Cognitive Psychology 109 (2019) 68–90

Contents lists available at ScienceDirect

Cognitive Psychology

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

Relations between numerical, spatial, and executive function skills and mathematics achievement: A latent-variable approach

Zachary Hawes[a,*], Joan Moss[b], Beverly Caswell[b], Jisoo Seo[b], Daniel Ansari[a]

[a] Department of Psychology, Brain and Mind Institute, University of Western Ontario, Canada
[b] Department of Applied Psychology and Human Development, Ontario Institute for Studies in Education, University of Toronto, Canada

ARTICLE INFO

Keywords:
Spatial skills
Spatial visualization
Numerical skills
Executive functions
Mathematics achievement
Latent-variable analysis (SEM)

ABSTRACT

Current evidence suggests that numerical, spatial, and executive function (EF) skills each play critical and independent roles in the learning and performance of mathematics. However, these conclusions are largely based on isolated bodies of research and without measurement at the latent variable level. Thus, questions remain regarding the latent structure and potentially shared and unique relations between numerical, spatial, EF, and mathematics abilities. The purpose of the current study was to (i) confirm the latent structure of the hypothesized constructs of numerical, spatial, and EF skills and mathematics achievement, (ii) measure their unique and shared relations with one another, and (iii) test a set of novel hypotheses aimed to more closely reveal the underlying nature of the oft reported space-math association. Our analytical approach involved latent-variable analyses (structural equation modeling) with a sample of 4- to 11-year-old children ($N = 316$, $M_{age} = 6.68$ years). Results of a confirmatory factor analysis demonstrated that numerical, spatial, EF, and mathematics skills are highly related, yet separable, constructs. Follow-up structural analyses revealed that numerical, spatial, and EF latent variables explained 84% of children's mathematics achievement scores, controlling for age. However, only numerical and spatial performance were unique predictors of mathematics achievement. The observed patterns of relations and developmental trajectories remained stable across age and grade (preschool – 4th grade). Follow-up mediation analyses revealed that numerical skills, but not EF skills, partially mediated the relation between spatial skills and mathematics achievement. Overall, our results point to spatial visualization as a unique and robust predictor of children's mathematics achievement.

1. Introduction

How do humans learn to think mathematically? What role do cognitive skills play in the ability to engage in abstract mathematical thought? During the past two decades, researchers from a wide variety of disciplines, including psychology, cognitive neuroscience, and education, have become increasingly interested in answering these and other related questions. This is due, in part, to the growing recognition of relations between mathematics skills and a range of desirable school and life outcomes (e.g., see Duncan et al., 2007; Parsons & Bynner, 2005). For instance, early mathematics skills strongly predict later mathematics achievement, as well as educational attainment more generally, and contribute to important life outcomes, such SES, health and personal well-being

* Corresponding author at: Department of Psychology, Brain and Mind Institute, University of Western Ontario, Western Interdisciplinary Research Building, Room 5180, 1151 Richmond Street North, London, ON N6A 5B7, Canada.
E-mail addresses: zhawes@gmail.com, zhawes@uwo.ca (Z. Hawes).

https://doi.org/10.1016/j.cogpsych.2018.12.002
Accepted 13 December 2018
Available online 04 January 2019

(Duncan et al., 2007; Parsons & Bynner, 2005; Ritchie & Bates, 2013). In short, there is a need to better understand factors that contribute to individual differences in the development of and achievement in mathematics.

Current evidence suggests that numerical, spatial, and executive function (EF) skills each play critical and independent roles in the learning and performance of mathematics (e.g., see Cragg & Gilmore, 2014; De Smedt, Noël, Gilmore, & Ansari, 2013; Mix & Cheng, 2012). However, these conclusions are largely based on isolated bodies of research and without measurement at the latent variable level. While prior research has examined latent relations between two of these constructs (e.g., spatial and mathematical abilities; Mix et al., 2016, 2017; spatial and EF abilities; Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001), relations between all four constructs has yet to be examined. Thus, questions remain regarding the extent to which these skills represent distinct constructs and whether numerical, spatial, and EF skills afford differentiated pathways to mathematics achievement. In this paper, we address this gap in the literature and examine the latent structure and interrelations between numerical, spatial, EF, and mathematics abilities in a sample of 4-to 11-year-olds.

In addition to providing a comprehensive cognitive model of children's mathematics achievement, this study was designed to more closely reveal insight into the underlying nature of the oft reported space-math association. Although there is extensive correlational evidence linking spatial and mathematical cognition, including decades of behavioral and neural research (Hubbard, Piazza, Pinel, & Dehaene, 2005; Mix & Cheng, 2012), relatively few efforts have been made to reveal potential mechanisms linking space and math. Here, we test the hypothesis that spatial visualization plays a critical role in mathematical problem solving and achievement. By testing a model of mathematical achievement that includes numerical, spatial, and EF factors, we were able to test and control for specific pathways connecting spatial visualization skills and mathematics achievement. This allowed us to examine the extent to which the space-math link is best explained by direct relations between spatial visualization and mathematics or whether the space-math link might be better explained by alternative mechanisms. Specifically, we test whether general intelligence and/or EF skills (including visual-spatial working memory) might better explain the space-math link. In addition, we test the hypothesis that spatial visualization skills are indirectly related to mathematics through basic numerical skills (e.g., see Gunderson, Ramirez, Beilock, & Levine, 2012; LeFevre et al., 2013). The findings related to these pathways are crucial in order to advance current theories of spatial and mathematical associations.

In the next section, we provide a more detailed review of space-math associations. We then operationalize spatial, numerical, and EF skills, as defined in the current study, and review evidence to suggest differentiated pathways from each of the targeted constructs to mathematics achievement.

2. Relations between spatial skills and mathematics

The scientific study of associations between spatial and mathematical thinking has a lengthy history, dating back to Sir Francis Galton's inquiries into the visualization of numerals in the late 1800′s (Galton, 1880). Indeed, a large body of research supports the finding that people with strong spatial skills also tend to do well in mathematics (Mix & Cheng, 2012). Of the various spatial skills identified, spatial visualization skills appear to play an especially important role in mathematics learning and achievement (Mix et al., 2016). Defined as the ability to generate, retrieve, maintain, and manipulate visual-spatial information (Lohman, 1996), spatial visualization skills have been linked to performance across a breadth of mathematics tasks, including arithmetic (Kyttälä & Lehto, 2008), word problems (Hegarty & Kozhevnikov, 1999), geometry (Delgado & Prieto, 2004), algebra (Tolar, Lederberg, & Fletcher, 2009), and highly advanced mathematics, including function theory, mathematical logic, and computational mathematics (Wei, Yuan, Chen, & Zhou, 2012). Moreover, the link between spatial visualization and mathematics is not limited to tasks that are inherently spatial, such as geometry or measurement. Research demonstrates that even basic number processing, such as comparing which digit is numerically larger (7 vs. 2), is closely associated with spatial visualization skills, such as mental rotation (Thompson, Nuerk, Moeller, & Kadosh, 2013; Viarouge, Hubbard, & McCandliss, 2014).

What explains the math-space link? One popular theory posits that numbers are represented spatially (de Hevia, Vallar, & Girelli, 2008). That is, humans come to conceive of and arrange numbers along a "mental number line," with small numbers belonging to the left and larger numbers extending to the right (Dehaene, Bossini, & Giraux, 1993). Interestingly, the left-to-right orientation of the mental number line appears to be culturally-specific and is reversed in cultures that read and write right-to-left (Göbel, Shaki, & Fischer, 2011). It is hypothesized that one of the ways children make sense of symbolic numbers is to learn to represent numbers according to their spatial relations (e.g., 1 and 2 are "close together," while 1 and 9 are "far apart."). Said differently, spatial skills are predicted to play an active role in the development of children's conceptualization and visualization of the various meanings of number (e.g., see Sella, Berteletti, Lucangeli, & Zorzi, 2017).

Moreover, the same spatial reasoning capacities that help ground various symbolic number relations may also help with the learning and representation of symbolic mathematics more generally (Lakoff & Núñez, 2000). That is, spatial visualization skills are hypothesized to play a critical role in one's ability to model, simulate, and form mathematical relationships. Accordingly, spatial visualization skills may provide a means to conceptualize numbers as ascending from left to right (akin to a number line), but also allow one to visualize and model various other mathematical transformations, such as the decomposition of 12 into a unit of 10 and 2. In sum, spatial visualization skills might represent one cognitive tool which children draw from to learn and make sense of not only basic numerical relations but novel and higher-level mathematics as well.

In the present study, we targeted spatial visualization skills by including measures of mental rotation and visual-spatial reasoning. These measures were selected as they were hypothesized to involve the recruitment of spatial visualization in the service of solving novel visual-spatial problems. More generally, we operationalized spatial visualization as a construct involving the "generation" or "creation" of visual-spatial solutions to problems (e.g., imagining how a folded and punctured piece of paper might appear when unfolded). We placed special emphasis on this aspect of spatial visualization (i.e., the need to generate mental images) in an effort to examine and better understand the hypothesized link between spatial thinking and mathematics described above. One reason for the consistent relations between spatial and mathematical thinking may be due to the shared task requirements involved in the generation and manipulation of visual-spatial representations and solutions to problems in both respective domains.

3. Relations between numerical skills and mathematics

Numbers – and their various relations with one another – lie at the heart of mathematics. As such, concerted efforts have been directed at studying how humans, and other species for that matter, perceive, represent, manipulate, and make sense of number. To date, the study of basic numerical skills has been approached through various paradigms that target the measurement of an individual's numerical magnitude representations. For example, the speed and accuracy in which individuals can compare and select the larger of two numerical magnitudes (5 vs. 3 or ⁙ vs. ⁚) is a commonly used approach to assess the precision of an individual's mental representation of number (Siegler, 2016). There is an extensive body of research linking individual differences in magnitude comparison tasks and various measures of mathematics. In general, children and adults who are faster and more accurate at comparing numerical symbols (5 vs. 3) and nonsymbolic number (⁙ vs. ⁚), tend to also do better on higher-level mathematical tasks, such as arithmetic (Schneider et al., 2017).

Another important marker of an individual's basic number skills relates to their understanding and processing of numbers as ordered sequences (i.e., ordinality). Performance on ordinality tasks, typically assessed by the speed and accuracy in which an individual can recognize ordered numerical sequences (e.g., 4-5-6; 5-7-9), are thought to index the strength of an individual's associations of numerical relations. Research indicates positive relations between children and adults' ordinality skills and mathematics performance (Lyons, Price, Vaessen, Blomert, & Ansari, 2014; Lyons, Vogel, & Ansari, 2016).

Taken together, current research in the field of numerical cognition point to magnitude and ordinal processing skills as foundational numerical competencies with strong links to more formal mathematics. For this reason, the current study included measures of children's magnitude comparison (symbolic and nonsymbolic) and ordinality skills as key indicators of the targeted construct of numerical ability.

The hypothesized causal link between basic numerical skills and higher-level mathematics is rather straightforward: Mathematics is a particularly hierarchical subject, where earlier learned concepts and skills are needed to give rise to new and more advanced mathematical knowledge, and thus, basic numerical skills represent and serve the role of fundamental building blocks.

4. Relations between executive functions and mathematics

The last decade has seen a sharp rise in research linking executive functioning (EF) and mathematics achievement (e.g., see Cragg & Gilmore, 2014). Although definitions vary, EF is most typically defined as a suite of highly related but separable cognitive control abilities that includes working memory, inhibitory control, and shifting or flexible attention (Friedman & Miyake, 2017; Miyake et al., 2000). Each sub-component has been found to both concurrently and longitudinally predict mathematics achievement (Cragg & Gilmore, 2014). Working memory, in particular, has been found to be a consistent predictor of mathematics (Friso-van den Bos, van der Ven, Kroesbergen, & van Luit, 2013; Fuchs et al., 2010), with especially strong relations between visual-spatial working memory (VSWM) and mathematics performance (Reuhkala, 2001). Despite evidence and theory to suggest that all three of these components are related and represent a unified construct (i.e., EF; Miyake et al., 2000), there is a scarcity of research studying relations between EF and mathematics at the latent variable level. Thus, it remains to be shown how EF – as a unified construct – relates to mathematics achievement.

In the present study, we attempted to measure EF by including measures of VSWM as well as a single behavioral measure of EF (i.e., the Head-Toes-Knees-Shoulders task; Ponitz, McClelland, Matthews, & Morrison, 2009); a comprehensive measure thought to tap into each sub-component of EF, but most notably inhibitory control (McClelland & Cameron, 2012). The decision to target VSWM also allowed us to test the extent to which spatial ability and VSWM represent potentially distinct constructs.

Given that many mathematics tasks are complex and involve multi-step solutions, EF skills have been theorized to play a critical, if not causal, role in mathematics learning and performance. Moreover, different components of EF have been proposed to play unique roles in the service of different mathematical goals. For example, working memory is called upon to remember the specifics of a given problem as well as to temporarily hold partially completed solutions in mind while performing other aspects of the problem. Inhibitory control is needed to ignore or suppress certain responses in favor of other more appropriate responses (e.g., inhibiting knowledge of whole number operations when dealing with fractions). Shifting or flexible attention is recruited when switching between different operations, such as problems that involve both addition and subtraction. From these examples, it can be seen how individual differences in EF skills may constrain one's capacity to learn and carryout various mathematical tasks.

5. Main questions and hypotheses

5.1. Rationale for testing a four-factor model

The first goal of this study was to carry out a confirmatory factor analysis (CFA) to examine the degree of evidence in favor of a four-factor model, with factors corresponding to numerical, spatial, EF, and mathematical constructs. Based on theory and a small body of empirical research suggesting the existence of highly-related, but *distinct* constructs, we had reason to expect similar findings.

For example, Mix et al. (2016) found evidence to suggest that spatial and mathematical abilities represent separate, but highly overlapping, constructs in a cross-sectional sample of children aged 5–13. In adults, Miyake et al. (2001) demonstrated that spatial abilities and EFs are highly related factors and distinguishable to the extent that the spatial abilities measured were theorized to involve high executive demands. For example, latent variables related to spatial visualization and EF skills were highly related to one another (0.91) compared to relations between spatial perceptual speed and EFs (0.43). The extent to which spatial abilities and EFs represent separate factors in children is currently unknown. The present study aims to shed light on this issue.

