown & Cudeck, 1993) of .043 (95% CI [.031, .056]), and standardized root mean residual (SRMR; Hu & Bentler, 1999) of .110. Cheung (2015) suggested RMSEA and SRMR to evaluate the second stage of TSSEM. RMSEA less than or close to 0.06 and SRMR less than or close to 0.08 support good fit (Hu & Bentler, 1999). The model fit indices showed somewhat mixed conclusions about the hypothesized path model. Given the exploratory nature of this analysis, we treat the interpretation of the path coefficients including direct and indirect effects among the skills as suggestive. Table 7 shows the results of the path model including direct and indirect effects. All direct and indirect paths were statistically significant ( $p < .05$ ). The strongest direct path was spatial skills to fluid reasoning. The direct path from spatial skills to mathematical skills remained significant, even after accounting for the indirect paths that included fluid reasoning or verbal skills. Fluid reasoning mediated more of the effect from spatial skills to mathematical skills than did verbal skills, and the relation that remained between spatial skills and mathematical skills was larger than both of these indirect relations. Figure 3 shows the path model with the path coefficients.

## Discussion

This study investigated the relations between spatial skills and mathematical skills. Furthermore, it examined the effect of gender and age on this association, and the role of domain-general reasoning skills in the relations between these two constructs. Results from synthesizing the reported effect sizes between spatial skills and mathematical skills from 45 studies revealed a positive moderate association between the two skill sets ( $r = .36$ , robust standard error = 0.035,  $\tau^2 = 0.039$ ).

**Table 6** Correlation estimates between spatial skills, mathematical skills, and verbal skills

Correlation	<i>r</i>	<i>SE</i>	<i>t</i>	df	<i>p</i>
Spatial – Math	0.35	0.03	11.21	43.21	<0.001
Spatial – Verbal	0.29	0.04	7.98	30.71	<0.001
Spatial – Fluid Reasoning	0.42	0.03	13.30	17.32	<0.001
Math – Verbal	0.31	0.05	6.95	31.26	<0.001
Math – Fluid Reasoning	0.33	0.03	10.49	21.28	<0.001
Verbal – Fluid Reasoning	0.28	0.04	7.31	7.19	<0.001

*Spatial* denotes spatial skills, *Math* denotes mathematical skills, and *Verbal* denotes verbal skills

The degrees of freedom (df) and *p* values are based on a Satterthwaite approximation for small-sample correction

However, no significant effect of gender or grade on the association was found.

By implementing the meta-analytic SEM (MASEM) approach, the current study was able to model directional paths among the variables with direct and indirect effects simultaneously. In addition, we utilized RVE with the CHE working model approach (Pustejovsky & Tipton, 2021) to handle dependent effect sizes within studies when estimating the pooled correlation matrix and variance-covariance matrix. Results indicated that fluid reasoning and verbal skills mediated the relationships between spatial skills and mathematical skills, and these indirect effects were small but statistically significant. In addition to the indirect effects, the direct relation from spatial skills to mathematical skills was also statistically significant and larger than the indirect effects.

Consistent with prior research, this meta-analysis confirms that spatial skills are significantly related to mathematical skills (Atit, Power, et al., 2020a; Casey et al., 1995; Mix

et al., 2016, 2017; Verdine et al., 2017). More specifically, our study verifies findings of previous studies conducted using experimental, factor analytic, and meta-analytic methods showing that spatial skills and mathematical skills have a direct positive association (e.g., Lombardi et al., 2019; Mix et al., 2016, 2017; Xie et al., 2019).

The magnitude of the relation between spatial skills and mathematical achievement found in our study ( $r = .36$ ) is meaningful as it is consistent with findings of existing meta-analyses examining the relations between various cognitive skills and mathematical achievement. For instance, Jacob and Parkinson (2015) used meta-analytic techniques to identify the relation between executive functioning skills and mathematical achievement in children aged 2–18 years. Their results showed that executive functioning skills and mathematical achievement are positively and moderately correlated ( $r = 0.31$ ). Similarly, Peng et al. (2019) used meta-analytic techniques to determine the relation between fluid intelligence and mathematical achievement in individuals aged 3–80 years. Their analyses revealed a positive moderate association between fluid intelligence and mathematical achievement ( $r = 0.41$ ). Lastly, our findings replicate the findings of the meta-analysis conducted by Xie et al. (2019) who also found a moderate positive relation between spatial skills and mathematical achievement ( $r = 0.27$ ). These findings suggest that spatial skills are similarly related to students' mathematical understanding as other cognitive processes, such as executive functioning skills and fluid intelligence. Future research should examine how to leverage these pertinent cognitive skill sets to bolster students' mathematics understanding and performance.

However, unlike many studies closely examining the relationship between spatial skills and mathematical achievement (e.g., Casey et al., 1997; Li & Geary, 2013; Stannard et al.,

**Table 7** The results of the analyzed path model

Path	Estimate	95% confidence interval	
		CI.L	CI.U
Direct effect			
Spatial → Math	.18	.13	.24
Indirect effect			
Spatial → Fluid Reasoning → Math	.08	.05	.11
Spatial → Fluid Reasoning	.41	.33	.48
Fluid Reasoning → Math	.19	.13	.25
Spatial → Verbal → Math	.02	.01	.05
Spatial → Verbal	.20	.13	.28
Verbal → Math	.12	.06	.19

*Spatial* denotes spatial skills, *Math* denotes mathematical skills, and *Verbal* denotes verbal skills

CI.L is the lower bound of the 95% confidence interval, CI.U is the upper bound of the 95% confidence interval

2001), we found that the association between these two skill sets was not influenced by participants' grade or gender. It is well established that there is more than one kind of spatial skill (e.g., Linn & Petersen, 1985; Newcombe & Shipley, 2015) and more than one kind of mathematical domain (American Mathematical Society, 2020). Many of the studies investigating the effects of gender and age/grade on the association between the two constructs focus only on one kind of spatial skill (e.g., Casey et al., 2017; Li & Geary, 2013) and/or one kind of mathematical domain (e.g., Lombardi et al., 2019). In our study, we did not examine whether the relations differed for different kinds of spatial skills or different kinds of mathematical concepts. The meta-analysis conducted by Xie et al. (2019), which did consider domains of spatial skills and mathematical skills in their analyses, did not examine the effect of participants' age/grade or gender for each subarea. Thus, the nuances of how gender and age/grade influence the relations between spatial skills and mathematical skills have yet to be understood. Future meta-analyses should aim to disentangle how different kinds of spatial skills are related to different kinds of mathematical skills, and identify the role of age/grade and gender in these various associations.