To our knowledge, researchers have yet to examine the latent structure of basic numerical skills (e.g., basic understandings of numerical symbols and their associated magnitudes) and its relations with constructs related to spatial, EF, and mathematical abilities. However, prior research using single indicator variables has revealed consistent relations between basic numerical and spatial skills (e.g., see Newcombe, Levine, & Mix, 2015), as well as basic numerical skills and EFs (e.g., see Cragg, Keeble, Richardson, Roome, & Gilmore, 2017). Effect sizes related to these studies are typically in the moderate range and provide reason to suspect that numerical skills will share both overlapping and unique variance with spatial and EF skills. However, the extent to which basic numerical skills and mathematics achievement represent distinct constructs remains an open question. Moreover, if such a distinction does exist, what construct might spatial ability share more variance with? How might EF skills modify these relations?

With these questions in mind, we examined the latent structure of constructs related to numerical, spatial, EF, and more general mathematical skills. Given empirical evidence suggesting overlapping but unique relations between each construct, along with general consensus that each construct does indeed refer to something specific, we predicted that results of a CFA would offer support for a four-factor model. However, it should be noted that this is the first CFA that we are aware of that tests evidence for all four factors in the same model. It is possible, given their high associations with one another, that a single factor (i.e., general intelligence or *g*) might emerge as the best model fit of the data. Thus, to rule out this possibility, we also ran an exploratory analysis of a single-factor (*g*) model for comparison purposes.

5.2. Which construct does visual-spatial working memory belong to: spatial or EF?

Another follow-up objective of testing the four-factor model was to examine the extent to which measures of visual-spatial working memory (VSWM) load more closely on the EF versus spatial factor. Currently, the decision to classify VSWM as a marker of spatial ability or EF is up to individual researchers and it is not always clear whether such a decision is theoretically guided or post hoc. Here, we make the prediction that VSWM is better characterized as an indicator of EF than spatial visualization ability. This prediction was theoretical and made a priori based on the following criteria. Although both constructs are similar in that both are presumed to place heavy demands on cognitive control, there are some important distinctions in task requirements that potentially result in differential relations with mathematics. The EF measures were selected based on their involvement of working memory and inhibitory control. These measures were selected as “recall-based” measures; they require the storage and recall of information. In contrast, the spatial reasoning measures were selected based on their heavy demands on spatial perception, reasoning, and most notably, spatial visualization. More specifically, each spatial task required participants to reason and visualize solutions to problems involving parts of objects in relation to their whole. Critically, the spatial measures are distinct from the EF measures in that they are “prospective” or “generative” in nature. Thus, the spatial measures place a low demand on recall of information and place a heavy demand on the visualization or modeling of problems and their solutions. As detailed further below, loading VSWM on the EF factor also allowed us to test the important question of whether spatial visualization makes unique contributions to mathematics over and above EF skills.

5.3. Differentiated pathways to mathematics achievement

Given sufficient evidence for the existence of a four-factor model, our second objective was to test the shared and unique contributions of each predictor variable with mathematics achievement. As reviewed above, separate bodies of research have identified numerical, spatial, and EF skills as robust and consistent predictors of mathematics achievement. For this reason, we expected the combination of factors to explain a large proportion of variance in children’s mathematics performance. To our knowledge, this is the first study to include each construct within the same model and to simultaneously examine the extent to which each cognitive construct uniquely relates to mathematics achievement. Therefore, it is currently unknown how each variable relates to one another and potentially afford differential pathways to mathematics achievement. However, as reviewed above, each cognitive construct in the current study has been posited to differentially explain individual differences in mathematical performance. A metaphor of building a house serves as an example: Whereas basic numerical skills represent the fundamental building blocks (bricks), spatial visualization skills are considered a tool in which to manipulate and assemble the bricks, and EF skills place certain constraints, such as rules and regulations, on the building process.

5.4. Does age moderate potential relations between constructs?

As a follow-up to the above objective and analyses, we were interested in testing the extent to which age might moderate the observed relations. One reason for targeting the selected age-range (4- to 11-year-olds) was to better understand how each one of these foundational skills develop and potentially interact with one another across the early to middle childhood years. While separate bodies of research suggest that each construct undergoes rapid development during this time frame (e.g., see Mix, Huttenlocher, & Levine, 2002; Newcombe & Huttenlocher, 2003; Zelazo, Carlson, & Kesek, 2008), we currently know very little about the potential influence of age on these relations. Given that children's numerical, spatial, and EF skills have all been posited to play a critical role in children's mathematical development, it is important to better understand the potential impact that age might have on these various relations. This information, in turn, may be useful when designing educational interventions. Thus, an important question concerns the extent to which relations between constructs remain consistent across time or show evidence of change during specific periods of development.

5.5. Uncovering the space-math association

Our third, and most theoretically-guided objective, involved working towards an improved understanding of the space-math association. Critically, the inclusion of the targeted constructs provided opportunities to test specific hypotheses about the underlying nature of this relationship. Namely, we sought to determine the potentially mediating roles of children's numerical and EF skills in the relation between spatial visualization and mathematics achievement. Reasons to suspect that numerical skills might mediate the space-math link includes research pointing to the fundamental importance of spatial thinking in the acquisition and development of basic numerical competencies (Dehaene, 2011; Geary, 2004). Indeed, there is evidence that both numerical and spatial cognition rely on highly similar neural networks (e.g., see Hubbard et al., 2005; Toomarian & Hubbard, 2018). This has led some to speculate that reasoning about symbolic number – a relatively recent cultural invention – is rooted in more evolutionarily adaptive neural networks specialized for performing various visual-spatial tasks, such as using and reasoning with objects and tools (Anderson, 2010; Dehaene & Cohen, 2007; Lakoff & Núñez, 2000). Moreover, not only do numerical and spatial processing appear to share biological underpinnings, but they also appear to share close conceptual links (Lakoff & Núñez, 2000; Marghetis, Núñez, & Bergen, 2014). One proposal is that the learning of the number system involves the mapping of numbers to space, a process that has been found to implicate higher-level spatial skills, such spatial visualization (Gunderson et al., 2012; Marghetis et al., 2014; Sella et al., 2017). For example, children's ability to estimate the locations of numbers along a physical number line, has been found to mediate relations between spatial skills and mathematics performance (Gunderson et al., 2012; LeFevre et al., 2013; Tam, Wong, & Chan, 2018). These findings dovetail with theoretical claims that the development of number knowledge corresponds to the refinement of one's 'mental number line' (Dehaene, 2011; Siegler & Booth, 2004; Siegler & Ramani, 2008). In the current study, we look to extend this finding by testing whether or not basic numerical skills more generally mediate the space-math link.

However, this is but one pathway in which spatial ability is potentially linked to mathematics. Moving beyond basic numerical-spatial associations, spatial skills may also be recruited and utilized across a breadth of mathematical tasks, including those that are more distally related to basic numerical processing, including geometric reasoning. Moreover, as discussed earlier, reasoning about numbers in novel contexts, as is required in word problems, algebra, or even arithmetic, may be augmented through the mapping and modeling of these various mathematical relations onto space. We suspect that the same spatial system that allows one to both map and conceptualize numbers along a 'mental number line,' is the same system that allows one to map, model, and conceptualize various other abstract mathematical relations (e.g., see Lakoff & Núñez, 2000; Marghetis et al., 2014). Accordingly, we predicted that spatial ability would relate to mathematics indirectly through its relation with basic numerical skills, but also directly, due to spatial processes that are not specific to number.

Another reason why spatial skills and mathematics may be linked is due to the high executive demands of spatial tasks (e.g., see Miyake et al., 2001). It is possible that spatial tasks are essentially a proxy for EF. Indeed, although a large body of research demonstrates close connections between spatial and mathematical thinking (Mix & Cheng, 2012), it remains to be tested whether EF skills might serve as the common and potentially explanatory source for this relationship. Said differently, it could be the case that spatial thinking is only related to mathematics insomuch as the spatial tasks also recruit and rely on executive functions, such as working memory and inhibitory control. By testing the mediating role of EF in the space-math association, we were able to test the extent to which the space-math link might be explained by individual differences in children's EFs. If the space-math link is fully attributable to children's EF skills than we should expect full mediation. If the space-math link is best explained by children's spatial skills, over and above EF skills, then we should not expect strong evidence of mediation. However, if spatial and EF skills represent distinct constructs with differentiated relations to mathematics achievement – as current theory suggests – we should expect to find evidence of both direct and indirect relations between spatial skills and mathematics achievement. This finding would provide evidence of both shared and unique relations with mathematics. Indeed, we predicted that EF skills would explain some of the shared variance between spatial ability and mathematics performance, but would not fully account for the space-math relation. As outlined above, we hypothesized that differences in the "generative" versus "recall" requirements of the spatial and EF tasks, respectively, should result in separate factors but also differential relations with mathematics performance.

5.6. Different pathways for different mathematical reasoning

Our final objective dealt with issues around the multidimensionality of mathematics. Mathematics is not a unified construct and represents multiple components and skills sets (Mix & Cheng, 2012). Yet, most researchers use arithmetic or calculation-based tasks as mathematics outcome measures. Although arithmetic represents a foundational mathematics skill, more comprehensive mathematics outcome measures are needed to capture the type of mathematics that is more representative of the subject as a whole. Furthermore, multiple measures of the different branches of mathematics are needed to better capture specific relations amongst cognitive skills and different aspects of mathematics. In the current study, numeration and geometry were selected as the two outcome measures of mathematics and used in combination to form the mathematics achievement factor. However, we were also interested in how numerical, spatial, and EF skills might differentially relate to numeration and geometry as separate outcome measures of mathematics. It was predicted that spatial skills will best predict geometry performance, numerical skills will best predict numeration performance, and EF skills will equally predict both. Although these predictions are relatively straightforward, they are a necessary first step in moving towards a more nuanced picture of the cognitive foundations of mathematics performance.

6. Methods

6.1. Participants

Three-hundred and sixteen 4- to 11-year-olds (kindergarten – 4th grade) participated in the study (M_{age} = 6.68 years, *SD* = 1.40: Females = 165). The mean age was the same for males (M_{age} = 6.74 years) and females (M_{age} = 6.62 years), $t(314) = -0.81$, $p = .42$. Table 4 provides a summary of the number of children and mean ages for each grade level. The sample was drawn from eight schools located in both rural (n = 6) and urban communities (n = 2) in northwestern and southwestern regions of Ontario, Canada. Based on 2016 Canadian census data, all participating schools serve communities with family income levels below the Canadian median ($70,336), ranging from $55,936 to $68,062. The schools represent a range of low-to-moderately high performing schools in mathematics based on available standardized provincial test scores. Exactly half of the sample identified as Indigenous peoples of Canada; 94% identified as Anishinaabe and 6% identified as Métis.[1] Based on available 2016 census data, the vast majority (> 95%) of the remaining population identified as Caucasian. Note that although all 316 participants were included in the analyses, data were incomplete or missing for performance on individual measures due to time restrictions (109 cases) or the child's inability to understand task requirements (46 cases). Missing data accounted for 4% of all cases. Written consent was provided by a parent/guardian for all participants and research was carried out in agreement with the ethics boards of the University of Toronto and University of Western Ontario.

6.2. Measures and testing procedures

Participants completed a cognitive test battery involving eleven separate measures (see Table 1). All measures were selected from previously published research. Participants completed the measures in pseudo-random order and in two approximately 30-minute sessions (1–5 days apart). Due to the nature of the tests, the following measures were presented within ordered blocks: Symbolic number comparison, nonsymbolic number comparison, and ordering; path span forward and path span reverse; and KeyMath numeration and KeyMath geometry. All tests were carried out in a quiet location of the child's school (e.g., empty classrooms or private testing rooms) and were administered one-to-one by trained experimenters. The details of each test are provided below.

6.3. Description of numerical indicators

The following three measures were adopted from Lyons, Bugden, Zheng, De Jesus, and Ansari (2018) and Lyons, Hutchison, Bugden, Goffin, and Ansari (2018) and presented to participants in a paper booklet (12 items per page). Participants marked their responses using a pencil. For all three tasks, children were provided with 1 min to complete as many items as possible. The tasks were presented to children in the order in which they are described below. Both the symbolic and nonsymbolic comparison tasks consisted of 72 items and the ordering task consisted of 48 items. For all three measures, the same scoring procedures were used: To adjust for potential speed-accuracy trade-offs or guessing behavior, adjusted raw scores were computed by subtracting the total number of incorrect items from the total number of correct items (see Lyons, Bugden, et al., 2018; Lyons, Hutchison, et al., 2018).

6.3.1. Symbolic number comparison

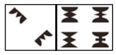
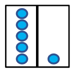
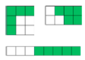
Participants were presented with pairs of Hindu-Arabic numerals (e.g., 4 | 9) and asked to indicate the larger of two numerals as quickly as possible. Numerals ranged from 1 to 9, with absolute numerical distances (N1 - N2) of 1 to 3. All 15 combinations of 1–9 with distances of 1 or 2 were included as well as three combinations with distance 3 (1|4; 3|6; 6|9). This resulted in 18 possible combinations. Trials were counterbalanced to ensure that the larger number appeared on the left and right side of the page an equal number of times.

6.3.2. Nonsymbolic number comparison

Participants were presented with pairs of dot arrays (e.g., : | ::) and asked to select the array with the most dots as quickly as possible. Dot arrays ranged from 1 to 9 dots and included the same numerical distances as those used in the symbolic task. That is, 18 combinations of dot arrays were used and were counterbalanced in the exact same order as the symbolic task. This was done to allow for direct comparison between the symbolic and non-symbolic versions of the task (e.g., see Lyons, Bugden, et al., 2018; Lyons, Hutchison, et al., 2018). Children were instructed not to count the dots. In an effort to control for the influence of the continuous properties of the dot stimuli on performance, both area and contour length were manipulated and controlled for across trials. More specifically, on half the trials dot area was positively correlated with numerosity and overall contour length was negatively correlated

[1] Note that information on Indigenous status was not collected at two of the participating schools due to prior knowledge that these schools predominantly serve Caucasian populations and an extremely low number of Indigenous students; e.g., the 2016 census listed the number of Indigenous families in these communities at zero. For this reason, we did not see a need to inquire about Indigenous status at these schools.