In line with findings on the importance of executive functioning skills in mathematical achievement (e.g., Green et al., 2017; McGrew & Wendling, 2010), our study found that fluid reasoning skills were an essential component of mediating the relation between spatial and mathematical skills. In a synthesis of studies investigating the concurrent relationships between various cognitive abilities and achievement measures, McGrew and Wendling (2010) found that fluid reasoning was one of the three broad cognitive abilities (fluid reasoning, verbal comprehension, and processing speed) that was consistently related to mathematical performance at all age ranges. Furthermore, in a longitudinal study, Green et al. (2017) found that fluid reasoning, spatial skills, and verbal skills accounted for 90% of the variance in future math achievement in individuals ranging from 5 to 15 years of age, with fluid reasoning skills being the only significant predictor. In our study, the indirect relation from spatial skills to fluid reasoning to mathematical skills was larger in magnitude compared to the indirect relation from spatial skills to verbal skills to mathematical skills. However, the direct path from spatial skills to mathematical skills was larger than the two indirect paths. These results are consistent with Kovacs and Conway's (2016) process overlap theory, which poses that domain-general executive processes, such as fluid reasoning, underlie performance on domain-specific cognitive tasks, such as spatial and mathematical measures. There were not enough studies that measured working memory capacity to include the construct in our estimated model. Thus, an area of focus for future research is to ascertain the role of additional executive functioning processes, such as working memory capacity and

inhibition (Miyake et al., 2000), on the relations between spatial skills and mathematical performance.

One limitation of this meta-analysis is that while we tried to include all relevant studies in our analysis, some eligible studies may have been missed due to our search strategy. We used generalized terms such as "spatial skills" or "spatial ability" or "mathematics" or "math skills." However, as many studies focus on a specific kind of spatial skill or mathematical skill, they may have only used such specific terms, such as mental rotation or algebra, in their articles. Therefore, future studies should consider using a broad variety of search terms to reduce the possibility of missing eligible articles. Secondly, the small number of effect sizes limited our ability to examine the relations between different kinds of spatial skills and different kinds of mathematical domains. Thus, more effect sizes between sub-constructs are needed to examine these relationships in future studies.

Another limitation of this research is that the studies included in these meta-analyses were largely from the USA and other Western nations. Only a small number of studies, specifically seven studies, were conducted in Eastern countries (i.e., China and Russia). The lack of international diversity of the participants in the selected studies potentially biases our results as existing data indicates that mathematical achievement varies greatly between students from eastern and western nations (Mullis et al., 2012). Thus, more primary studies examining the relations between spatial skills and mathematical skills need to be conducted in samples across a broader range of countries. Furthermore, the search and inclusion criteria for literature to be examined in future meta-analyses needs to be revised, perhaps by expanding the literature search to studies published in languages other than English, so that a broader range of international studies are represented in the analyses.

The findings from these meta-analyses have multiple implications for mathematics education research and practice. First, this research highlights that integrating practices that develop and support students' spatial skills could benefit students' mathematical understanding at all educational levels. Many of the efforts aimed at improving students' mathematical achievement by way of bolstering their spatial skills have been carried out with students at the preschool/primary and secondary levels (e.g., Cheng & Mix, 2014; Hawes et al., 2015; Lowrie et al., 2017; Lowrie et al., 2019; Schmitt et al., 2018). Only a handful of efforts have been made to bolster students' mathematical outcomes by improving their spatial skills at the postsecondary level (e.g., Sorby, 2007; Sorby et al., 2013). However, our research shows that the relations between spatial skills and mathematical skills do not vary by participants' grade level. Therefore, future research should examine how to support mathematics instructors at the postsecondary level in developing and supporting their students' spatial skills.



Second, while our findings underline the importance of developing students' spatial skills to support their mathematics learning and achievement, efforts should also focus on developing students' domain-general cognitive processes, such as their verbal skills and fluid reasoning skills. Our research shows that domain-general cognitive processes account for some of the variance between spatial skills and mathematical skills. This suggests that bolstering spatial skills in conjunction with fluid reasoning and verbal skills may provide a greater boost to students' mathematical achievement than developing spatial skills alone. Thus, future research should aim to identify pedagogical practices focused on developing students' domain-general reasoning skills as well as their spatial skills, which can be integrated into preschool to postsecondary mathematics curricula.

## Conclusions

In conclusion, the current meta-analysis affirms that spatial skills and mathematical skills are positively related and, moreover, there is a direct relationship between the two constructs that did not vary based on grade or gender. Furthermore, we found that other cognitive processes, specifically fluid reasoning skills and verbal skills, mediated the relations between spatial skills and mathematical skills, and yet a direct relation between the two constructs independent of other general cognitive processes remained. These findings indicate that efforts to improve mathematical skills should include bolstering spatial skills but may also consider support for other pertinent cognitive skill sets.

**Acknowledgements** We would like to acknowledge Dr. Martha Carr who contributed to the initial stages of this research but sadly passed away during the summer of 2017.