Table 1
Summary of measures used in the study.

Measures	Task description	Example items
Numerical measures		
Symbolic number comparison	● Participants select the numerically larger of two Hindu-Arabic numerals ● 1 min to complete as many items as possible	3 | 4
Nonsymbolic number comparison	● Participants select the numerically larger of two dot arrays ● 1 min to complete as many items as possible	
Ordering	● Participants indicate whether or not a sequence of numerals are in numerical order ● 1 min to complete as many items as possible	1 2 3 ✓ ✗
Spatial measures		
Visual-spatial reasoning	● Participants are presented with 4 different types of 'spatial puzzles' requiring participants to visualize solutions to partially completed puzzles, composition/decomposition tasks, and mental paper folding challenges	
2D mental rotation	● Participants select amongst four options a given shape that can be made by mentally rotating and translating two separated shapes	
Raven's matrices	● Participants are presented with a partially completed image or visual-spatial pattern and must select amongst 6 options the piece that best completes the image/pattern	
Executive function measures		
Head-toes-knees-shoulders	● Participants touch the opposite body part of the one instructed	*"When I say touch your head, I really want you to touch your toes"*
VSWM - forward path span	● Participants are presented with a random sequence of green dots on an iPad screen and watch as individual dots light up one at a time ● Participants recall the exact sequence	
VSWM - reverse path span	● Participants are presented with a random sequence of green dots on an iPad screen and watch as individual dots light up one at a time ● Participants recall the exact sequence but in reverse order in which they occurred	
Mathematics measures		
Numeration	● Comprehensive suite of questions targeting numeration, including questions related to counting, ordering, operations, place value, fractions/proportions/decimals	 *"How many more dots are needed to make ten?"*
Geometry	● Comprehensive suite of questions targeting geometry, including questions related to shape recognition, positional language, transformations, measurement, angles, proofs and formulas	 *"Here are three shapes. Which shape will have the most green squares when it's filled completely?"*

Note that for copyright reasons the example items for Raven's matrices, numeration, and geometry measures were reproduced and do not constitute direct replicas of the actual items. Also note that the for the actual 2D mental rotation task, the bisected shape was presented above the four response items. VSWM = visual-spatial working memory.

with numerosity. On the other half of the trials the opposite was true. Thus, relying on either area or contour length to would result in chance performance (Gebuis & Reynvoet, 2012).

6.3.3. Ordering task

Participants were presented with a sequence of numerals (e.g., 2 – 3 – 4) and asked to indicate whether or not the sequence was in numerical order (i.e., are the numerals in an ascending sequence?). Numerals ranged from 1 to 9, with absolute numerical distances of 1 (e.g., 2 – 3 – 4) or 2 (e.g., 2 – 4 – 6). There were an equal number of correct and incorrect sequences of distances 1 and 2. For half of the items, the sequences were in correct 'ascending order' (e.g., 2 – 3 – 4 or 3 – 5 – 7) and for the other half, the sequences were in incorrect order (e.g., 2 – 4 – 3 or 5 – 3 – 7). Participants put a line through a checkmark to indicate when the sequence was believed to be in order and a line through an 'X' when the order was not believed to be in order.

6.4. Description of spatial indicators

6.4.1. 2D mental rotation

This measure was adapted from Levine, Huttenlocher, Taylor, and Langrock (1999) Children's Mental Transformation Task

(CMTT); a widely used measure of young children's spatial visualization skills, namely mental rotation (Ehrlich, Levine, & Goldin-Meadow, 2006; Gunderson et al., 2012; Hawes, LeFevre, Xu, & Bruce, 2015). Children were presented with two halves of a shape, bisected either along the horizontal or vertical line of symmetry (e.g., a diamond that has been divided into two triangles) and separated and rotated 60° from one another on either the same plane (direct rotation items) or diagonal plane (diagonal rotation items). Four response items (2D shapes) were presented in a 2 × 2 array below the bisected shape. For each item, children were asked to point to the shape that could be made by putting the two pieces together (e.g., a diamond can be made by rotating and translating two triangles). There were 16 items in total; half of which required direct rotations and half of which required diagonal rotations. Note that we modified the original measure by only including items that involved mental rotation (we eliminated items that required translations only). This modification has been shown to make the task more difficult and more appropriate for our targeted grade range (K-3; e.g., see Casey et al., 2018; Hawes et al., 2015). Each item had one correct response. Children were awarded one point for each correct response.

6.4.2. Visual-spatial reasoning

This measure was adapted from Hawes, Moss, Caswell, Naqvi, and MacKinnon (2017) and was designed as a comprehensive measure of children's spatial visualization skills. The test consists of 20 items divided into four different problem types: missing puzzle pieces (two variations), mental paper folding, and composition/decomposition of 2D shapes. For each problem, children were asked to identify the correct answer among four options. One point was awarded for each correct response.

6.4.3. Raven's progressive matrices

This is a widely used measure of children's visual-spatial analogical reasoning (Raven, 2008). Previous research has shown that performance on the task can be linked to a latent spatial visualization factor (Lynn, Backhoff, & Contreras-Niño, 2004; also see Kunda, McGreggor, & Goel, 2010). For each item, participants are presented with a partially completed visual-spatial pattern and must select from amongst six alternatives the puzzle piece that will complete the pattern. The test consists of 36 items. One point was awarded for each correct response.

6.5. Description of executive function indicators

6.5.1. Head-Toes-Knees-Shoulders task (HTKS)

This measure was adapted from Ponitz et al. (2009). The task requires children to engage in flexible attention, working memory, and inhibitory control (McClelland & Cameron, 2012) and closely aligns with Miyake et al. (2000) model of executive functioning. For each item, children listen to an instruction to touch a body part (e.g., "Touch your head") and then must touch a paired "opposite" body part (e.g., toes). Head and toes represented one pair and shoulders and knees represented the other pair. The test was divided into two sections. In the first section, participants were only asked to deal with one pair of body parts (head and toe pairings or shoulder and knee pairings). These pairings were counterbalanced across tests and participants were randomly administered a test version that started with the head and toe pairings or one that started with shoulder and knee pairings. The second section included both pairings. Both sections included 10 items. For each item participants were given a score of 0, 1, or 2; a score of 0 corresponded to incorrect body movements (touching one's head when asked to touch their head), a score of 1 corresponded to a self-corrected body movements (initiating movement towards the wrong body part and then making a correction), and a score of 2 corresponded to correct body movements (touching one's toes when asked to touch their head). Children were given a total score out of 40.

6.5.2. Forward path span

This task was completed on an iPad and used to measure children's working memory, a key component of executive functioning (Miyake et al., 2000). Participants were presented with a set of nine randomly arranged green circles and instructed to watch as the circles lit up one at a time. Each circle was presented for 0.6 s, with 0.5 s of wait time between presentation. Participants then attempted to recall the sequence by touching/tapping the circles in the same order in which they were presented. Following a practice trial, participants began by attempting two trials at a sequence length of two. Upon successful recall of one or two sequences the child progressed to the next level. The task was discontinued when the child failed to recall both sequences at any given level. Children were assigned a score based on the total number of correct sequences recalled.

6.5.3. Reverse path span

This task was identical to the one above but required participants to recall the given sequence in reverse order. For this reason, this task is considered to place even more demands on executive control. For more information and to access both path span tasks see: http://hume.ca/ix/pathspan.html.

Note that we did not include a manifest measure of shifting ability. This decision was based on research indicating that the working memory and inhibitory components of EF are stronger predictors of mathematics than shifting (e.g., see Cragg & Gilmore, 2014). Moreover, shifting skills are presumably implicated in the HTKS task (McClelland & Cameron, 2012).

6.6. Description of mathematics achievement indicators

To assess children's mathematics achievement, we used the Numeration and Geometry subtests of KeyMath (Connolly, 2007). We selected KeyMath as our mathematics outcome measure because it is a standardized Canadian normed test and the items represent a broad range of content knowledge closely aligned with the Ontario mathematics curriculum. The test is administered with an easel booklet and each problem refers to information presented in the form of an image and/or writing. The test is adaptive in that it begins by establishing baseline performance and continues with questions of increasing difficulty. The test is discontinued when the child answer four questions incorrectly in a row. Thus, the test captures a range of children's mathematics skills and not all children are administered the exact same questions. Moreover, because the test is adaptive and continues until a ceiling level of performance is established children are almost inevitably presented with novel mathematical content. The majority of questions are also novel in that they require children to apply their knowledge of mathematical concepts, facts, and procedures within contexts likely unfamiliar to the students (e.g., rather than solving a standard arithmetic problem children must apply their knowledge of arithmetic to solve a problem dealing with combinations of block structures). The majority of items in both subtests require knowledge of the symbolic number system. Children were awarded a total raw score by subtracting the total number of incorrect responses from the maximum item number reached.

6.6.1. Numeration test

This measure includes a total of 49 questions related to counting, comparing quantities, recognizing and ordering number symbols, operations, place value, and proportions/fractions/decimals.

6.6.2. Geometry test

This measure includes a total of 36 questions related to shape recognition, positional language, geometrical transformations (e.g., rotations), measurement, grid coordinates, angles, geometric proofs.

6.7. Analytic approach

Analyses were carried out using the recommended two-step approach to structural equation modeling (SEM; see Kline, 2015). The first step involved testing the measurement model using confirmatory factor analyses (CFA). The purpose of the measurement model is to test and observe the relations between the observed variables (aka indicator or manifest variables) and the relations these variables have with the hypothesized construct or constructs (aka factors or latent variables). Failure to obtain adequate fit statistics at this stage may indicate the need to reconsider the model and/or make modifications to the model. The second step involved analyses of the full structural equation model(s). The purpose of this step is to test hypothesized interrelations between constructs/factors and is similar in some ways to general linear regression models. However, a major advantage of SEM over general linear models is that SEM takes error variances into account (regression analyses assume variables are measured without error: see Weston & Gore, 2006) and allows one to model both variability common to a latent variable (i.e., error-free scores) as well as the variability not explained by the latent variable (i.e., error). Moreover, SEM allows for the creation of weighted aggregate variables of targeted constructs. That is, latent variables are not merely an average of scores obtained across different measures but a composite score that has been weighted according to the various contributions that each indicator variable makes to the construct of interest.

Analyses were performed with Mplus Version 7.4 (Muthén & Muthén, 1998, 2015) using the default maximum likelihood estimation (MLE) procedures. All analyses were conducted on raw (continuous) scores. Modification indices were requested for chi-squared values equal to or > 10. Missing data (4% of all cases) were treated with full information likelihood (FIML) estimation procedures (the default option in Mplus). Confidence intervals were computed using Mplus' bias corrected bootstrapped confidence interval procedure. Note that the following link provides an annotated copy of the Mplus scripts used for each of our analyses along with the corresponding output/results (https://osf.io/2y7xu/).

We used three goodness-of-fit statistics to compare our CFA models and determine model fit: (1) Root Mean Square Error of Approximation (RMSEA), (2) Comparative Fit Index (CFI), and (3) Standardized Root Mean Residual (SRMR). Decisions about what constitutes acceptable or 'good' model fit were based on the following recommendations: RMSEA values of < 0.10, and CFI values > 0.95, and SRMR values < 0.08 (Kline, 2015). Note that we also report chi-squared (χ^2) values for comparison purposes but due to the large sample size (> 200) did not interpret statistically significant results in any meaningful way (see Kline, 2015).

Power analyses were conducted to determine the minimum sample size needed to detect a medium effect size with an alpha of = 0.05 and power = 0.95 (Soper, 2018). Using a SEM with four latent variables and 11 indicator variables, the results indicated a recommended sample size of 241 participants.

7. Results

7.1. Part I: Measurement model

Table 2 shows the descriptive statistics for each measure. As can be seen, the kurtosis and skewness values of each indicator variable fall within the acceptable limits of ± 2 (Field, 2009). Table 3 shows the bivariate zero-order correlations between all variables. As can be seen, there were moderate to high correlations between all measures included in the measurement model (0.42–0.83): Note that age was included as covariate in the structural models, but not the measurement model, as we had little reason to suspect measurement invariance across age (see Section 7.2.2 for analyses related to the moderating effects of age). Scatter plots were used to visualize the data distributions between all variables and no concerns were noted (e.g., nonmorality, lack of homoscedastic, outliers). Furthermore, data were screened to ensure normal distributions of performance for each task across each grade level. These analyses indicated relative normal distributions across tasks and grades. Table 4 provides a summary of the number of

Table 2
Descriptive statistics for all measures and reliability estimates.

	N	Mean score (SD)	Range (min - max)	Kurtosis	Skewness	Reliability
Demographic measures						
Age (years)	316	6.68 (1.40)	4.08–10.92	−0.398	0.331	N/A
Gender	316	52% female	0–1	N/A	N/A	N/A
Numerical measures						
Symbolic number comparison	304	21.91 (13.31)	−6 to 64	−0.398	0.143	0.84
Nonsymbolic number comparison	305	15.52 (7.61)	−5 to 35	−0.272	−0.203	0.71
Ordering	285	7.24 (5.44)	−10 to 21	−0.096	−0.118	0.72
Spatial measures						
Visual-spatial reasoning	313	9.75 (3.66)	2–19	−0.642	0.294	0.70
2D mental rotation	309	8.83 (3.35)	1–16	−0.762	−0.175	0.66
Raven's matrices	309	20.31 (6.05)	4–33	−0.523	−0.084	0.68
EF measures						
Head-toes-knees-shoulders	292	27.52 (10.99)	0–40	0.423	−1.21	0.73
VSWM - forward path span	286	4.06 (2.25)	0–12	−0.005	0.244	0.66
VSWM - reverse path span	285	3.03 (2.45)	0–9	−0.692	0.581	0.60
Mathematics measures						
Numeration	314	9.63 (5.08)	0–22	−0.736	0.436	0.91
Geometry	311	10.75 (4.78)	0–24	−0.089	0.205	0.60

Note that reliability estimates represent test-retest coefficients and were based on a subsample of participants ($N = 106$) who were administered the measures at two time points (~4 months apart).

Table 3
Zero-order correlations between variables.

Measures	1	2	3	4	5	6	7	8	9	10	11	12	13
Demographic measures													
1. Age	–												
2. Gender	0.05	–											
Numerical measures													
3. Symbolic number comparison	0.76	0.02	–										
4. Nonsymbolic number comparison	0.72	−0.03	0.83	–									
5. Ordering	0.51	0.03	0.66	0.64	–								
Spatial measures													
6. Visual-spatial reasoning	0.59	0.02	0.58	0.56	0.47	–							
7. 2D mental rotation	0.64	−0.03	0.63	0.61	0.48	0.64	–						
8. Raven's matrices	0.67	0.00	0.66	0.65	0.51	0.65	0.67	–					
EF measures													
9. Head-toes-knees-shoulders	0.49	−0.07	0.61	0.61	0.47	0.48	0.51	0.52	–				
10. VSWM – forward path span	0.56	−0.03	0.57	0.56	0.46	0.42	0.48	0.52	0.50	–			
11. VSWM – reverse path span	0.56	0.07	0.61	0.56	0.52	0.54	0.54	0.59	0.51	0.57	–		
Mathematics measures													
12. Numeration	0.70	0.06	0.76	0.72	0.69	0.69	0.67	0.72	0.58	0.57	0.63	–	
13. Geometry	0.56	−0.01	0.62	0.55	0.54	0.68	0.61	0.65	0.54	0.46	0.55	0.75	–

Note: With the exception of the correlations between gender and the other variables ($ps > .05$), all other correlations were statistically significant at $p < .001$. Gender was dummy coded: 0 = females and 1 = males.

children and mean scores for each grade level.

In total, five different CFA models were run on the data and associated covariance matrix (see Table 5 for summary of each model run). Of primary interest was to test the hypothesized four-factor measurement model. The results of this model indicated good fit statistics, with the RMSEA value (0.057) below the recommended cut-off of 0.10 and the CFI value (0.983) above the recommended threshold of 0.95 (Kline, 2015). The high CFI value suggests that the model is superior to a "null" model or one that assumes zero correlations between the variables. Importantly, the recommended modification indices were relatively low (MIs < 17) and inconsequential to the overall model fit. These recommendations ran contrary to the theoretical model (later to be tested with SEM), as they implicated cross-loadings between the mathematics outcome variables and individual indicator variables. Overall, the results provide support for a four-factor model.

Although the four-factor model demonstrated good fit statistics, we tested four alternative models. These modifications served the purpose of hypothesis testing as well as attempts to improve the overall model fit. The first of these modifications included the removal of the path span reverse indicator (a measure of VSWM) from the Executive Function Factor and including it as an indicator

Table 4
Descriptive statistics showing the number of children and mean scores for each grade level.

	Grade	VisSpat (max 20)	MR2D (max 16)	Raven's (max 36)	SymNum (max 72)	NonSym (max 72)	Ordering (max 48)	HTKS (max 40)	PathFor (max 18)	PathRev (max 18)	Numeration (max 49)	Geometry (max 36)
N	JK	32	32	32	32	32	32	32	31	31	32	32
	SK	71	69	70	67	66	54	59	65	65	72	72
	1	77	76	76	75	76	73	73	72	72	76	76
	2	74	73	74	73	74	69	72	67	66	74	73
	3	46	46	45	45	45	45	44	39	39	47	45
	4	13	13	12	12	12	12	12	12	12	13	13
Mean	JK	6.06	4.75	13	4.84	5.16	2.81	17.9	1.71	0.71	4.41	7.09
	SK	7.48	6.8	16.1	11.6	10	4.39	21.6	2.71	1.57	5.75	7.38
	1	9.55	8.82	20.6	22	16.2	6.88	28.5	4.35	3.32	9.63	10.9
	2	11.4	10.2	22.6	26.1	18.1	8.3	30	4.81	3.55	10.9	12.3
	3	12.7	11.4	26.1	37	23	11.3	35.1	6.08	5.05	15.5	15.1
	4	12.7	12.7	26.8	41.8	25.5	12.8	34.1	5.08	5.67	15.4	13.9
SD	JK	1.58	1.68	3.66	5.07	4.11	3.2	12.3	1.27	0.864	1.7	3.32
	SK	2.52	2.42	4.37	8.88	5.56	4.2	12.5	1.8	1.51	2.79	3.72
	1	3.16	2.84	4.4	7.97	5.06	4.41	9.34	1.91	2.26	3.66	3.64
	2	3.36	2.84	5.02	9.02	5.48	5.52	9.02	1.87	2.16	4.66	4.48
	3	3.12	2.35	4.06	9.71	4.99	5.27	4.01	2.08	2.55	3.49	4.22
	4	3.12	2.59	4	4.59	4.1	4.13	2.43	1.44	2.23	3.73	3.33

Note: JK = Junior Kindergarten (equivalent of pre-school in the U.S.); SK = Senior Kindergarten (equivalent of kindergarten in the U.S. and other countries). SD = standard deviation. Max = maximum possible score on measure. VisSpat = Visual-Spatial Reasoning; MR2D = 2D Mental Rotation; SymNum = Symbolic Number Comparison; NonSym = Non-Symbolic Number Comparison; Ordering = Ordering Task; HTKS = Head-Toes-Knees-Shoulders Task; PathFor = Path span forward; PathRev = Path span reverse.

Table 5
CFA (measurement model) goodness-of-fit statistics for original hypothesized model and four modified alternative models.

Model	χ^2 (df)	RMSEA (90% CIs)	CFI	SRMR	AIC
1. Hypothesized four-factor model	77.31(38) $p < .001$	0.057 (0.039–0.075)	0.983	0.024	17959.95
2. Loading VSWM (path spanreverse) with Spatial Factor	89.19(38) $p < .001$	0.065 (0.048–0.083)	0.978	0.029	17971.83
3. Loading both VSWM measures with Spatial Factor and HTKS as single indicator	98.56(39) $p < .001$	0.070 (0.053 –0.087)	0.975	0.031	17979.19
4. Three-factor model where spatial and EF measures load on same factor	109.87(41) $p < .001$	0.073 (0.057–0.090)	0.971	0.033	17986.51
5. Single, *g*-factor, model	182.34(43) $p < .001$	0.101 (0.086–0.117)	0.941	0.038	18054.97

Note: VSWM = visual spatial working memory; HTKS = head-toes-knees-shoulders.

of the Spatial Factor. This modification was justified based on the grounds that the task demands share some features with the other spatial ability measures (i.e., require storage and manipulation of visual-spatial information) and has been used in prior investigations as a measure of spatial ability. The modification did not improve the model. As can be seen in Table 5, the change resulted in a slight increase in the chi-square value and a small increase in the RMSEA values. As a follow-up to this analysis, and based on a similar rationale, another modification was made in which both VSWM measures were made to load on the Spatial Factor. The HTKS variable was converted into a single indicator latent variable. This was achieved by multiplying the mean/variance associated with the HTKS measure (120.307) with a reasonable estimate of assumed error variance (0.20; see Kline (2015) for more details on this approach). This modification did not improve the model fit. Thus, there appears to be little difference between a model that includes a well-defined Spatial Factor with three indicator variables and a more comprehensive, yet less defined, Spatial Factor with five indicator variables. To further test the relative separability of the Spatial Factor from the EF factor, a three-factor model was carried out in which the spatial and EF measures were made to load on the same factor. Although this model demonstrated good fit (see Table 5), it failed to achieve the same quality of fit statistics of the four-factor model. Results of a nested chi-square difference test revealed statistically significant differences between the four-factor (Model 1) and the three-factor model (Model 4); χ^2 (3) = 32.56, $p < .001$. These results suggest that spatial and EF skills – as measured in the present study – represent distinct constructs.

Finally, a measurement model was evaluated in which all predictor variables were made to load on a single general factor (i.e., *g*). The rationale for such a modification was based on recent research suggesting that general intelligence, or a *g*-factor, might be responsible for previously observed relations between cognitive variables and academic achievement (e.g., see Ritchie, Bates, & Deary, 2015). Moreover, this modification was justified based on the relatively high correlations between all predictor variables. The results of this model indicated marginally acceptable fit (RMSEA = 0.101, CFI = 0.94), but worse fit statistics compared to the hypothesized four-factor model. Results of a nested chi-square difference test revealed statistically significant differences between the four-factor (Model 1) and the single-factor model (Model 5); χ^2 (5) = 105.03, $p < .001$. This suggest that the four-factor model fits the data significantly better than the one-factor model.

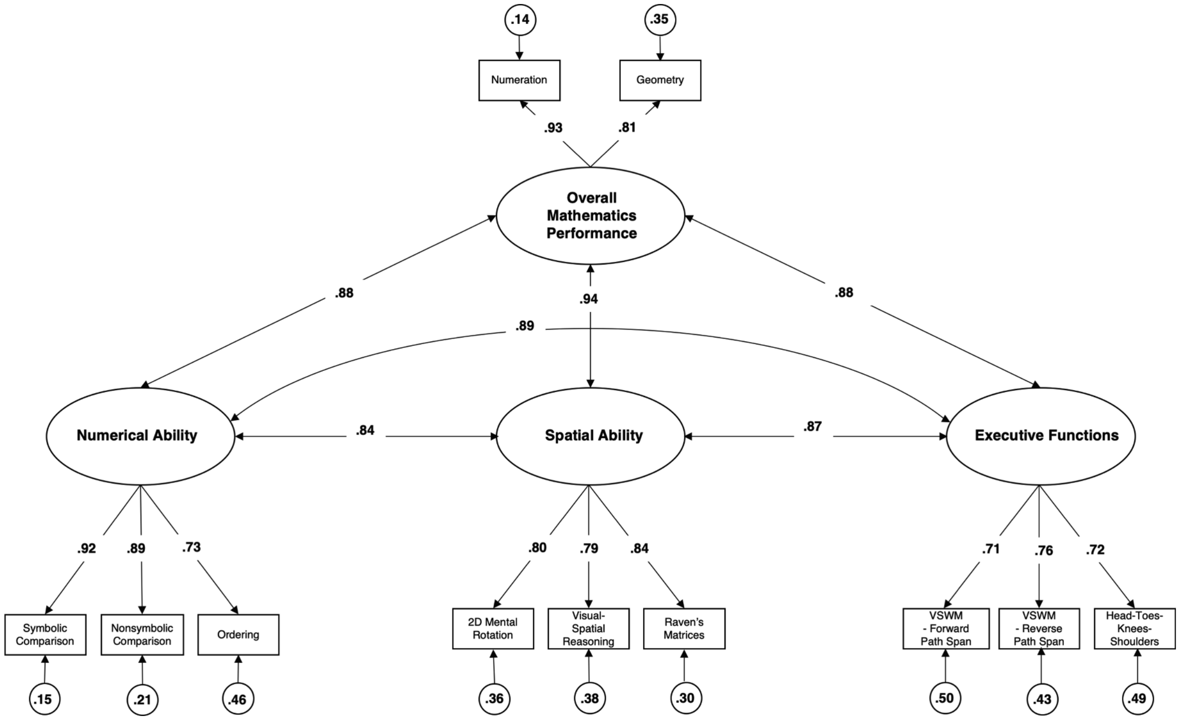


Fig. 1. The four-factor measurement model and the one retained for further structural analyses. Double-headed arrows represent correlations between factors and single-head (unidirectional) arrows indicate factor loadings (interpreted as regression coefficients). The smaller circles with arrows leading to each indicator variable represents unexplained variance or residual/error terms (interpreted as proportions of unexplained variance). Note that correlations can be squared to determine the proportion of shared variance between variables (e.g., the proportion of shared variance between spatial ability and mathematics is $0.94^2 = 0.88$).

7.1.1. *Summary of results*

Overall, the five measurement models evaluated demonstrated acceptable fit statistics and any one of them could technically be retained and considered decent representations of the data. However, based on a priori theoretical decisions and the finding that models 1–4 were all comparable in fit, the four-factor model was retained and used in all subsequent analyses.

Fig. 1 shows the final measurement model and the relations between factors as well as the relations between indicators and their residuals in relation to each factor. As can be seen, the correlations between factors are extremely strong (> 0.84), with a range of correlation values between 0.84 (spatial with numerical) and 0.94 (spatial and mathematics achievement). Despite high correlations between factors, multicollinearity analyses at the latent variable level revealed acceptable tolerance and VIF statistics (Spatial factor; tolerance = 0.402, VIF = 2.488: Numerical factor; tolerance = 0.370, VIF = 2.699: EF factor; tolerance = 0.423, VIF = 2.364). Note that mathematics was entered as the dependent variable in the model that was used to derive these statistics. Concerns of multicollinearity occur when the VIF statistic exceeds 10 and the tolerance statistic is below 0.10 (e.g., see O'brien, 2007). Accordingly, multicollinearity of factors does not appear to present a problem in the present study and further structural analyses were planned to potentially reveal the unique relations between factors.

The indicator variables also appear to adequately reflect the factors of interest as can be seen by the relatively low residuals. All indicators explain at least 50% of the variance with their associated factor. This is a desirable outcome and provides further evidence that the factors are adequately represented by their hypothesized indicator variables (Kline, 2015).

In sum, the four-factor model provides a good fit of the data, despite extremely high correlations between each factor. Thus, there is evidence to suggest that numerical skills, spatial ability, EF, and mathematics achievement are highly correlated but separable constructs.