**Availability of data and materials** The dataset generated during and/or analyzed during the current study are available at the Open Science Framework, [osf.io/h3w7j](https://osf.io/h3w7j).

**Code availability** The coding scheme and script used in conducting the analysis is also available from the corresponding author upon reasonable request. The dataset and analysis code associated with this project are available at the Open Science Framework, [osf.io/h3w7j](https://osf.io/h3w7j).

**Funding** This research was partially supported by a grant from the US Department of Education (award number R305A150365).

## Declarations

**Conflicts of interest/Competing interests** The authors have no financial or non-financial interests to disclose.

**Ethics approval** NA

**Consent to participate** NA

## References

- Ackerman, P. L., & Wolman, S. D. (2007). Determinants and validity of self-estimates of abilities and self-concept measures. *Journal of Experimental Psychology: Applied*, 13(2), 57–78. <https://doi.org/10.1037/1076-898X.13.2.57>
- Altman, D. G. (1990). *Practical Statistics for Medical Research*. CRC Press.
- American Mathematical Society (2020). *Homepage*. Retrieved November 24, 2020, from <https://www.ams.org/home/page>
- Atit, K., Gagnier, K., & Shipley, T. F. (2015). Student gestures aid penetrative thinking. *Journal of Geoscience Education*, 63(1), 66–72. <https://doi.org/10.5408/14-008.1>
- Atit, K., Power, J. R., Veurink, N., Uttal, D. H., Sorby, S., Panther, G., Msall, C., Fiorella, L., & Carr, M. (2020a). Examining the role of spatial skills and mathematics motivation on middle school mathematical achievement. *International Journal of STEM Education*, 7(1), 38. <https://doi.org/10.1186/s40594-020-00234-3>
- Atit, K., Uttal, D. H., & Stieff, M. (2020b). Situating space: Using a discipline-focused lens to examine spatial thinking skills. *Cognitive Research: Principles and Implications*, 5(1), 19. <https://doi.org/10.1186/s41235-020-00210-z>
- Aud, S., Hussar, W., Johnson, F., Kena, G., Roth, E., Manning, E., Wang, X., & Zhang, J. (2012). The condition of education 2012. NCES 2012-045. *National Center for Education Statistics*. <http://files.eric.ed.gov/fulltext/ED532315.pdf>
- Battista, M. T. (1990). Spatial visualization and gender differences in high school geometry. *Journal for Research in Mathematics Education*, 21(1), 47–60. <https://doi.org/10.2307/749456>
- Berg, D. H. (2008). Working memory and arithmetic calculation in children: The contributory roles of processing speed, short-term memory, and reading. *Journal of Experimental Child Psychology*, 99(4), 288–308. <https://doi.org/10.1016/j.jecp.2007.12.002>
- Berman, N. G., & Parker, R. A. (2002). Meta-analysis: Neither quick nor easy. *BMC Medical Research Methodology*, 2, 10. <https://doi.org/10.1186/1471-2288-2-10>
- Bethell-Fox, C. E., & Shepard, R. N. (1988). Mental rotation: Effects of stimulus complexity and familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 12–23. <https://doi.org/10.1037/0096-1523.14.1.12>
- Bonny, J. W., & Lourenco, S. F. (2015). Individual differences in children's approximations of area correlate with competence in basic geometry. *Learning and Individual Differences*, 44, 16–24. <https://doi.org/10.1016/j.lindif.2015.11.001>
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive Meta-analysis (Version 2)*. Biostat.
- Brown, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage Focus Editions*, 154, 136–136.
- Burte, H., Gardony, A. L., Hutton, A., & Taylor, H. A. (2017). Think3d!: Improving mathematics learning through embodied spatial training. *Cognitive Research: Principles and Implications*, 2(1), 13. <https://doi.org/10.1186/s41235-017-0052-9>
- Carr, M., Steiner, H. H., Kyser, B., & Biddlecomb, B. (2008). A comparison of predictors of early emerging gender differences in mathematics competency. *Learning and Individual Differences*, 18(1), 61–75. <https://doi.org/10.1016/j.lindif.2007.04.005>
- Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. Cambridge University Press. [https://market.android.com/details?id=book-jp9dt4\\_0\\_cIC](https://market.android.com/details?id=book-jp9dt4_0_cIC)
- Casey, M. B., Nuttall, R., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology*, 31(4), 697–705. <https://doi.org/10.1037/0012-1649.31.4.697>

- Casey, M. B., Nuttall, R. L., & Pezaris, E. (1997). Mediators of gender differences in mathematics college entrance test scores: a comparison of spatial skills with internalized beliefs and anxieties. *Developmental Psychology*, 33(4), 669–680. <https://doi.org/10.1037/0012-1649.33.4.669>
- Casey, B. M., Dearing, E., Vasilyeva, M., Ganley, C. M., & Tine, M. (2011). Spatial and numerical predictors of measurement performance: The moderating effects of community income and gender. *Journal of Educational Psychology*, 103(2), 296–311. <https://doi.org/10.1037/a0022516>
- Casey, B. M., Pezaris, E., Fineman, B., Pollock, A., Demers, L., & Dearing, E. (2015). A longitudinal analysis of early spatial skills compared to arithmetic and verbal skills as predictors of fifth-grade girls' math reasoning. *Learning and Individual Differences*, 40, 90–100. <https://doi.org/10.1016/j.lindif.2015.03.028>
- Casey, B. M., Lombardi, C. M., Pollock, A., Fineman, B., & Pezaris, E. (2017). Girls' spatial skills and arithmetic strategies in first grade as predictors of fifth-grade analytical math reasoning. *Journal of Cognition and Development: Official Journal of the Cognitive Development Society*, 18(5), 530–555. <https://doi.org/10.1080/15248372.2017.1363044>
- Cattell, R. B. (1971). *Abilities: Their Structure, Growth, and Action*. Houghton Mifflin.
- Ceci, S. J., & Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences of the United States of America*, 108(8), 3157–3162. <https://doi.org/10.1073/pnas.1014871108>
- Cheng, Y. L., & Mix, K. S. (2014). Spatial training improves children's mathematics ability. *Journal of Cognition and Development*, 15(1), 2–11.
- Cheung, M.W.L. (2015). {metaSEM}: An R package for meta-analysis using structural equation modeling. *Frontiers in Psychology*, 5, 1–7. <https://doi.org/10.3389/fpsyg.2014.01521>
- Cheung, M. W.-L., & Chan, W. (2005). Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods*, 10(1), 40–64. <https://doi.org/10.1037/1082-989X.10.1.40>
- Cirino, P. T., Tolar, T. D., Fuchs, L. S., & Huston-Warren, E. (2016). Cognitive and numerosity predictors of mathematical skills in middle school. *Journal of Experimental Child Psychology*, 145, 95–119. <https://doi.org/10.1016/j.jecp.2015.12.010>
- Clifford, E. (2008). *Visual-spatial processing and mathematics achievement: the predictive ability of the visual-spatial measures of the Stanford-Binet intelligence scales, and the Wechsler intelligence scale for children*. University of South Dakota. <http://search.proquest.com/openview/7909403d7b16e45eca563dbc3826e3eb/1?pq-origsite=gscholar&cbl=18750&diss=y>
- College Board (2020). *SAT Math Test*. <https://collegereadiness.collegeboard.org/sat/inside-the-test/math>
- Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, 7(12), 547–552. <https://doi.org/10.1016/j.tics.2003.10.005>
- Corsi, P. M. (1972). *Human Memory and the Medial Temporal Region of the Brain*. McGill University.
- Cox, J. W. (1928). *Mechanical aptitude*. Methuen.
- Cui, J., Zhang, Y., Cheng, D., Li, D., & Zhou, X. (2017). Visual form perception can be a cognitive correlate of lower-level math categories for teenagers. *Frontiers in Psychology*, 8, 1336. <https://doi.org/10.3389/fpsyg.2017.01336>
- Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, 35(1), 13–21.
- Delgado, A. R., & Prieto, G. (2004). Cognitive mediators and sex-related differences in mathematics. *Intelligence*, 32(1), 25–32. [https://doi.org/10.1016/S0160-2896\(03\)00061-8](https://doi.org/10.1016/S0160-2896(03)00061-8)
- Edens, K., & Potter, E. (2007). The relationship of drawing and mathematical problem solving: Draw for math tasks. *Studies in Art Education*, 48(3), 282–298. <https://doi.org/10.1080/00393541.2007.11650106>
- Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629–634. <https://doi.org/10.1136/bmj.315.7109.629>
- Ekstrom, R. B., French, J. W., Harman, H., & Derman, D. (1976). *Kit of Factor-referenced cognitive tests* (revised edition). Educational Testing Service, Princeton, NJ.
- Embretson, S. E. (1995). The role of working memory capacity and general control processes in intelligence. *Intelligence*, 20(2), 169–189. [https://doi.org/10.1016/0160-2896\(95\)90031-4](https://doi.org/10.1016/0160-2896(95)90031-4)
- Flanagan, D. P. & Kaufman, A. S. (2004). *Essentials of WISC-IV Assessment*. John Wiley & Sons.
- Frick, A. (2019). Spatial transformation abilities and their relation to later mathematics performance. *Psychological Research*, 83(7), 1465–1484. <https://doi.org/10.1007/s00426-018-1008-5>
- Fry, A. F. & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science*, 7(4), 237–241.
- Ganley, C. M., & Vasilyeva, M. (2011). Sex differences in the relation between math performance, spatial skills, and attitudes. *Journal of Applied Developmental Psychology*, 32(4), 235–242. <https://doi.org/10.1016/j.appdev.2011.04.001>
- Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A five-year longitudinal study. *Developmental Psychology*, 47(6), 1539–1552. <https://doi.org/10.1037/a0025510>
- Geer, E. A., & Ganley, C. M. (2021). Sex differences in social and spatial perspective taking: A replication and extension of Tarampi et al., (2016). Unpublished manuscript.
- Geer, E. A., Quinn, J. M., & Ganley, C. M. (2019). Relations between spatial skills and math performance in elementary school children: A longitudinal investigation. *Developmental Psychology*, 55(3), 637–652. <https://doi.org/10.1037/dev0000649>
- Gilles, P.-Y., & Bailleux, C. (2001). Personality traits and abilities as predictors of academic achievement. *European Journal of Psychology of Education*, 16(1), 3–15. <https://doi.org/10.1007/BF03172991>
- Gilligan, K. A., Flouri, E., & Farran, E. K. (2017). The contribution of spatial ability to mathematical achievement in middle childhood. *Journal of Experimental Child Psychology*, 163, 107–125. <https://doi.org/10.1016/j.jecp.2017.04.016>
- Gottfredson, L. S., & Deary, I. J. (2004). Intelligence predicts health and longevity, but why? *Current Directions in Psychological Science*, 13(1), 1–4. <https://doi.org/10.1111/j.0963-7214.2004.01301001.x>
- Green, C. T., Bunge, S. A., Briones Chiongbian, V., Barrow, M., & Ferrer, E. (2017). Fluid reasoning predicts future mathematical performance among children and adolescents. *Journal of Experimental Child Psychology*, 157, 125–143. <https://doi.org/10.1016/j.jecp.2016.12.005>
- GrÉGoire, J., & Van Der Linden, M. (1997). Effect of age on forward and backward digit spans. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, 4(2), 140–149. <https://doi.org/10.1080/13825589708256642>
- Gunderson, E. A., Ramirez, G., Beilock, S. L., & Levine, S. C. (2012). The relation between spatial skill and early number knowledge: the role of the linear number line. *Developmental Psychology*, 48(5), 1229–1241. <https://doi.org/10.1037/a0027433>
- Haciomeroglu, E. S. (2015). The role of cognitive ability and preferred mode of processing in students' calculus performance. *Eurasia Journal of Mathematics, Science & Technology Education*, 11(5), 1165–1179. <http://www.ejmste.com/pdf/51586-11492?filename=The%20Role%20of%20Cognitive.pdf>
- Haciomeroglu, E. S. (2016). Object-spatial visualization and verbal cognitive styles, and their relation to cognitive abilities and mathematical performance. *Educational Sciences: Theory and Practice*, 16(3), 987–1003. <https://eric.ed.gov/?id=EJ1115147>