7.2. *Part II: Structural models*

7.2.1. *Cognitive predictors of mathematics achievement*

Our first set of structural analyses tested the unique and shared relations between the numerical, spatial, and EF factors and their predictive relations with mathematics achievement. Fig. 2 shows the various relations with one another after controlling for age.[2]

[2] Hereafter, all analyses were carried out with age as a covariate at the individual variable level (i.e., scores on each manifest variable was regressed on each child's age in months). Fig. 2 represents the retained four-factor model; all subsequent analyses involve a deconstruction of this

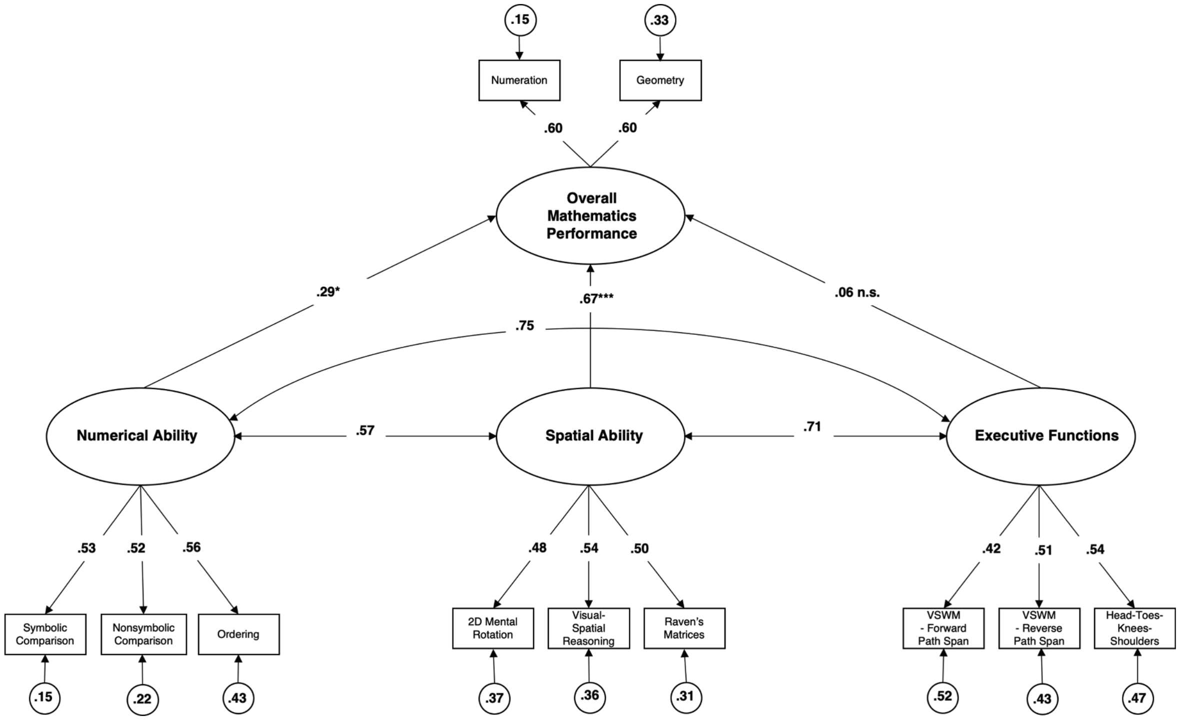


Fig. 2. Relations between cognitive predictors and overall mathematics achievement controlling for age. All pathways are significant (*p*s $<$ .05) except for the path between EF and mathematics. All values represent standardized estimates.

Overall, the model explained a large proportion, 0.84, of the variance in mathematics achievement controlling for age. Both the numerical and spatial factors were unique predictors of mathematics achievement, $\beta = 0.289$, $SE = 0.12$, $p = .013$, 95% *CI* [0.06, 0.52], 99% *CI* [−0.01, 0.59] and $\beta = 0.673$, $SE = 0.11$, $p < .001$, 95% *CI* [0.46, 0.88], 99% *CI* [0.40, 0.95], respectively. Accordingly, a 1-unit increase on the numerical factor was associated with a 0.29 standard deviation unit increase on the mathematics factor, controlling for the effects of age, spatial ability, and EF skills. A 1-unit increase on the spatial factor was associated with a 0.67 standard deviation unit increase on the mathematics factor, controlling for the effects of age, numerical, and EF skills. The relation between the spatial and mathematics factors remains robust even at 99% CIs, whereas the relation between the numerical and mathematics factors is no longer statistically significant at 99% CIs. There were no unique relations between the EF and mathematics factor once the numerical and spatial factors were taken into account, $\beta = 0.056$, $SE = 0.17$, $p = .734$, 95% *CI* [−0.27, 0.38], 99% *CI* [−0.37, 0.48].

A follow-up test was carried out to examine the possibility that the above results may have been driven by the inclusion of Raven's Matrices as an indicator of spatial ability. Given that matrix reasoning is typically considered a measure of nonverbal intelligence and not necessarily a measure of spatial ability proper, it was important to determine whether the above results remained once this measure and its contributions were removed. Moreover, this allowed us to further isolate and more narrowly examine the effects of spatial visualization on mathematics achievement. The results were highly consistent with those reported above. Both the numerical and spatial factors remained independent predictors of mathematics, $\beta = 0.302$, $SE = 0.12$, $p = .014$, 95% *CI* [0.06, 0.54], 99% *CI* [−0.01, 0.62] and $\beta = 0.665$, $SE = 0.12$, $p < .001$, 95% *CI* [0.44, 0.89], 99% *CI* [0.37, 0.96], respectively. There was no unique relation between EF and mathematics, $\beta = 0.08$, $SE = 0.17$, $p = .630$, 95% *CI* [−0.25, 0.42], 99% *CI* [−0.36, 0.53]. Notably, highly similar results were obtained when Raven's Matrices was used as a control measure. Both numerical and spatial performance were strongly related to mathematics performance, $\beta = 0.38$, $SE = 0.13$, $p = .004$, 95% *CI* [0.12, 0.64], 99% *CI* [0.04, 0.72], and $\beta = 0.60$, $SE = 0.12$, $p < .001$, 95% *CI* [0.37, 0.83], 99% *CI* [0.30, 0.90]. Again, there was no unique relation between EF and mathematics, $\beta = 0.10$, $SE = 0.17$, $p = .556$, 95% *CI* [−0.23, 0.42], 99% *CI* [−0.33, 0.52]. These results demonstrate that the relation between spatial and mathematics performance remains strong when Raven's Matrices is excluded from the analyses altogether, but also when it is included as general covariate in the model. Therefore, the relation between spatial ability and mathematics appears to be related to spatial visualization skills.

Taken together, the results indicate that in combination, numerical, spatial, and EF skills explain a large proportion of the variance in mathematics performance. More specifically, the results reveal significant unique relations between numerical and spatial skills with mathematics achievement and no unique relations between EF and mathematics. Spatial ability appears to be an especially

(*footnote continued*)
model in an attempt to better understand the potentially explanatory pathways that give rise to these observed relations.

strong contributor to mathematics achievement, over and above contributions from EF and numerical skills.[3]

7.2.2. Stability of performance across age

To examine the extent to which scores on the various factors were stable across age, composite 'factor' scores were computed for each individual. Importantly, these scores are weighted according to each indicator's contributions (i.e., performance on each test) to the latent construct of interest. In this way, composite scores are not merely an average score on a given number of measures. To test whether factor scores vary as a function of development, a grade (6) by factor (4) repeated measures ANOVA was carried out. Results revealed that mean factor scores differed significantly between grades, $F(13.849, 667.544) = 1.830$, $p = .032$, ${\eta_p}^2 = 0.037$. Follow-up Bonferroni corrected comparisons revealed statistically significant differences between factor scores only amongst the youngest grade tested (i.e., junior kindergarten); scores differed significantly from one another on the spatial versus mathematics factor ($p = .007$) and numerical versus mathematics factor ($p = .037$). There were no other significant differences in factor scores in kindergarten through 4th grade. A Bayesian repeated measures ANOVA was carried out to further examine the strength of evidence for the presence of grade by factor interactions. These results revealed a Bayes Factor of 0.003, indicating extremely weak support for the hypothesis of grade by factor interactions. Overall, despite a statistically significant grade by factor interaction, a closer look at the data reveals highly similar developmental trajectories of each construct across age. Note that Mix et al. (2016) also reported consistent relations between space and mathematics across grades K, 3, and 6. Fig. 3 provides an illustration of the relation between age and children's individual factor scores. An analysis of potential differences in the slopes of each factor revealed statistically insignificant results, $F(3, 1161) = 0.980$, $p = .402$. There was also no statistically significant differences in the intercepts, $F(3, 1164) = 0.294$, $p = .830$. Overall, the results suggest fairly consistent relations between factors over developmental time.

7.2.3. Mediation analyses: numerical and EF skills as mediators of the space-math link

Next, mediation analyses were carried out based on theory to suggest that numerical and EF skills potentially mediate the shared relations between spatial and mathematical processing. That is, the analyses reported above were further decomposed to test whether EF and numerical skills independently mediate the relation between spatial thinking and mathematics achievement. For example, to test the mediating role of numerical skills in the space-math relationship, we removed the EF factor from the model presented in Fig. 2; conversely, to test the mediating role of EF skills we removed the numerical factor from the model. Thus, mediation models were achieved by removing irrelevant pathways from the four-factor model and retaining only the pathways of specified interest. Note that the rationale and interpretation of each mediation analysis differed somewhat according to construct and question of interest (numerical vs EF). Numerical skills were targeted as a potential mediator for reasons that are best described by the *causal steps approach* to mediation (Barron & Kenny, 1986; Judd & Kenny, 1981), while EF skills were entered into the model for reasons that align with the *confounding variables approach* (e.g., see MacKinnon, Fairchild, & Fritz, 2007). While the causal steps approach involves testing a causal chain of events and typically assumes temporal precedence (e.g., spatial skills → numerical skills → mathematics achievement), the confounding variables approach – although a mathematically equivalent model – is used to test the influence of a potentially confounding or third variable in a given bi-variate relationship (e.g., testing the extent to which EF skills might explain the space-math link; MacKinnon et al., 2007; also see Fiedler, Harris, & Schott, 2018).

As shown in Fig. 4, numerical skills were found to partially mediate the relation between spatial skills and mathematics achievement, $\beta = 0.182$, $SE = 0.04$, $p < .001$, 95% CI [0.10, 0.27], 99% CI [−0.07, 0.29]. The direct effect between spatial skills and mathematics remained robust, $\beta = 0.697$, $SE = 0.08$, $p < .001$, 95% CI [0.55, 0.85], 99% CI [0.50, 0.89]. Fig. 5 shows the results of EF as a mediator between spatial ability and mathematics. As can be seen, EF failed to mediate these relations, $\beta = 0.159$, $SE = 0.09$, $p = .064$, 95% CI [−0.01, 0.33], 99% CI [−0.06, 0.38]. The direct effect between spatial skills and mathematics remained robust, $\beta = 0.740$, $SE = 0.11$, $p < .001$, 95% CI [0.52, 0.96], 99% CI [0.45, 1.03].[4] In sum, numerical skills, but not EF, were found to partially mediate the relation between spatial ability and mathematics.

7.2.4. Predictive relations with different components of mathematics

The final set of analyses examined the relations between the cognitive predictors and each mathematics outcome measure.

Numeration and geometry were entered as single indicator outcome variables and two separate structural models were run. These analyses were carried out to determine how the various relations previously observed potentially vary as a function of the mathematics activity in question. Fig. 6 shows a summary of the results when numeration was used as a single outcome variable. Similar to the results of the full model, scores on the numerical and spatial factors were both uniquely related to numeration performance, $\beta = 0.252$, $SE = 0.07$, $p = .001$, 95% CI [0.11, 40], 99% CI [0.06, 44] and, $\beta = 0.342$, $SE = 0.07$, $p < .001$, 95% CI [0.21, 0.48],

[3] Due to concerns over the age-appropriateness of some measures, we re-ran the main analysis (3-factor prediction model) without the 3rd and 4th grade children. This resulted in the removal of 60 children, leaving a total of 256 K-2 children in the database. The pattern of results from this analysis were identical to those obtained with the larger sample. These results suggest that students in grades 3 and 4 are not skewing the results potentially due to ceiling effects on any of the measures.

[4] As a follow-up, we also conducted an analysis in which we partialled out the effects of EF skills from *both* spatial and mathematics skills. (The mediation model is more akin to a partial correlation in which the influence of the mediator is partialled out from the outcome variable only.) These results further confirmed strong relations between spatial and mathematics performance even after the influence of EF on both variables was taken into account. More specifically, spatial skills explained 81% of the variance in mathematics before taking EF into account and 62% of variance after EF was taken into account. Therefore, EF skills explained approximately 23% (i.e., 19/81) of the variance in the space-math link. In short, regardless of analytical approach, the space-math link does not appear to be explained by children's EF skills.

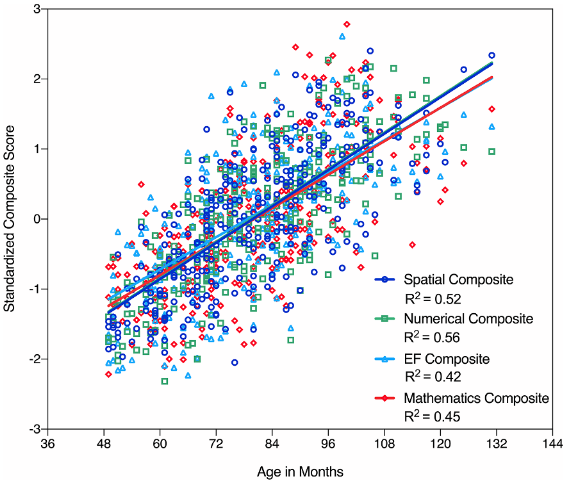


Fig. 3. Scatterplot of individual participants' age and standardized composite score for each factor. Each column represents one year.

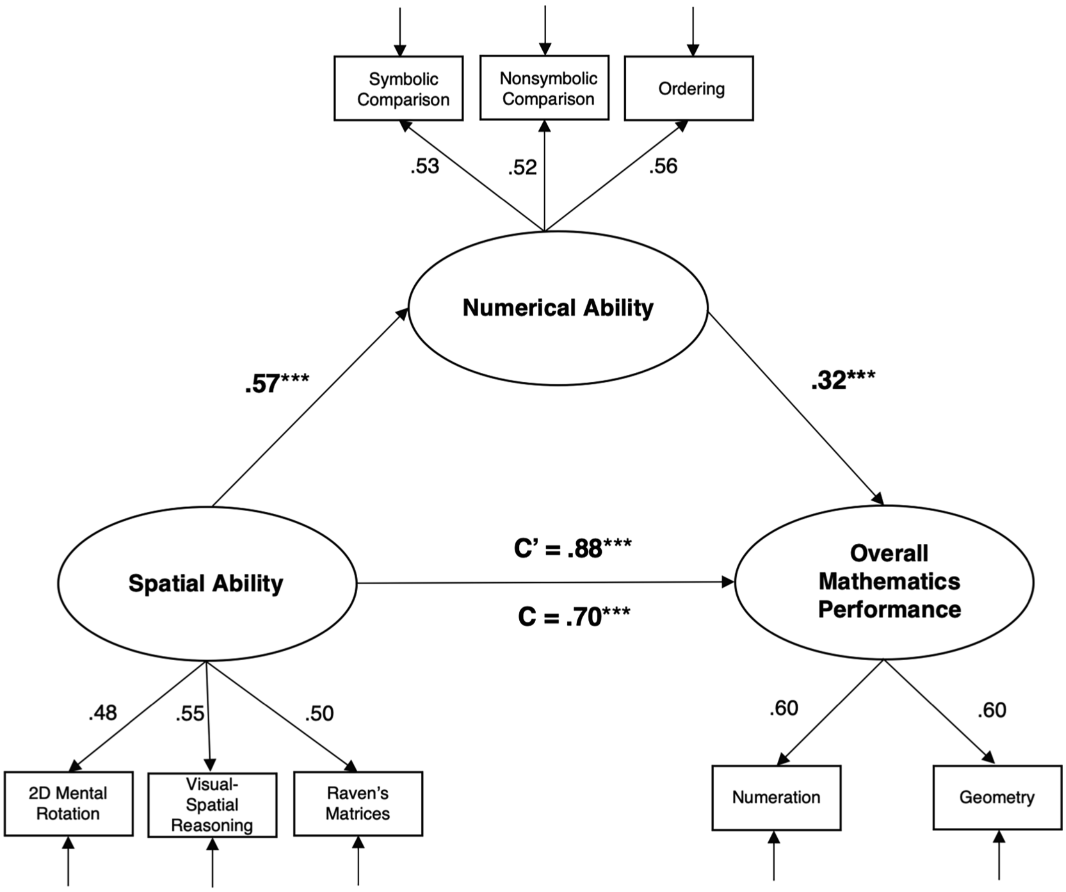


Fig. 4. Mediation model of numerical ability in the relation between spatial ability and overall mathematics performance. $^{***}p < .001$. Values represent standardized coefficients.

99% *CI* [0.16, 52] respectively. Scores on the EF factor were not statistically predictive of performance on the numeration test, $\beta = 0.024$, $SE = 0.11$, $p = .821$, 95% *CI* [−0.18, 0.23], 99% *CI* [−0.25, 30].

Fig. 7 shows the results when geometry was used as a single outcome variable. Scores on the spatial factor were strongly related to performance in geometry, $\beta = 0.532$, $SE = 0.09$, $p < .001$, 95% *CI* [0.36, 0.71], 99% *CI* [0.30, 0.77]. Scores on the numeration and EF factor did not predict performance on the geometry test, $\beta = -0.003$, $SE = 0.10$, $p = .979$, 95% *CI* [−0.19, 0.19], 99% *CI*

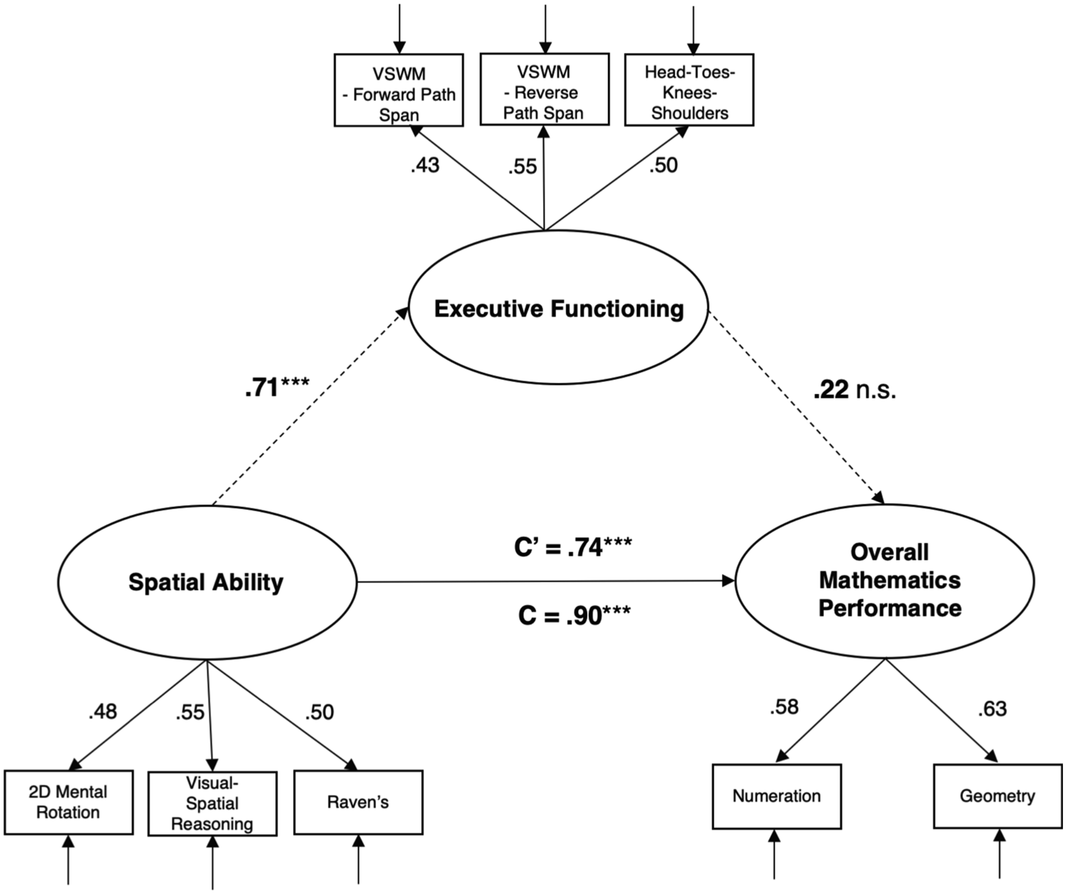


Fig. 5. Mediation model of EF in the relation between spatial ability and overall mathematics performance. $^{***}p < .001$. n.s. = non-significant. Values represent standardized coefficients.

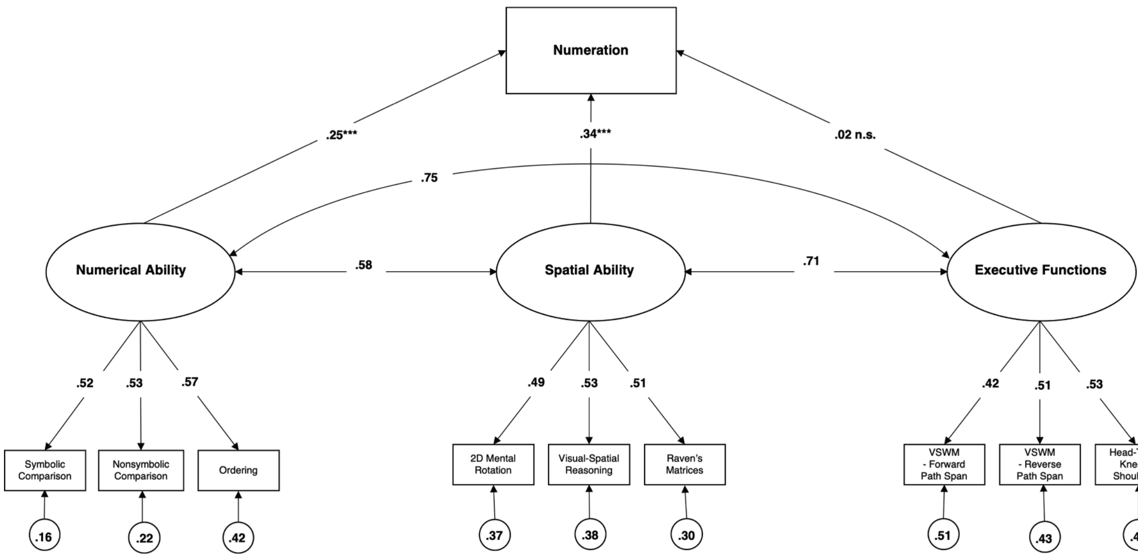


Fig. 6. Cognitive predictors of numeration as a single test outcome measure. Values represent standardized coefficients. $^{***}p = \leq 0.001$.

[−0.25, 0.25] and, $\beta = 0.064$, $SE = 0.14$, $p = .648$, 95% CI [−0.21, 0.34], 99% CI [−0.29, 0.43].

Taken together, both spatial and numerical skills predicted performance on the numeration test, but only spatial skills predicted performance on the geometry test. Executive functioning skills did not explain any unique variance on either measure.

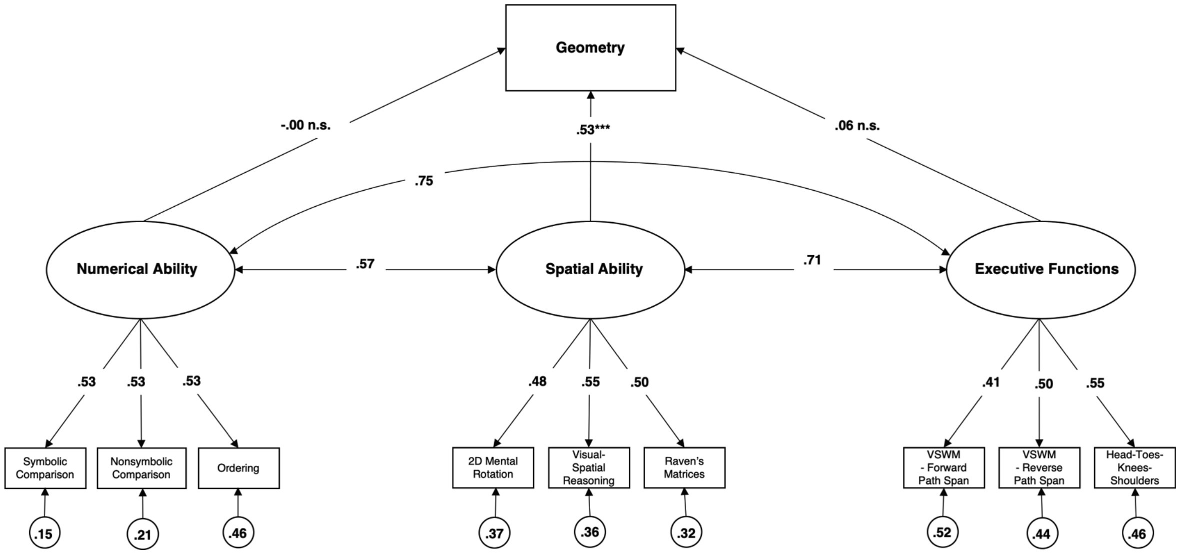


Fig. 7. Cognitive predictors of geometry as a single test outcome measure. Values represent standardized coefficients. $^{***}p < .001$.

8. Discussion

The current study examined the cognitive foundations of early mathematics achievement in a sample of 4- to 11-year-olds. Analyses were first carried out to test the psychometric properties associated with a hypothesized four-factor model, with cognitive constructs related to numerical, spatial, and executive function skills and mathematics achievement. The four-factor model revealed robust correlations between each factor while also demonstrating good fit statistics; a finding that suggests that numerical, spatial, EF, and mathematics abilities are highly related but separable constructs. Importantly, the original four-factor model achieved better fit statistics than several alternative models, including a model in which a general (*g*) factor was used to link each individual predictor variable with mathematics achievement.

Given evidence of a four-factor model, our primary analyses aimed to more closely reveal the structure and underlying relations between numerical, spatial, EF, and mathematics skills. To this regard, we had several goals: (i) to examine the shared and unique contributions of children's numerical, spatial, and EF skills to mathematics achievement, (ii) to determine the relative stability of these relations across childhood, (iii) to test the potentially mediating roles of numerical and EF skills in the oft reported space-math link, and, lastly, (iv) to examine the extent to which relations between the predictor variables and mathematics vary as a function of the mathematics task in question (i.e., numeration vs. geometry).

Results revealed that children's numerical, spatial, and EF skills collectively explained 84% of the variance in mathematics achievement, controlling for the effects of age. These results provide evidence of a fairly comprehensive model of children's mathematics achievement. However, only the numerical and spatial factors were uniquely predictive of mathematics achievement. The observed relations between factors remained stable across age and grade, appearing to undergo highly parallel growth trajectories. Follow-up mediation analyses revealed that numerical skills, but not EF skills, partially mediated the relation between spatial skills and mathematics achievement. Our last set of analyses examined how the predictive utility of the model potentially varies as a function of the mathematics task being assessed, that is, numeration vs. geometry. Scores on the numerical and spatial factors were uniquely related to numeration performance, while only spatial ability was a unique predictor of geometry performance.

In the following sections, we provide a more detailed review of the main findings just described. We begin by discussing the results of the CFA analysis and then review and offer interpretations of the findings related to the structural models employed. We focus much of our attention on the space-math link and more carefully consider the role of spatial visualization in children's mathematics performance.

8.1. Evidence of a four-factor model

We found evidence to suggest that numerical, spatial, EF, and more general mathematics skills are highly related but separable constructs. The correlations amongst these factors were strikingly high and similar in strength (*r*s 0.84–0.94) and indicate higher relations at the latent variable level than what would be predicted by examining the relations amongst the single indicator variables alone. This finding in itself demonstrates the potential utility of forming and testing the relations between latent variables, as they offer a more comprehensive model of the targeted constructs; one not defined by a single measure – but rather a combination of measures – and less influenced by measurement error.

Subsequent analyses revealed that the four-factor model achieved better fit statistics than a single-factor (*g*) model. While the four-factor model demonstrated good fit, the single-factor model straddled the boundary of what is considered acceptable fit statistics. Overall, our results suggest the need to be cautious in interpreting each factor as fully independent constructs. Instead, numerical, spatial, EF, and general mathematics achievement appear to strongly overlap with one another and yet are distinct enough to represent separable constructs. This finding supports Mix et al. (2016), who found evidence of highly related but separable factors associated with spatial and mathematical domains in a large sample (N = 854) of 5-to 13-year-olds. Moreover, these authors found evidence of strong cross-domain loadings for certain spatial and mathematical tasks, suggesting that a common cognitive network might underlie certain spatial and mathematical tasks. Notably, Mix et al. (2016) also presented evidence showing that spatial and mathematical tasks loaded on to a single factor in an orthogonal EFA model. Thus, our findings, like those reported by Mix et al. (2016), suggest a tight coupling of spatial and mathematical thinking. The current findings suggest that numerical skills and EFs might also be implicated in this same cognitive network.