- Halford, G. S., Wilson, W. H., Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, 21(6), 803–831.
- Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, 8(1), 1–51. <https://doi.org/10.1111/j.1529-1006.2007.00032.x>
- Hanline, M. F., Milton, S., & Phelps, P. C. (2010). The relationship between preschool block play and reading and maths abilities in early elementary school: A longitudinal study of children with and without disabilities. *Early Child Development and Care*, 180(8), 1005–1017. <https://doi.org/10.1080/03004430802671171>
- Harter, C. A. & Ku, H. (2008) The effects of spatial contiguity within computer-based instruction of group personalized two-step mathematics word problems. *Computers in Human Behavior*, 24(4), 1668–1685. <https://doi.org/10.1016/j.chb.2007.06.006>
- Hawes, Z. & Ansari, D. (2020). What explains the relationship between spatial and mathematical skills? A review of evidence from brain and behavior. *Psychonomic Bulletin & Review*, 27, 465–482. <https://doi.org/10.3758/s13423-019-01694-7>
- Hawes, Z., Moss, J., Caswell, B., & Poliszczuk, D. (2015). Effects of mental rotation training on children’s spatial and mathematics performance: A randomized controlled study. *Trends in Neuroscience and Education*, 4(3), 60–68.
- Hawes, Z., Moss, J., Caswell, B., Seo, J., Ansari, D. (2019). Relations between numerical, spatial, and executive function skills and mathematical achievement: A latent-variable approach. *Cognitive Psychology*, 109, 68–90.
- Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, 1(1), 39–65. <https://doi.org/10.1002/jrsm.5>
- Hegarty, M., & Kozhevnikov, M. (1999). Types of visual–spatial representations and mathematical problem solving. *Journal of Educational Psychology*, 91(4), 684–689. <https://doi.org/10.1037/0022-0663.91.4.684>
- Holyoak, K. J. (2012). Analogy and relational reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning* (pp. 234–259). Oxford University Press.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. In *Nature Reviews Neuroscience* (Vol. 6, Issue 6, pp. 435–448). <https://doi.org/10.1038/nrn1684>
- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: a meta-analysis. *Psychological Bulletin*, 107(2), 139–155. <https://doi.org/10.1037/0033-2909.107.2.139>
- Imbo, I., Vandierendonck, A., & De Rammelaere, S. (2007). The role of working memory in the carry operation of mental arithmetic: number and value of the carry. *Quarterly Journal of Experimental Psychology*, 60(5), 708–731. <https://doi.org/10.1080/17470210600762447>
- Jacob, R. & Parkinson, J. (2015). The potential for school-based interventions that target executive function to improve academic achievement: A review. *Review of Educational Research*, 85(4), 512–552.
- Jensen, A. R. (1998). *The Science of Mental Ability*. Praeger Publishers.
- Johnson, C. M. (2017). *Mathematics performance of sixth-grade students in single-gender and mixed classrooms in a large urban school system* (Publication No. 10604090) [Doctoral dissertation, Grand Canyon University]. ProQuest Dissertation Publishing.
- Kirby, J. R., & Boulter, D. R. (1999). Spatial ability and transformational geometry. *European Journal of Psychology of Education*, 14(2), 283. <https://doi.org/10.1007/BF03172970>
- Klingberg, T. (2006). Development of a superior frontal-intraparietal network for visuo-spatial working memory. *Neuropsychologia*, 44(11), 2171–2177.
- Kovacs, K., & Conway, A. R. A. (2016). Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry*, 27(3), 151–177. <https://doi.org/10.1080/1047840X.2016.1153946>
- Krawczyk, D. C. (2012). The cognition and neuroscience of relational reasoning. *Brain Research*, 1428, 13–23.
- Kyttälä, M., & Lehto, J. E. (2008). Some factors underlying mathematical performance: The role of visuospatial working memory and non-verbal intelligence. *European Journal of Psychology of Education*, 23(1), 77. <https://doi.org/10.1007/BF03173141>
- Kyttälä, M., Aunio, P., Lehto, J. E., Van Luit, J., & Hautamäki, J. (2003). Visuospatial working memory and early numeracy. *Educational and Child Psychology*, 20(3), 65–76.
- Li, Y., & Geary, D. C. (2013). Developmental gains in visuospatial memory predict gains in mathematical achievement. In *PLoS ONE* (Vol. 8, Issue 7, p. e70160). <https://doi.org/10.1371/journal.pone.0070160>
- Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: a meta-analysis. *Child Development*, 56(6), 1479–1498. <https://www.ncbi.nlm.nih.gov/pubmed/4075870>
- Lombardi, C. M., Casey, B. M., Pezaris, E., Shadmehr, M., & Jong, M. (2019). Longitudinal analysis of associations between 3-D mental rotation and mathematics reasoning skills during middle school: Across and within genders. *Journal of Cognition and Development: Official Journal of the Cognitive Development Society*, 20(4), 487–509. <https://doi.org/10.1080/15248372.2019.1614592>
- Lowrie, T., Logan, T., & Ramful, A. (2017). Visuospatial training improves elementary students’ mathematics performance. *British Journal of Educational Technology: Journal of the Council for Educational Technology*. <https://doi.org/10.1111/bjep.12142>
- Lowrie, T., Logan, T., & Hegarty, M. (2019). The influence of spatial visualization training on students’ spatial reasoning and mathematics performance. *Journal of Cognition and Development*. <https://doi.org/10.1080/15248372.2019.1653298>
- McGrew, K. S., & Wendling, B. J. (2010). Cattell–Horn–Carroll cognitive-achievement relations: What we have learned from the past 20 years of research. *Psychology in the Schools*, 47(7), 651–675. <https://doi.org/10.1002/pits.20497>
- McKenzie, B., Bull, R., & Gray, C. (2003). The effects of phonological and visual-spatial interference on children’s arithmetical performance. *Educational and Child Psychology*, 20(3), 93–108.
- Merriam Webster (n.d.) *Definition of MATHEMATICS*. Retrieved November 18, 2020, from <https://www.merriam-webster.com/dictionary/mathematics>
- Mix, K. S. (2019). Why are spatial skill and mathematics related? In *Child Development Perspectives* (Vol. 13, Issue 2, pp. 121–126). <https://doi.org/10.1111/cdep.12323>
- Mix, K. S., & Cheng, Y.-L. (2012). The relation between space and math: developmental and educational implications. *Advances in Child Development and Behavior*, 42, 197–243. <https://doi.org/10.1016/b978-0-12-394388-0.00006-x>
- Mix, K. S., Levine, S. C., Cheng, Y.-L., Young, C., Hambrick, D. Z., Ping, R., & Konstantopoulos, S. (2016). Separate but correlated: The latent structure of space and mathematics across development. *Journal of Experimental Psychology: General*, 145(9), 1206–1227. <https://doi.org/10.1037/xge0000182>
- Mix, K. S., Levine, S. C., Cheng, Y.-L., Young, C. J., Hambrick, D. Z., & Konstantopoulos, S. (2017). The latent structure of spatial skills and

- mathematics: A replication of the two-factor model. *Journal of Cognition and Development: Official Journal of the Cognitive Development Society*, 18(4), 465–492. <https://doi.org/10.1080/15248372.2017.1346658>
- Miyake, A., & Shah, P. (1999). Models of working memory: Mechanisms of active maintenance and executive control. Cambridge University Press.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. <https://doi.org/10.1006/cogn.1999.0734>
- Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology. General*, 130(4), 621–640. <https://doi.org/10.1037/0096-3445.130.4.621>
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4, 1. <https://doi.org/10.1186/2046-4053-4-1>
- Möhring, W., Newcombe, N. S., & Frick, A. (2015). The relation between spatial thinking and proportional reasoning in preschoolers. *Journal of Experimental Child Psychology*, 132, 213–220. <https://doi.org/10.1016/j.jecp.2015.01.005>
- Morosanova, V. I., Fomina, T. G., Kovas, Y., & Bogdanova, O. Y. (2016). Cognitive and regulatory characteristics and mathematical performance in high school students. *Personality and Individual Differences*, 90, 177–186. <https://doi.org/10.1016/j.paid.2015.10.034>
- Mullis, I. V. S., Martin, M. O., Foy, P., Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*. International Association for the Evaluation of Educational Achievement. <https://eric.ed.gov/?id=ed544554>
- National Center for Education Statistics (2018). Annual reports program 2018. [https://nces.ed.gov/programs/digest/d18/figures/fig\\_01.asp?referrer=figures](https://nces.ed.gov/programs/digest/d18/figures/fig_01.asp?referrer=figures)
- Neale, M. C., & Miller, M. B. (1997). The use of likelihood-based confidence intervals in genetic models. *Behavior genetics*, 27(2), 113–120.
- Newcombe, N. S. (2010). Picture this: Increasing math and science learning by improving spatial thinking. *American Educator*, 34(2), 29. <https://eric.ed.gov/?id=EJ889152>
- Newcombe, N. S., & Shipley, T. F. (2015). Thinking about spatial thinking: New typology, new assessments. *Studying Visual and Spatial Reasoning for Design Creativity*, 179–192. [https://doi.org/10.1007/978-94-017-9297-4\\_10](https://doi.org/10.1007/978-94-017-9297-4_10)
- Novack, M. A., Congdon, E. L., Hemani-Lopez, N., & Goldin-Meadow, S. (2014). From action to abstraction: using the hands to learn math. *Psychological Science*, 25(4), 903–910. <https://doi.org/10.1177/0956797613518351>
- Nuttall, R. L., Casey, M. B., & Pizaris, E. (2005). Spatial ability as a mediator of gender differences on mathematics tests: A biological-environmental framework. In A. M. Gallagher (Ed.), *Gender differences in mathematics: An integrative psychological approach*. (Vol. 351, pp. 121–142). Cambridge University Press, xvi. <https://psycnet.apa.org/fulltext/2005-04568-006.pdf>
- Olver, A. (2013). *Investigating early spatial and numerical skills in junior kindergarten children learning in an inquiry-and play-based environment* [Master's Thesis, University of Toronto]. Tspace. <https://tspace.library.utoronto.ca/handle/1807/42637>
- Oostermeijer, M., Boonen, A. J. H., & Jolles, J. (2014). The relation between children's constructive play activities, spatial ability, and mathematical word problem-solving performance: a mediation analysis in sixth-grade students. *Frontiers in Psychology*, 5, 782. <https://doi.org/10.3389/fpsyg.2014.00782>
- Organization for Economic Cooperation and Development (2012). *Programme for International Student Assessment (PISA) Results from PISA 2012*. <https://www.oecd.org/unitedstates/PISA-2012-results-US.pdf>
- Organization for Economic Development (2016). *Country Note: Key Findings from PISA 2015 for the United States*. [Oecd.org/pisa/PISA-2015-United-States.pdf](https://www.oecd.org/pisa/PISA-2015-United-States.pdf)
- Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods*, 47(4), 1343–1355. <https://doi.org/10.3758/s13428-014-0543-2>
- Paterson, D. G., Elliot, R., Anderson, L. D., Toops, H. A., & Heidbreder, E. (1930). Minnesota mechanical ability tests: The report of a research investigation subsidized by the committee on human migrations of the national research council and conducted in the department of psychology of the University of Minnesota. University of Minnesota Press.
- Paunonen, S. V., & Hong, R. Y. (2010). Self-efficacy and the prediction of domain-specific cognitive abilities. *Journal of Personality*, 78(1), 339–360. <https://doi.org/10.1111/j.1467-6494.2009.00618.x>
- Payne, C. M. (2008). *So Much Reform, So Little Change: The Persistence of Failure in Urban Schools*. Harvard Education Press.
- Peng, P., Yang, X., & Meng, X. (2017). The relation between approximate number system and early arithmetic: The mediation role of numerical knowledge. *Journal of Experimental Child Psychology*, 157, 111–124. <https://doi.org/10.1016/j.jecp.2016.12.011>
- Peng, P., Wang, T., Wang, C., Lin, X. (2019). A meta-analysis on the relation between fluid intelligence and reading/mathematics: Effects of tasks, age, and social economics status. *Psychological Bulletin*, 145(2), 189–236.
- Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test: different versions and factors that affect performance. *Brain and Cognition*, 28(1), 39–58. <https://doi.org/10.1006/brcg.1995.1032>
- Pustejovsky, J. E. (2020). clubSandwich: Cluster-Robust(Sandwich) Variance Estimators with Small-Sample Corrections. R package version 0.5.1. <https://CRAN.R-project.org/package=clubSandwich>
- Pustejovsky, J. E., & Rodgers, M. A. (2019). Testing for funnel plot asymmetry of standardized mean differences. *Research Synthesis Methods*, 10(1), 57–71. <https://doi.org/10.1002/jrsm.1332>
- Pustejovsky, J. E., & Tipton, E. (2021). Meta-analysis with robust variance estimation: Expanding the range of working models. Manuscript under review. *Osfi.io*. <https://osfi.io/vyfcj/download>
- Pyke, C. L. (2003). The use of symbols, words, and diagrams as indicators of mathematical cognition: A causal model. *Journal for Research in Mathematics Education*, 34(5), 406–432. <https://doi.org/10.2307/30034794>
- Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, 20(2), 110–122. <https://doi.org/10.1016/j.lindif.2009.10.005>
- Reuhkala, M. (2001). Mathematical skills in ninth graders: Relationship with visuo-spatial abilities and working memory. *Educational Psychology Review*, 21(4), 387–399. <https://doi.org/10.1080/01443410120090786>
- Rodán, A., Contreras, M. J., Elosúa, M. R., & Gimeno, P. (2016). Experimental but not sex differences of a mental rotation training program on adolescents. *Frontiers in Psychology*, 7, 1050. <https://doi.org/10.3389/fpsyg.2016.01050>
- Rodgers, M. A., & Pustejovsky, J. E. (2020). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods*. <https://doi.org/10.1037/met0000300>