Several follow-up analyses were carried out to further confirm evidence of a four-factor model as well as to test specific theoretical distinctions in measurement. Of primary interest was whether measures of VSWM would more strongly load on the EF or spatial factor. Our results indicated better model fit when the VSWM measures were made to load on the EF factor. Moreover, the four-factor model fit the data better than a three-factor model in which the spatial and EF measures were made to load on the same construct. This suggest that spatial and EF skills – as measured in the current study – represent distinct constructs.

Taken together, our results suggest that VSWM is better defined as a measure of EF than spatial ability. This finding has important implications as VSWM and spatial visualization skills appear to represent different constructs and, as further discussed below, share different relations with measures of mathematics achievement. Follow-up research is needed to further test the extent to which differences in constructs are potentially due to the amount that the respective tasks emphasize the need to 'recall' visual-spatial information as opposed to self-generate and manipulate visual-spatial information. Do these differences in recall- versus generative-based tasks represent shared or distinct underlying cognitive mechanisms? Moreover, assuming VSWM and spatial visualization do represent distinct mechanisms, and our data suggest that they might, how do individual differences in these areas relate to mathematics achievement? Interestingly, while our data point to spatial visualization (i.e., generative spatial reasoning) as a more important contributor to mathematics achievement, it is possible that VSWM (i.e., recall-based spatial reasoning) may play a more important role in mathematics tasks that emphasize fluency, such as the retrieval of arithmetic facts. Answering questions such as these will contribute to a more nuanced understanding of when and how spatial and mathematical thinking interact.

In summary, the results of the CFA provided support for the hypothesized four-factor model in which performance on numerical, spatial, EF, and mathematics tasks emerged as separate but highly related factors. Although we retained this model for all subsequent pathway analyses, some caution is warranted as our results also suggest strong cross-loadings between factors. Future research is need to replicate the current findings and to further test the extent to which each factor is indeed independent from the other. Furthermore, research is needed that seeks to better explain the underlying mechanisms that give rise to similarities and differences across constructs.

8.2. Predictors of mathematics achievement

Our results indicated that only the numerical and spatial factors explained unique variance in children's mathematics achievement. Children's scores on the EF factor failed to explain performance in mathematics once the other two factors and age were taken into account. Spatial ability was an especially strong predictor of children's mathematics achievement.

Given that the majority of the mathematics test items required the understanding and/or manipulation of symbolic mathematics (e.g., 34 = 30 + _), it is somewhat surprising that the spatial factor, and not the numerical factor, was the best predictor of mathematics achievement. Critically, the relations between spatial ability and mathematics could not be explained by the inclusion of matrix reasoning as an indicator of spatial ability. The same pattern of results was obtained when matrix reasoning was eliminated from, as well as controlled for, in the analyses. These findings provide support for specific relations between spatial visualization and mathematics.

One explanation for this finding, and one not unique to our original hypothesis, has to do with the role of spatial visualization in mathematical problem solving. Indeed, it has been hypothesized that spatial visualization plays a critical role in how one mentally organizes, models, and ultimately makes sense of novel mathematical problems (Ackerman, 1988; Mix et al., 2016; Uttal & Cohen, 2012). Accordingly, the 'spatial modelling hypothesis,' as we have come to refer to it, predicts especially strong relations between spatial visualization and performance on novel mathematical tasks compared to highly familiar tasks. For example, spatial visualization would be expected to play a more important role when one is first learning arithmetic compared to when one has mastered their arithmetic facts. Interestingly, recent findings of Mix et al. (2016) provide support for the spatial modelling hypothesis, in which it was found that spatial skills were more related to novel mathematics problems than familiar ones. In the current study, the mathematics tests predominantly featured applied problems, lending further support for the role of spatial visualization in solving novel problems. This hypothesis dovetails nicely with the metaphor of spatial visualization as a cognitive tool used to construct spatial-numerical/mathematical relations.

Interestingly, our findings also provide evidence of relations between basic numerical skills and spatial visualization (for similar findings see Thompson et al., 2013; Viarouge et al., 2014). More specifically, our results indicate considerable overlap at the latent variable level ($r = 0.84$) as well as evidence that basic numerical skills partially mediate relations between spatial visualization and overall mathematics achievement. These results suggest that spatial visualization might also be involved in processing familiar and well-learned mathematical content, such as making rapid judgments about numerical symbols. Thus, our results implicate spatial visualization in numerical tasks that are solved both quickly and with seemingly little effort as well as tasks that require deliberate and effortful reasoning.

Although these results appear to run counter to the spatial modeling hypothesis (i.e., spatial visualization plays a greater role in novel mathematics), one possibility is that the association between basic number skills and spatial visualization is an artefact of numerical-spatial relations formed earlier in development. Spatial visualization may play an important role in early number learning as children actively construct spatial-numerical associations. Eventually, over development, these early conceptual groundings become increasingly more automatic and give rise to procedural fluency. Findings from our mediational analysis offer preliminary – albeit far from causal – support for this possibility and suggest that spatial visualization skills may facilitate numerical development. However, the relation between spatial visualization and mathematics achievement appears to be much stronger than the one shared between spatial visualization and basic numerical skills. Thus, the relation between spatial visualization and numerical skills cannot explain the robust relationship between spatial visualization and mathematics achievement more broadly and lends support to the spatial modelling hypothesis. This suggest that although numbers and spatial processes are linked at a relatively basic level, the association is even stronger at higher levels of numerical and mathematical processing.

Taken together, our findings suggest that spatial visualization skills play an important role in both basic numerical skills as well as more advanced numerical and mathematical reasoning. However, there appears to be an asymmetry in these relations, as spatial visualization was found to be more strongly related to novel or much less practiced mathematical tasks compared to tasks assessing numerical fluency. Future research efforts are needed to further disentangle when, why, and how spatial visualization is implicated in both basic and advanced mathematical reasoning.

8.3. Effects of age and grade on observed relations

Our findings suggest that the relations between numerical, spatial, and EF skills, and mathematical achievement develop in parallel and maintain relatively stable relations during early childhood (4-to-11 years of age). On the one hand, these findings are to be expected based on prior research showing strong and consistent relations between these variables in isolated studies of both children and adults (e.g., see Miyake et al., 2001; Mix & Cheng, 2012). On the other hand, these findings run counter to the idea that certain cognitive skills, such as spatial visualization, share stronger relations during initial learning of academic content, such as symbolic number, as compared to when the content has become more procedural and automatic (e.g., see Holmes & Adams, 2006; Huttenlocher, Jordan, & Levine, 1994; Mix et al., 2016; Rasmussen & Bisanz, 2005). For example, prior research has demonstrated that the learning of new mathematical content relies more on VSWM and less on verbal working memory (Rasmussen & Bisanz, 2005). However, with learning and development, the role of verbal working memory becomes increasingly more important for representing the learned material and the role of VSWM appears to become less important (Huttenlocher et al., 1994; Rasmussen & Bisanz, 2005). Indeed, this 'spatial' to 'verbal' shift is thought to correspond to changes in how the content is conceptualized; that is, as information that is initially grounded and understood in terms of concrete, spatial, and embodied experiences, but over time and experience, becomes increasingly more abstract and verbal in its representation (Bruner, 1966; Lakoff & Núñez, 2000). Notably, this shift also corresponds to a decrease in the need to exhibit effortful top-down executive control, suggesting that the role of EFs is dampened with mastery of content in a given area. Paradoxically, when one is confronted with the learning of new mathematics material, the role of EFs, most notably inhibitory control, is needed to inhibit prior learning experiences (e.g., overcoming the 'whole number bias' when introduced to fractions; $2/3 > 4/7$; Gómez, Jiménez, Bobadilla, Reyes, & Dartnell, 2015).

Taken together, the research above helps to shed light on why we may have observed consistent relations between spatial ability, EF skills, and mathematics achievement across such a wide age range of children. So long as the mathematics tasks are adaptive and requires the use of spatial skills and EFs to make sense of new or rarely encountered problems, relatively stable correlations between constructs are predicted to emerge. According to this view, successful performance on novel or difficult mathematical problems requires both independent and integrated contributions from both spatial and EF skills. Interestingly, our data offer only partial support for this hypothesis as only spatial skills were found to uniquely predict mathematics performance. Future research efforts are needed to further investigate this hypothesis using a different suite of EF measures. It is possible that the EF measures enlisted in the current study were too closely related to the spatial ability measures and any remaining variance was simply not enough to detect individual contributions of EF skills to mathematics achievement.

Based on the hypothesis stated above, we should expect to see tighter relations between spatial, EF, and numerical skills earlier in development and a gradual divergence of relations between spatial and EF skills and their relations with basic numerical skills over development. As symbolic number skills become more automatic the roles of higher cognitive skills should be minimized. However, this does not mean that the correlations between these various skills should necessarily become minimized. On the contrary, and to use the relation between EFs and mathematics as an example, if a child enters the learning of new mathematics material with strong EF skills, he/she should be able to harness these skills to better learn the new task(s). There is little reason to suspect that the relations would weaken over time, despite fundamental changes in the recruitment and reliance on EFs as the learner progresses from novice to 'expert.' This same explanation might underlie the relative stable relations in the current study between spatial visualization and basic numerical skills as well as more sophisticated mathematics tasks.

Longitudinal research is needed to further test the stability of the factors at the individual level. This approach will provide further insight into the relative stability and change that occurs in performance over development. For example, is it the case that children who start low on any given factor are likely to remain low throughout development? Moreover, how are improvements on any one factor associated with improvements across the other factors, perhaps most notably, mathematics achievement? In short, longitudinal research provides a means to better understand directional relations between the various factors. This information, in turn, has implications for educational design and intervention.

8.4. Mediating roles of EFs and numerical skills

Our results indicated that basic numerical skills, but not EFs, partially mediated the relation between spatial visualization and mathematics achievement. As noted above, the finding that numerical skills mediated the relation is in line with prior theoretical and empirical support that spatial skills facilitate numerical development (e.g., see Gunderson et al., 2012; Tam et al., 2018).

Our failure to reveal a mediating role of EF skills in the space-math link is a novel and surprising finding. Recall that the decision to test EF as a mediator in the relation between space and maths was based on the proposal that EF skills may be driving the space-math link due to the shared recruitment and reliance on top-down effortful control mechanisms. However, our findings indicate that although the two constructs were highly related, they were found to differentially relate to mathematics achievement. Spatial visualization appears to share a more direct link to our measures of mathematics than children's EF skills. This finding supports the longitudinal findings of Verdine, Irwin, Golinkoff, and Hirsh-Pasek (2014), who found that children's spatial skills at the age of three uniquely predicted children's mathematics performance one year later, explaining an additional 27% of the variance over and above children's EF skills.

However, it also worth considering an alternative explanation for why EF skills were not uniquely related to mathematics achievement and similarly failed to mediate the space-math link. Rather than assuming that spatial visualization and EFs represent distinct constructs, as we have done, it is possible that the spatial measures enlisted may in fact better represent indicators of EF than the measures enlisted to represent EF. Our attempt to separate spatial ability from EFs based on distinctions, in part, between the need to 'generate' versus 'recall' information may have resulted in a misrepresentation of EF. For example, it could be argued that the best measures of EF used in the current study were those used to measure spatial visualization, as these measures required a greater degree of manipulation of information in the service of a task. Future studies are needed to further investigate this possibility. It is possible that EF tasks that require greater amounts of planning and manipulation, as opposed to more recall-based tasks, would potentially result in stronger relations with both spatial visualization tasks but also mathematics achievement. In order to further test our claims made about the 'spatial modeling hypothesis' this is a critical next step: Is it the ability to generate and model *visual-spatial* solutions to problems that is most important to mathematical problem solving? Or is it a more general ability to generate solutions to problems, including verbally mediated processes, that matters most? As it stands, our data suggest that mathematics performance is best explained by an underlying construct related to the ability to generate and reason about visual-spatial images compared to a construct related to visual-spatial recall and inhibitory control.

8.5. Predicting numeration vs. geometry performance

As just alluded to, mathematics is not a unitary construct but rather a varied and complex one. For this reason, it has been suggested that any attempt to predict mathematical behavior should first consider the task requirements of the particular mathematics in question (see Mix & Cheng, 2012). In the present study, we used separate tests of numeration and geometry as examples of outcome measures that were expected to call upon different cognitive resources. More specifically, we predicted that numeration would best be predicted by basic number skills and geometry would best be predicted by spatial ability. These predictions were only partially supported. Although basic numerical skills did predict performance on the numeration test, spatial ability was found to be an even stronger predictor. Only spatial ability was a unique predictor of geometry.

Why might spatial ability better explain performance in both these areas of mathematics? One possibility is that basic numerical skills are necessary but not sufficient in order to perform well on both tests of mathematics. To do well requires not only a familiarity and fluency with numbers, but perhaps more importantly, knowledge and skills in the *use* and *application of* numbers within broader mathematical contexts. As hypothesized above, spatial visualization skills might serve as an important cognitive tool in this regard. To further illustrate this point, we return to the building metaphor in which numbers might be seen as the building blocks and spatial visualization as a tool used to manipulate and assemble the building blocks. Our findings suggest that key differences in mathematics performance are explained by both one's fluency with basic numerical skills but also – and perhaps to a greater extent – one's ability to operate on, use, and apply numbers within and across various mathematical problems. Ultimately, mathematics performance likely rests on one's ability to coordinate multiple representations and uses of number and various other mathematical symbols. Future research efforts are needed to better understand how different cognitive skills not only differentially relate to different branches of mathematics but also potentially different numerical and mathematical concepts, procedures, and facts within each branch.