- Schmitt, S. A., Korucu, I., Napoli, A. R., Bryant, L. M., & Purpura, D. J. (2018). Using block play to enhance preschool children's mathematics and executive functioning: A randomized controlled trial. *Early Childhood Research Quarterly, 44*, 181–191.
- Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 99–144). The Guilford Press.
- Seitz, K., & Schumann-Hengsteler, R. (2002). Phonological loop and central executive processes in mental addition and multiplication. *Psychologische Beiträge, 44*(2), 275–303.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(3972), 701–703. <https://doi.org/10.1126/science.171.3972.701>
- Simms, V., Clayton, S., Cragg, L., Gilmore, C., & Johnson, S. (2016). Explaining the relationship between number line estimation and mathematical achievement: The role of visuomotor integration and visuospatial skills. *Journal of Experimental Child Psychology, 145*, 22–33. <https://doi.org/10.1016/j.jecp.2015.12.004>
- Skagerlund, K., & Träff, U. (2016). Processing of space, time, and number contributes to mathematical abilities above and beyond domain-general cognitive abilities. *Journal of Experimental Child Psychology, 143*, 85–101. <https://doi.org/10.1016/j.jecp.2015.10.016>
- Sorby, S. A. (2007). Developing 3D spatial skills for engineering students. *Australasian Journal of Engineering Education, 13*(1), 1–11. <https://doi.org/10.1080/22054952.2007.11463998>
- Sorby, S., Casey, B., Veurink, N., & Dulaney, A. (2013). The role of spatial training in improving spatial and calculus performance in engineering students. *Learning and Individual Differences, 26*, 20–29. <https://doi.org/10.1016/j.lindif.2013.03.010>
- Stannard, L., Wolfgang, C. H., Jones, I., & Phelps, P. (2001). A longitudinal study of the predictive relations among construction play and mathematical achievement. *Early Child Development and Care, 167*(1), 115–125. <https://doi.org/10.1080/0300443011670110>
- Stevenson, H. W., Parker, T., Wilkinson, A., Hegion, A., & Fish, E. (1976). Longitudinal study of individual differences in cognitive development and scholastic achievement. *Journal of Educational Psychology, 68*(4), 377–400. <https://doi.org/10.1037/0022-0663.68.4.377>
- Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 96*(3), 471–491. <https://doi.org/10.1037/0022-0663.96.3.471>
- Tartre, L. A., & Fennema, E. (1995). Mathematics achievement and gender: A longitudinal study of selected cognitive and affective variables [grades 6–12]. *Educational Studies in Mathematics, 28*(3), 199–217. <https://doi.org/10.1007/bf01274173>
- Taub, G. E., Floyd, R. G., Keith, T. Z., & McGrew, K. S. (2008). Effects of general and broad cognitive abilities on mathematical achievement. *School Psychology Quarterly, 23*, 187–198.
- Thurstone, L. L., & Thurstone, T. G. (1941). Factorial studies of intelligence. *Psychometric Monographs, 2*, 94. <https://psycnet.apa.org/fulltext/1941-03743-001.pdf>
- Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods, 20*(3), 375–393. <https://doi.org/10.1037/met0000011>
- Tolar, T. D., Lederberg, A. R., & Fletcher, J. M. (2009). A structural model of algebra achievement: computational fluency and spatial visualization as mediators of the effect of working memory on algebra achievement. *Educational Psychology Review, 29*(2), 239–266. <https://doi.org/10.1080/01443410802708903>
- Uttal, D. H., Miller, D. I., & Newcombe, N. S. (2013). Exploring and enhancing spatial thinking. *Current Directions in Psychological Science, 22*(5), 367–373. <https://doi.org/10.1177/0963721413484756>
- Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology, 28*(2), 187–204. <https://www.sciencedirect.com/science/article/pii/S0361476X02000073>
- van Garderen, D. (2006). Spatial visualization, visual imagery, and mathematical problem solving of students with varying abilities. *Journal of Learning Disabilities, 39*(6), 496–506. <https://doi.org/10.1177/00222194060390060201>
- Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*(2), 599–604. <https://doi.org/10.2466/pms.1978.47.2.599>
- Vendetti, M. S., & Bunge, S. A. (2014). Evolutionary and developmental changes in the lateral frontoparietal network: a little goes a long way for higher-level cognition. *Neuron, 84*(5), 906–917. <https://doi.org/10.1016/j.neuron.2014.09.035>
- Ventura, M., Shute, V., Wright, T., & Zhao, W. (2013). An investigation of the validity of the virtual spatial navigation assessment. *Frontiers in Psychology, 4*, 852. <https://doi.org/10.3389/fpsyg.2013.00852>
- Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., & Chang, A. (2014). Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. *Child Development, 85*(3), 1062–1076. <https://doi.org/10.1111/cdev.12165>
- Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., & Newcombe, N. (2017). Links between spatial and mathematical skills across the preschool years. *Monographs of Society for Research in Child Development, 82*(1), 7–30. <https://doi.org/10.1111/mono.12280>
- 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><https://lirias.kuleuven.be/1059637?limo=0>
- Voyer, D. (1998). Mathematics, gender, spatial performance, and cerebral organization: A suppression effect in talented students. *Roeper Review, 20*(4), 251–258. <https://doi.org/10.1080/02783199809553902>
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*(4), 817–835. <https://doi.org/10.1037/a0016127>
- Walberg, H. J. (1984). Improving the productivity of America's schools. *Educational Leadership, 41*, 19–27.
- Wang, M.-T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review, 29*(1), 119–140. <https://doi.org/10.1007/s10648-015-9355-x>
- Wechsler, D. (1976). *Wechsler intelligence scale for children-revised*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1981). WAIS-R manual: Wechsler adult intelligence scale-revised. Psychological Corporation.
- Wechsler, D. (2012). Wechsler preschool and primary scale of intelligence—fourth edition. The Psychological Corporation San Antonio.
- Wei, W., Yuan, H., Chen, C., & Zhou, X. (2012). Cognitive correlates of performance in advanced mathematics. *The British Journal of Educational Psychology, 82*(1), 157–181. <https://doi.org/10.1111/j.2044-8279.2011.02049.x>
- Wilson, S. J., Polanin, J. R., & Lipsey, M. W. (2016). Fitting meta-analytic structural equation models with complex datasets. *Research Synthesis Methods, 7*(2), 121–139. <https://doi.org/10.1002/jrsm.1199>
- Wolfgang, C. H., Stannard, L. L., & Jones, I. (2001). Block play performance among preschoolers as a predictor of later school achievement in mathematics. *Journal of Research in Childhood Education:*