8.6. Limitations

There are several limitations worth pointing out. First, this study was carried out in low SES populations, living in mostly rural areas. For this reason, one must be careful not to generalize the current findings to the general population. It is possible that children of lower SES backgrounds may rely more heavily on informal approaches to mathematics problem solving compared to their higher SES peers who may rely more heavily on formal learning experiences (e.g., see Jordan, Huttenlocher, & Levine, 1994). Accordingly, children in higher SES populations may rely less on spatial visualization skills and more on symbolic numerical skills (e.g., see Butterworth, Reeve, & Reynolds, 2011). However, recent evidence challenges this prediction. Reeve, Reynolds, Paul, and Butterworth (2018) demonstrated that the predictors of arithmetic in Indigenous and non-Indigenous children in rural and urban Australia were highly comparable and driven by similar visual-spatial factors. This finding highlights the importance of visual-spatial abilities for early numerical cognition regardless of SES and cultural divides. Given the mixed results to date, future work of this sort should strive to use a more economically diverse sample and seek to better understand the potentially moderating effects of SES on the observed relations.

Second, another concern with the current study has to do with the issue of common method variance (Kline, 2015); when variables are measured in highly similar ways. One reason we may have found separate factors for each construct might be partially explained by the common measurement approaches used to test each construct. For example, the numerical measures were all timed tests and involved highly similar task demands (e.g., crossing out the correct response). The spatial and mathematical tasks were all untimed and involved pointing to the correct response. Consequently, the spatial and mathematics measures may have been more closely related because individuals who were careful, took their time, and double-checked their work in the spatial measures may have also been more likely to do so in the mathematics measures. Issues of common measurement variance should be carefully considered in any future efforts to replicate the current findings.

Third, our results should be interpreted with acknowledging that all four constructs were highly correlated. Although a four-factor model was found to best fit the data, other models also fit the data to a satisfactory degree as well. So, although we can be confident that in combination, numerical, spatial, and EF skills provide a robust model of mathematics achievement, we are less confident of the more specific relations observed. The current model of mathematics, including the individual pathways, needs to be replicated.

Finally, our data were cross-sectional and limit any conclusions we can make about the directionality of the mediation analyses. Future research is needed to test longitudinal relations between numerical, spatial, and EF skills and their relations with mathematics achievement. Given their high correlations with one another, it seems germane to study the extent to which these variables interact with one another over time and potentially develop in part due to synergistic effects of one domain on the other. Said differently, does growth or improvement in one domain predict growth in the other domains? Intervention studies that target each construction in isolation but also in combination with one another will be critical in order to arrive at a better understanding of causal pathways between variables. Moreover, these efforts have the potential to eventually inform educational practice.

9. Conclusion

Results of a CFA demonstrated that numerical, spatial, EF, and mathematics skills are highly related, yet separable, constructs. Follow-up structural analyses revealed that numerical, spatial, and EF latent variables explained 84% of children's mathematics achievement scores while controlling for age. These results further highlight the potential importance of numerical, spatial, and EF skills in the learning and performance of foundational mathematics competencies, such as numeration and geometry. Further analyses revealed spatial reasoning as a particularly strong contributor to mathematics achievement. It is hypothesized that this relation rests on the critical role that spatial visualization plays in forming the problem and potential solutions to novel mathematics tasks. This study contributes to the growing need to further understand the dynamic interplay of basic cognitive skills and performance in various branches of mathematics.

Acknowledgement

We are grateful for the funding and research opportunities made possible through the following organizations. The Dr. Eric Jackman Institute of Child Study at the University of Toronto, Ontario Institute for Studies in Education. The Robertson Program for Inquiry-Based Teaching in Mathematics and Science at the University of Toronto, Ontario Institute for Studies in Education. Natural Sciences and Engineering Council of Canada (NSERC), the Canadian Institutes of Health Research (CIHR), the Canada Research Chairs Program, an E.W.R Steacie Memorial Fellowship from the Natural Sciences and Engineering Council of Canada (NSERC) to DA. Social Sciences and Humanities Research Council of Canada (SSHRC) Doctoral Fellowship to ZH.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cogpsych.2018.12.002.

References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117*, 288–318. https://doi.org/10.1037/0096-3445.117.3.288.

Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences, 33*(4), 245–266.

Barron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182.

Bruner, Jerome (1966). *Toward a theory of instruction.* Cambridge, MA: Belknap Press.

Butterworth, B., Reeve, R., & Reynolds, F. (2011). Using mental representations of space when words are unavailable: Studies of enumeration and arithmetic in indigenous Australia. *Journal of Cross-Cultural Psychology, 42*(4), 630–638.

Casey, B. M., Lombardi, C. M., Thomson, D., Nguyen, H. N., Paz, M., Theriault, C. A., & Dearing, E. (2018). Maternal support of children's early numerical concept learning predicts preschool and first-grade math achievement. *Child Development, 89*(1), 156–173.

Connolly, A. J. (2007). *KeyMath diagnostic assessment* (3rd ed.). Minneapolis, MN: Pearson.

Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition, 162*, 12–26.

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education, 3*(2), 63–68.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics.* New York, NY: Oxford University Press.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*(3), 371–396.
Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56*(2), 384–398.
de Hevia, M. D., Vallar, G., & Girelli, L. (2008). Visualizing numbers in the mind's eye: The role of visuo-spatial processes in numerical abilities. *Neuroscience & Biobehavioral Reviews, 32*(8), 1361–1372.
Delgado, A. R., & Prieto, G. (2004). Cognitive mediators and sex-related differences in mathematics. *Intelligence, 32*(1), 25–32.
De Smedt, B., Noël, M. P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education, 2*(2), 48–55.
Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... Sexton, H. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428–1446.
Ehrlich, S. B., Levine, S. C., & Goldin-Meadow, S. (2006). The importance of gesture in children's spatial reasoning. *Developmental Psychology, 42*(6), 1259–1268.
Fiedler, K., Harris, C., & Schott, M. (2018). Unwarranted inferences from statistical mediation tests – An analysis of articles published in 2015. *Journal of Experimental Social Psychology, 75*, 95–102.
Field, A. (2009). *Discovering statistics using SPSS.* London, UK: Sage Publications.
Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex, 86*, 186–204.
Friso-van den Bos, I., van der Ven, S. H., Kroesbergen, E. H., & Van Luit, J. E. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review, 10*, 29–44.
Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. D. (2010). The contributions of numerosity and domain-general abilities to school readiness. *Child Development, 81*(5), 1520–1533.
Galton, F. (1880). Visualised numerals. *Nature, 21*(533), 252–256.
Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37*(1), 4–15.
Göbel, S. M., Shaki, S., & Fischer, M. H. (2011). The cultural number line: A review of cultural and linguistic influences on the development of number processing. *Journal of Cross-Cultural Psychology, 42*(4), 543–565.
Gebuis, T., & Reynvoet, B. (2012). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General, 141*(4), 642–648.
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.
Hawes, Z., LeFevre, J. A., Xu, C., & Bruce, C. D. (2015). Mental rotation with tangible three-dimensional objects: A new measure sensitive to developmental differences in 4- to 8- year-old children. *Mind, Brain, and Education, 9*(1), 10–18.
Hawes, Z., Moss, J., Caswell, B., Naqvi, S., & MacKinnon, S. (2017). Enhancing children's spatial and numerical skills through a dynamic spatial approach to early geometry instruction: Effects of a 32-week intervention. *Cognition and Instruction, 35*(3), 236–264.
Hegarty, M., & Kozhevnikov, M. (1999). Types of visual–spatial representations and mathematical problem solving. *Journal of Educational Psychology, 91*(4), 684–689.
Holmes, J., & Adams, J. W. (2006). Working memory and children’s mathematical skills: Implications for mathematical development and mathematics curricula. *Educational Psychology, 26*(3), 339–366.
Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience, 6*(6), 435–448.
Huttenlocher, J., Jordan, N. C., & Levine, S. C. (1994). A mental model for early arithmetic. *Journal of Experimental Psychology: General, 123*(3), 284–296.
Jordan, N. C., Huttenlocher, J., & Levine, S. C. (1994). Assessing early arithmetic abilities: Effects of verbal and nonverbal response types on the calculation performance of middle-and low-income children. *Learning and Individual Differences, 6*(4), 413–432.
Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review, 5*(5), 602–619.
Kline, R. B. (2015). *Principles and practice of structural equation modeling.* New York, NY: Guilford Publications.
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), 77–94.
Kunda, M., McGreggor, K., & Goel, A. (2010). Taking a look (literally!) at the Raven’s intelligence test: Two visual solution strategies. *Proceedings of the Cognitive Science Society, 32*(32), 1690–1696.
Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being.* New York, NY: Basic Books.
LeFevre, J. A., Jimenez Lira, C., Sowinski, C., Cankaya, O., Kamawar, D., & Skwarchuk, S. L. (2013). Charting the role of the number line in mathematical development. *Frontiers in Psychology, 4*, 641. https://doi.org/10.3389/fpsyg.2013.00641.
Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology, 35*(4), 940–949.
Lohman, D. F. (1996). Spatial ability and G. In I. Dennis, & P. Tapsfield (Eds.). *Human abilities: Their nature and assessment* (pp. 97–116). Hillsdale, NJ: Lawrence Erlbaum.
Lynn, R., Backhoff, E., & Contreras-Niño, L. A. (2004). Sex differences on g, reasoning and visualisation tested by the progressive matrices among 7–10 year olds: Some normative data for Mexico. *Personality and Individual Differences, 36*(4), 779–787.
Lyons, I. M., Bugden, S., Zheng, S., De Jesus, S., & Ansari, D. (2018). Symbolic number skills predict growth in nonsymbolic number skills in kindergarteners. *Developmental Psychology, 54*(3), 440–457.
Lyons, I. M., Hutchison, J. E., Bugden, S., Goffin, C., & Ansari, D. (2018). *Kindergarteners reliability mis-classify ordered sequences of non-adjacent numbers. Paper presented at the Mathematical Cognition and Learning Society conference*. UK: Oxford University.
Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1–6. *Developmental Science, 17*(5), 714–726.
Lyons, I. M., Vogel, S. E., & Ansari, D. (2016). On the ordinality of numbers: A review of neural and behavioral studies. *Progress in Brain Research, 227*, 187–221.
MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology. 58*, 593–614.
Marghetis, T., Núñez, R., & Bergen, B. K. (2014). Doing arithmetic by hand: Hand movements during exact arithmetic reveal systematic, dynamic spatial processing. *The Quarterly Journal of Experimental Psychology, 67*(8), 1579–1596.
McClelland, M. M., & Cameron, C. E. (2012). Self-regulation in early childhood: Improving conceptual clarity and developing ecologically valid measures. *Child Development Perspectives, 6*(2), 136–142.
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.
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.
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.
Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). *Quantitative development in infancy and early childhood.* New York, NY: Oxford University Press.
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, 18*(4), 465–492.
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.
Muthén, L. K., & Muthén, B. O. (1998–2015). *Mplus (Version 7.4).* Los Angeles, CA: Muthén & Muthén.
Newcombe, N. S., & Huttenlocher, J. (2003). *Making space: The development of spatial representation and reasoning.* Cambridge, MA, US: MIT Press.
Newcombe, N. S., Levine, S. C., & Mix, K. S. (2015). Thinking about quantity: The intertwined development of spatial and numerical cognition. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(6), 491–505.
O’brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41*(5), 673–690.
Parsons, S., & Bynner, J. (2005). *Does numeracy matter more?* London, UK: National Research and Development Centre for Adult Literacy and Numeracy, Institute of

Education.
Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology, 45*(3), 605.
Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of Experimental Child Psychology, 91*(2), 137–157.
Raven, J. (2008). *Coloured progressive matrices and Chrichton vocabulary scale.* London, UK: Pearson.
Reeve, R. A., Reynolds, F., Paul, J., & Butterworth, B. L. (2018). Culture-independent prerequisites for early arithmetic. *Psychological Science, 29*(9), 1383–1392.
Reuhkala, M. (2001). Mathematical skills in ninth-graders: Relationship with visuo-spatial abilities and working memory. *Educational Psychology, 21*(4), 387–399.
Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological Science, 24*(7), 1301–1308.
Ritchie, S. J., Bates, T. C., & Deary, I. J. (2015). Is education associated with improvements in general cognitive ability, or in specific skills? *Developmental Psychology, 51*(5), 573–582.
Sella, F., Berteletti, I., Lucangeli, D., & Zorzi, M. (2017). Preschool children use space, rather than counting, to infer the numerical magnitude of digits: Evidence for a spatial mapping principle. *Cognition, 158*, 56–67.
Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science, 20*(3), https://doi.org/10.1111/desc.12372.
Siegler, R. S. (2016). Magnitude knowledge: The common core of numerical development. *Developmental Science, 19*(3), 341–361.
Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*(2), 428–444.
Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science, 11*(5), 655–661.
Soper, D. S. (2018). A-priori sample size calculator for structural equation models [software]. Available from http://www.danielsoper.com/statcalc.
Tam, Y. P., Wong, T. T. Y., & Chan, W. W. L. (2018). The relation between spatial skills and mathematical abilities: The mediating role of mental number line representation. *Contemporary Educational Psychology, 56*, 14–24.
Thompson, J. M., Nuerk, H. C., Moeller, K., & Kadosh, R. C. (2013). The link between mental rotation ability and basic numerical representations. *Acta Psychologica, 144*(2), 324–331.
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, 29*(2), 239–266.
Toomarian, E. Y., & Hubbard, E. M. (2018). On the genesis of spatial-numerical associations: Evolutionary and cultural factors co-construct the mental number line. *Neuroscience & Biobehavioral Reviews.*.
Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why and how. *Psychology of Learning and Motivation, 57*, 147–181. https://doi.org/10.1016/B978-0-12-394293-7.00004-2.
Verdine, B. N., Irwin, C. M., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Contributions of executive function and spatial skills to preschool mathematics achievement. *Journal of Experimental Child Psychology, 126*, 37–51.
Viarouge, A., Hubbard, E. M., & McCandliss, B. D. (2014). The cognitive mechanisms of the SNARC effect: An individual differences approach. *PloS ONE, 9*(4), e95756.
Wei, W., Yuan, H., Chen, C., & Zhou, X. (2012). Cognitive correlates of performance in advanced mathematics. *British Journal of Educational Psychology, 82*(1), 157–181.
Weston, R., & Gore, P. A., Jr. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist, 34*(5), 719–751.
Zelazo, P. D., Carlson, S. M., & Kesek, A. (2008). The development of executive function in childhood. In C. A. Nelson, & M. Luciana (Eds.). *Developmental cognitive neuroscience. Handbook of developmental cognitive neuroscience* (pp. 553–574). Cambridge, MA, US: MIT Press.