- JRCE / Association for Childhood Education International*, 15(2), 173–180. <https://doi.org/10.1080/02568540109594958>
- Wong, W. I. (2017). The space-math link in preschool boys and girls: Importance of mental transformation, targeting accuracy, and spatial anxiety. *The British Journal of Developmental Psychology*, 35(2), 249–266. <https://doi.org/10.1111/bjdp.12161>
- Woodcock, R. W., McGrew, K. S., Mather, N., & Schrank, F. A. (2001). *Woodcock Johnson III Tests of Cognitive Abilities*. Riverside Publishing.
- Wrigley, J. (1958). Factorial nature of ability in elementary mathematics. *British Journal of Educational Psychology*, 28, 61–78.
- Xie, F., Zhang, L., Chen, X., & Xin, Z. (2019). Is spatial ability related to mathematical ability: a meta-analysis. *Educational Psychology Review*. <https://doi.org/10.1007/s10648-019-09496-y>
- Zhang, X., & Lin, D. (2017). Does growth rate in spatial ability matter in predicting early arithmetic competence? *Learning and Instruction*, 49, 232–241. <https://doi.org/10.1016/j.learninstruc.2017.02.003>
- Zhang, X., & Lin, D. (2018). Cognitive precursors of word reading versus arithmetic competencies in young Chinese children. *Early Childhood Research Quarterly*, 42, 55–65. <https://doi.org/10.1016/j.ecresq.2017.08.006>
- Zhang, X., Koponen, T., Räsänen, P., Aunola, K., Lerkkanen, M.-K., & Nurmi, J.-E. (2014). Linguistic and spatial skills predict early arithmetic development via counting sequence knowledge. *Child Development*, 85(3), 1091–1107. <https://doi.org/10.1111/cdev.12173>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.